ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Alternative Splicing In Fibronectin: Regulation And Basic Mechanism

Thesis submitted for the Degree of

Doctor Philosophiae

Candidate: Massimo Caputi

Supervisor: Prof. Francisco Ernesto Baralle

Academic Year 1994/1995

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

PREFACE

The work described in this thesis was carried out at the International Center for Genetic Engineering and Biotechnology (ICGEB), Trieste, between August 1991 and September 1995.

I would like to thank my supervisor, Prof. F. E. Baralle, under whose supervision the experimental work was carried out for his excellent guidance and support throughout this project. I am also indebted to Maurizio Romano and Andres Muro, for their help, ideas, discussion and friendship throughout my time in this laboratory and to Carlos Melo that encouraged and helped me in the early stages of this work. Thanks are also due to Carlo Gregori for photographic work.

A special thank goes to Franco, Giuliano, Riccardo, Roby, Sergio, Stefano and my brother Fabrizio for their support, patience and friendship in these years. Finally, I would like to thank my parents for their encouragement and aid throughout.

TABLE OF CONTENTS

Chapter 1

Introduction

In 1977 two pioneristic works (Berget et al., 1977; Chow et al., 1977) revolutionised our understanding of the biochemical pathway between a gene in the nucleus and its mRNA in the cytoplasm. The comparison of the adenovirus cytoplasmatic mRNA sequences with the viral DNA sequences from which it was transcribed resulted in the discovery of the split gene structure. Furthermore in the same year Jeffreys and Flavell found mammalian coding regions also interupted by large insertions. It is now well established that most eucaryotic genes which encode messenger RNA are split genes, that is, the protein coding sequences (exons) are interrupted by non-coding sequences (introns). Expression of these genes requires the accurate excision of introns from the primary transcript (pre-mRNA) and the ligation of exon RNA sequences, which occurs in a process termed pre-mRNA splicing. Interrupted genes are found in all types of eukaryotic cells studied. Although the proportion of intron-containing genes is generally lower in unicellular organisms than in metazoan.

Splicing occurs in the nucleus, as the other modifications that are made to newly synthesised RNAs. The substrate for these processes is the ribonucleoprotein particle (hnRNP), in which the RNA synthesised by the transcription machinery, the so called heterogeneous nuclear RNA (hnRNA), is bound by proteins. The transcript is capped at the 5′ end, has the introns removed and is polyadenilated at the 3′ end. The RNA is then transported through nuclear pores to the cytoplasm, where it is available to be translated.

The nuclear hnRNA processing is steel unclear in many points: does splicing occur at a precise location in the nucleus and is it connected with the other events?

Is the hnRNA processing used to regulate gene expression by discriminating among the available precursors? Are introns excised from a precursor in a particular order?

With regard to the splicing reaction itself one of the main questions is how its specificity is controlled. Up to now the complex mechanism that ensures a precise excision of the intronic sequences is still unknown. This point is particularly important in higher vertebrate were intronic sequences are particularly long and splicing undergoes a complex regulation. In this study I have tried to clarify this point using the fibronectin messenger as a model, analysing both its expression and the splicing mechanism regulating it in a tissue specific manner.

1.1 THE MECHANISM OF SPLICING

1.1.1 Different Introns Different Splicing Mechanisms

Depending from the presence of conserved nucleotide sequences and RNA structure introns have been classified in at least four groups. A different splicing mechanism has been discovered for each group (Fig. 1.1).

a) Group I introns are found in genes of fungal mitochondria and in rRNA of lower eukaryotes such as Tetrahymena and Physarium. These introns have a peculiar secondary structure and can be spliced autocatalytically in two steps in presence of guanosine nucleotide as cofactor (Fig. 1.1.A).

b) Group II introns are found in fungal mitochondrial genes. As group I introns the group II share a peculiar secondary structure and could be self-spliced in a two step reaction. They do not need guanosine as cofactor and their splicing intermediate is similar to the one of the nuclear mRNA introns (Fig. 1.1.B).

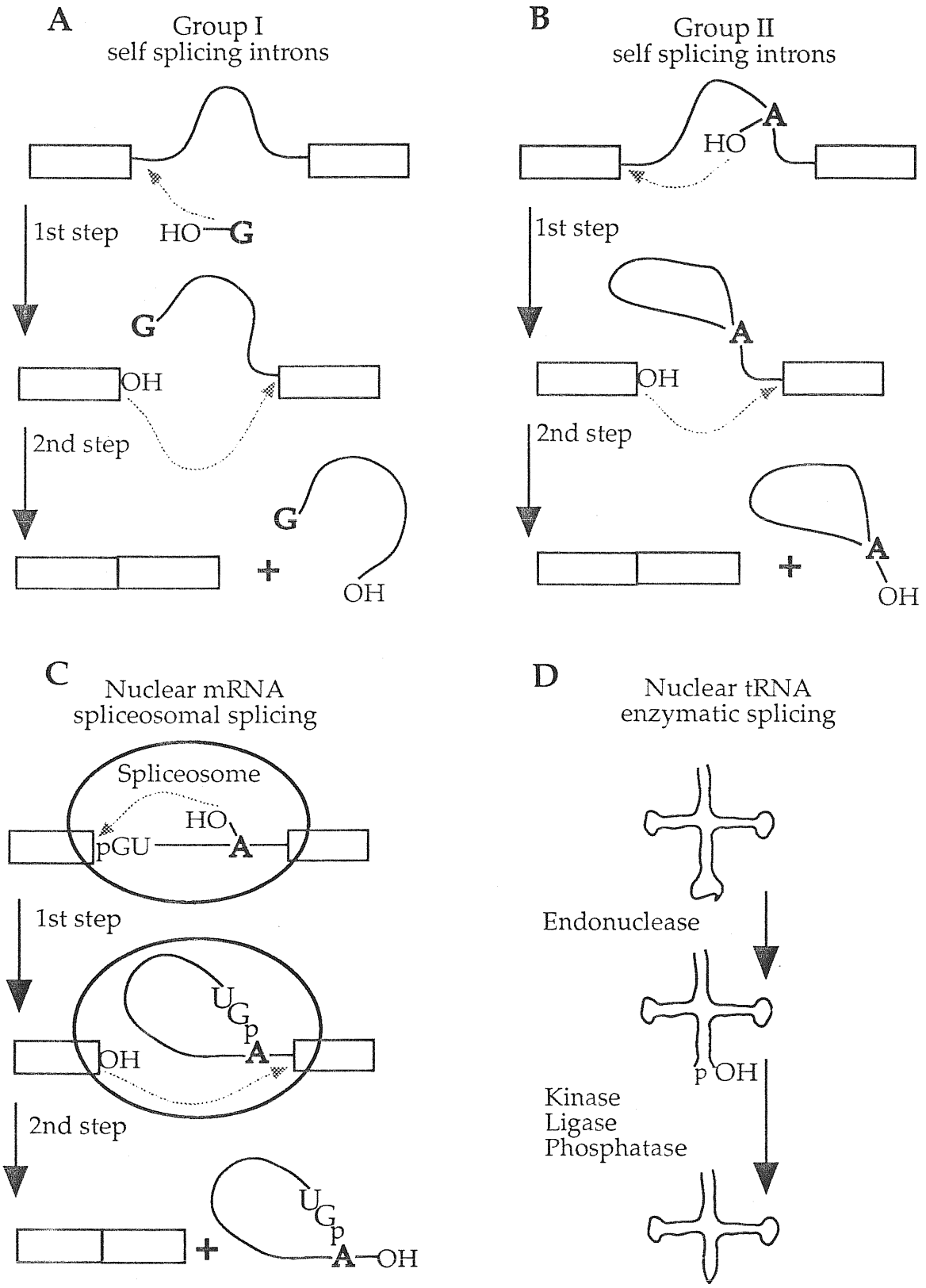


Figure 1.1. Splicing mechanism of the four major groups of introns.
The basic splicing mechanism is outlined for precursors containing two exons separated by an intron.

c) Nuclear mRNA in higher eukaryotes is processed by a system that recognise short consensus at the intron-exon junctions and within the intron. The reaction requires a large splicing apparatus constituted by proteins and ribonucleoproteins and undergoes to complex regulation. As autocatalytic introns nuclear mRNA is also spliced in a two steps reaction (Fig. 1.1.C).

d) Splicing of yeast tRNA precursors involves enzymatic activities that cleave and rejoin regions of the precursors (Fig. 1.1.D).

1.1.2 The Nuclear Pre-mRNA Splicing

1.1.2.1 Summary of the steps of nuclear mRNA splicing

The development of in vitro splicing systems (Padget et al., 1983) contributed enormously to our knowledge of the mechanistic and biochemistry of the splicing process, nevertheless there are still open questions concerning both the catalytic step itself and the regulative mechanism controlling the process.

Splicing of nuclear pre-mRNAs occurs in a two steps pathway (Fig. 1.1). The first step involves cleavage of the 5′ splicing site and the formation of a 2′-5′-phosphodiester bond between the 5′-terminal G residue of the intron and an A residue normally located between 18 and 40 nucleotides upstream from the 3′ splice site, namely the branch point. A free 5′ exon and a intron-3′ exon in a so called "lariat" structure are formed. The second step involves cleavage of the 3′ splice site, ligation of the two exons and release of the intron which is still in the lariat form. The highly conserved structural elements, typical of pre-RNA introns that are autocatalytically spliced (group I and II introns), appears to be compensated by the involvement of a large number of trans-acting factors in pre-mRNA splicing. The trans-acting factors can be divided in two classes. There are the snRNPs, evolutionary highly conserved nuclear RNA-protein complexes, and the so called non-snRNP splicing factors. The snRNPs assemble in an ordered fashion onto the pre-mRNA to form a complex, the spliceosome (reviewed in Lamond, 1993), were the two steps catalytic reaction takes place.

1.1.2.2 The splicing substrate

The substrate of the splicing reaction is the nuclear pre-mRNA that is known to be bound in a ribonucleoprotein complex termed heterogeneous nuclear ribonucleoprotein (hnRNP) particles (reviewed in Dreyfuss et al., 1993). An hnRNP particle takes the form of beads connected by a fibre. The beads are globular particles of diameter ~200Å, sediment at ~40s, and contain 100-800 bases of RNA associated with a set of proteins that range in size from 34,000 to 120,000 daltons. A hnRNP particle contains ~20 different proteins 6 of which (A1, A2, B1, B2, C1, C2) are called the core proteins. There are strong evidence for distinct RNA binding sequence specificity for many hnRNP proteins (Bennet et al., 1992), however up to now the exact role of the hnRNP in splicing remain to be determined.

Three short sequences (5′ and 3′ splice sites and the branch point region) in the pre-mRNA intron are essential, though not sufficient, for the intron to be accurately and efficiently spliced. In mammals the consensus sequences for the 5′ and 3′ splice sites are the following (Shaphiro and Senaphaty, 1987):

	3' splice site	EXON	5' splice site
	Y Y Y Y Y Y Y Y Y Y Y Y Y Y N C A G /	R.....N	A G / G T R A G T
	76 76 80 80 76 76 80 80 79 78 82 82 85 84 72 100 100	76	56 78 100 100 95 69 82 53

The Y indicate a pyrimidine, the R a purine and N any nucleotide. The numbers indicate the percentage of each base in the consensus.

There is considerable flexibility in the sequence requirements for the branch site in higher eukaryotes, the consensus sequence determined is the following: UNCUR<u>A</u>C (Nelson and Green 1989), where the "A" residue underlined is that at which branch formation occurs. Usually this site is locate 18-40 nucleotides upstream of the point of cleavage of the RNA. In yeast the sequence requirements for splicing are somewhat different to those in higher eukaryotic genes although the basic reaction mechanism is similar. The 5′ splice site sequence (G/GTATGT) and the branch site (TACTA<u>A</u>C) are strongly conserved. Few changes are tolerated in this sequence, mutation here can abolish splicing (Langford et al.,

1984), in contrast to mutations in mammalian branch sites which appear only to slow the rate of splicing. The acceptor site in comparison is less well conserved between genes although there is a consensus sequence (C/TAG).

1.1.2.3 The spliceosome

The two-step nuclear pre-mRNA splicing reaction is carried out within a multi-component complex called the spliceosome. The spliceosome is comprised of RNA-protein subunits, the U1, U2, U4/U6 and U5 small nuclear ribonucleoprotein particles (snRNPs). The U1, U2, U4 and U5 snRNAs are transcribed by RNA polymerase II, they rapidly migrate to the cytoplasm were they assemble with the core snRNP proteins. The snRNPs are reimported into the nucleus. U6 snRNA is transcribed by RNA polymerase III. U6 remains in the nucleus and assemble with the U4 snRNP to form a larger U4/U6 ribonucleoprotein complex (Luhrmann et al., 1990). U5 assemble in an ATP dependent reaction with the U4/U6 snRNP forming a [U4/U6.U5] three snRNP particle (Konarska and Sharp, 1987).

A group of of eight polypeptides (B, B′, D1, D2, D3, E, F, G) form a common core in each snRNP particle. Core proteins interact with a conserved sequence present in the U1, U2, U4 and U5 snRNA. In addition to the core proteins the U1 and U2 snRNPs also contain a group of specific proteins. The A, C and 70 kDa proteins are specific for the U1 snRNP. U1 70K and U1A appear to be involved in splice site recognition and selection (Flickinger and Salz, 1994; Kaltz et al., 1994). U1A is also involved in linking the splicing and polyadenilation machinery (Gunderson et al., 1994). A′ and B′ are specific for the U2 snRNP. Both proteins are involved in the U2 snRNA binding specificity (Scherly et al., 1990). No additional proteins have, as yet, been found in purified U4/U6 snRNPs in mammals. In yeast a specific U4/U6 snRNP associated protein, named PRP4, has been described (Banroques and Abelson, 1989). Nevertheless when U5 and U4/6 are associated several specific polypeptides have been detected in the [U4/U6.U5] three snRNP.

1.1.2.4 The commitment complex

Spliceosome assembly on the pre-mRNA substrate is a complicated multistep process which general features are common to mammals and yeast. While some steps of the assembly pathway are not defined yet a general picture has been pointed out in recent reviews (Moore et al., 1993; Sharp, 1994; Green, 1991) and is summarised in figure 1.2.

The first step is a formation of an ATP independent commitment complex (CC in figure 1.2) on the pre-mRNA substrate, this is defined as a complex that commit the pre-mRNA to the splicing pathway, it is, the pre-mRNA will be spliced preferentially in presence of an excess amount of competitor RNA. Commitment complex have been identified both in yeast (Seraphyn and Rosbash, 1989) and mammals were it has also been named E or early (Michaud and Reed, 1991). The commitment complex contain a stable bound U1 snRNP and, at least in some cases in mammals (Jamison and Garcia-Blanco, 1992), the U2 snRNP is also weakly associated to the complex. The binding of U1 involves the formation of Watson-Crick base pairing between the 5′ splice site region consensus sequence (CAGGUAAGU) and a complementary region at the 5′ end of the U1 snRNA.

1.1.2.5 U2 snRNP binding and A complex formation

The U2 snRNP bind the pre-mRNA at an early stage of spliceosome assembly, following the commitment complex formation, forming the stable complex A. U2 snRNP binding requires ATP hydrolysis and a base-pairing interaction between the branch point sequence and a complementary sequence in the U2 snRNA (Wu and Manley, 1989). This interaction is particularly important in yeast where there is an high conservation of the branch point sequence, it is less important in mammalian system where branch point sequences are less conserved. In mammalian systems the stable binding of U2 snRNP requires also a non snRNP protein: the U2 auxiliary factor (U2AF) (Ruskin et al., 1988). It is constituted by two subunits of 65 kDa and 35 kDa. The 65 kDa subunits bind tightly with the polypyrimidine tract. An additional set of polypeptides associated with the U2 snRNP are needed for A complex assembly and the further splicing steps, among those the better characterised are SF1 and SF3 (Brosi et al., 1993).

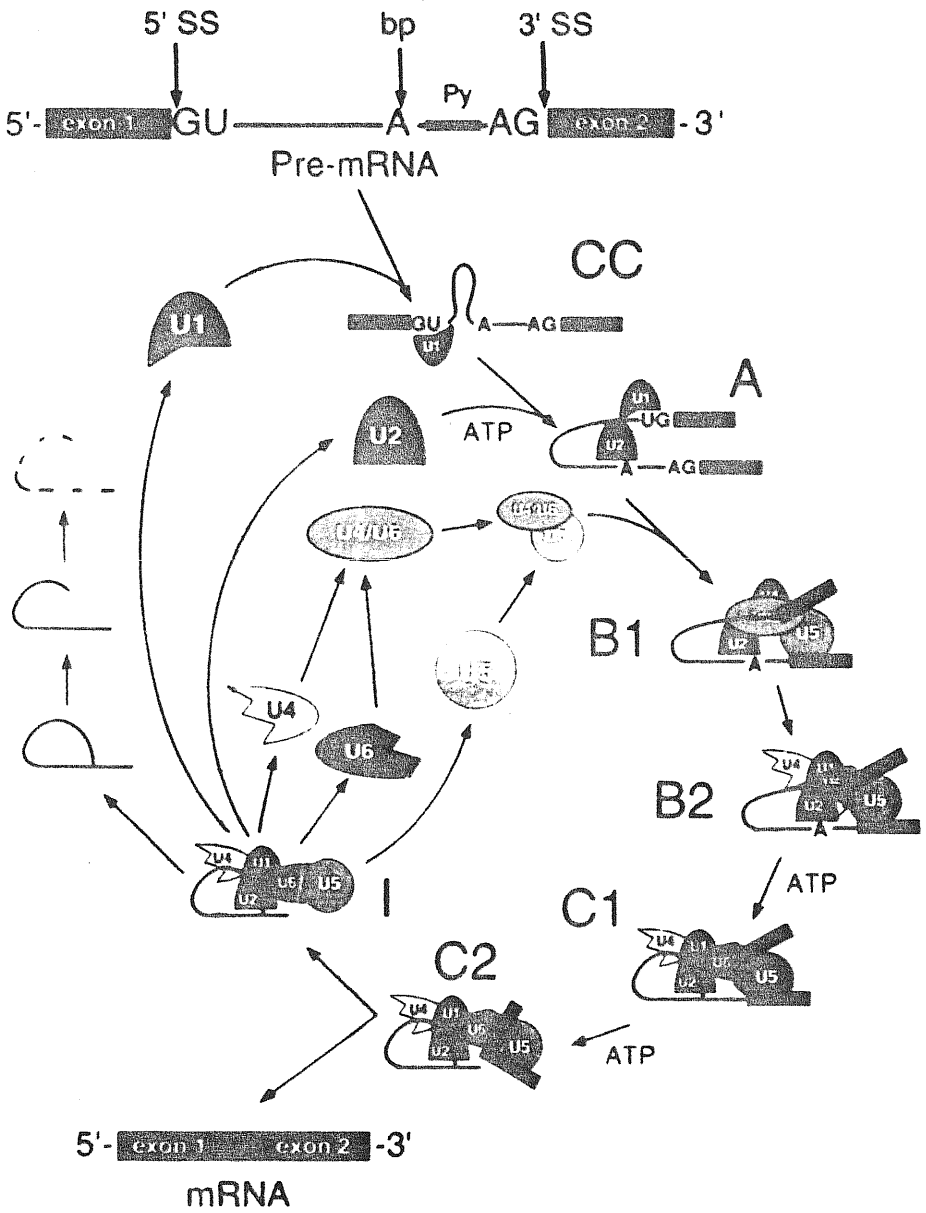


Figure 1.2. Schematic representation of the spliceosome cycle in nuclear pre-mRNA splicing.

A basic pre-mRNA substrate containing two exons separated by an intron (top figure) is processed in a two step reaction inside a large ribonucleoprotein complex (the spliceosome) to give a two exon mRNA and an excised lariat intron (bottom figure). The individual small nuclear ribonucleoprotein that enter the cycle are indicated (U1, U2, U4, U5 and U6). The different macromolecular complexes formed during the splicing pathway (A, B1, B2, C1, C2, I) are indicated as described in section 1.1.2.

(Sharp, 1994).

1.1.2.6 [U4/U6.U5] snRNP binding

U4/U6 and U5 enter the spliceosome as a trimer [U4/U6.U5] to form complex B (Lamm et al., 1992; Seraphin et al., 1991). At this stage a significant modification of the intermolecular base pairing snRNA:pre-mRNA and snRNA:snRNA takes place. U1 snRNA is unwound from the pre-mRNA and replaced by U5 snRNA that has the same RNA binding specificity (Newman and Norman, 1992) and after conformational changes gain a firmer grip on the 5′ splice site (Sontheimer and Steitz, 1993; Blencowe et al., 1989). The U4/U6 snRNAs base pairing is also weakened, lost its base pairing with U4 the U6 snRNA is now able to pair with a complementary region in the U2 snRNA (Madhani end Guthrie, 1992). This base pairing produces RNA structures similar of those of group II self splicing introns. Furthermore direct indications for a direct role of the U6 snRNA in the first catalytic step (Wolff and Bindereif, 1993) enforce the hypothesis that the splicing reaction of nuclear pre-mRNA is RNA catalysed.

1.1.2.7 The catalytic steps

After spliceosome reorganisation occur in the late B complex (B2) the first catalytic step take place. Cleavage of the 5′ splice site with the concomitant liberation of the free 5′ exon and production of the lariat intermediate define the formation of the C1 complex (as in Moore et al., 1993). The second step involves cleavage of the 3′ splice site and ligation of the adjacent exons. The transition from the first to the second catalytic step is accompanied by a second major conformational change in the spliceosome (Schwer and Guthrie, 1992; Sontheimer and Steitz, 1993) that lead to the C2 complex.

The ligated exons and the lariat-intron are then rapidly released in an active process involving proteins (Company et al., 1991). The lariat-intron is rapidly degraded while the spliced product will end its processing pathway and will be transferred to the cytoplasm throughout the nuclear pores. It is not well known the fate of the single snRNP when the spliceosome is disassembled but since they have a long half life it is likely that they reassemble in a new spliceosome.

1.1.2.8 Non-snRNP splicing factors

The spliceosome contains several non-snRNP protein splicing factors. only in few cases those factors have been purified to homogeneity and their cDNA have been isolated. Most of the factors identified acts in the early step of spliceosome formation and are associated to the 5′ splice site and to the polypyrimidine tract/3′ splice site. U2AF is one of the best characterised, it binds the polypyrimidine tract via its 65 kDa subunit. Several hnRNPs are somehow involved in pre-mRNA splicing but their role is not clear. A similar specificity is found in the polypyrimidine-tract binding protein (PTB), a 62 kDa polypeptide also called hnRNP I (Garcia-Blanco et al., 1989). The hnRNP C has polypyrimidine tract binding property as well but with distinct specificity of PTB (Bennet et al., 1992). The hnRNP A1 preferentially interacts with 5′ splice site and can play a role in alternative splicing events (see section 1.4) (Mayeda and Krayner, 1992). Another factor termed PSF also bind the polypyrimine tract (Patton et al., 1993), it is associated in the C complex with at least 20 others non-snRNP proteins and is essential for the second catalytic step. A possible function for PSF is to replace the 65 kDa U2AF subunit which is destabilized during spliceosome assembly (Gozani et al., 1994).

A family of arginine/serine rich proteins (SR proteins) involved in the early step of splicing has been recently characterised, SR protein characteristic and functions will be discussed in section 1.4.

1.1.3 Yeast Nuclear Pre-mRNA Splicing

Nuclear pre-mRNA splicing in yeast is a more rare event that in higher eukariotes. Usually introns are short with strong consensus sequences and up to now only few examples of regulated splicing have been discovered. Many aspects of the two step splicing mechanism have been elucidated by a genetic approach in yeast. The basic feature of the spliceosome formation are similar for both yeast and mammals but in yeast there is a clear formation of an early complex in which U1 snRNP is the only snRNP present. At least 30 genes essential for RNA processing

have been identified from screens of temperature sensitive mutants (reviewed in Ruby and Abelson, 1991), these genes have been named PRP (precursor RNA processing). A number of these PRP gene products appear to be either snRNP components or snRNP-associated factors (Fig. 1.3). Precise biochemical functions have not been assigned to all the PRP genes and only recently a number of mammalian PRP homologues have been cloned (Brosi et al., 1993). Some of them have RNA recognition motifs (RRMs) or zinc finger-motifs and a number share sequence homologies with the members of the so called "DEAD" or "DEAH" box ATP-dependent RNA helicases family (Schmid and Linder, 1992).

Surprisingly, none of the PRP genes identified so far encodes proteins that have the SR motif found in many mammalian splicing factors (see section 1.4). This absence could be explained with the lack of regulatory events in the yeast splicing, exception made for the meiotic specific splicing of the MER2 gene (Nandabalan et al., 1993) and for the ribosomal protein L32 (Vilardell and Warner, 1994). SR proteins seems, indeed, to be exential in the alternative splicing regulation in higher eukariots.

It has been estimated that in yeast at least 100 genes encodes for products important for the splicing apparatus, that is 1-2% of the total genome. It is likely that most of the PRP gene encodes for proteins performing multiple roles in the nuclear pre-RNA processing pathway with RNA binding, RNA helicase and protein:protein properties.

1.2 REGULATION OF SPLICE SITE SELECTION

In eukariotes an additive mechanism to regulate gene expression for many genes is the use of alternative splice sites (reviewed in McKeow, 1992). Regulated alternative splicing can lead to the production of different proteins from a single pre-mRNA or can function as an additive on-off switch for some genes in which one of the alternative forms does not encodes for a functional protein. Genes can be alternatively spliced in a temporal, tissue specific or developmental fashion.

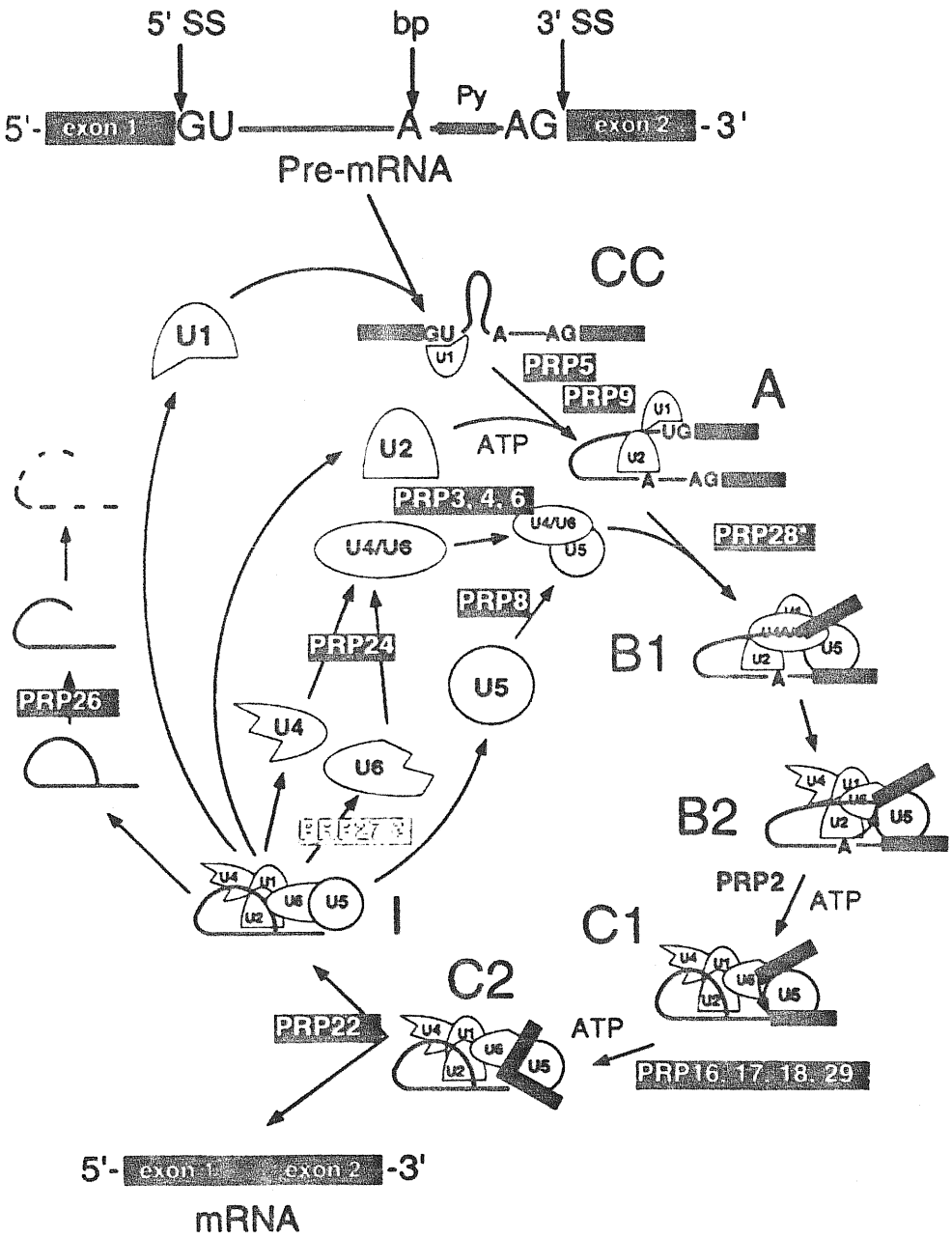


Figure 1.3. PRP proteins required in the yeast spliceosome cycle.
The PRP mutant required in each step of spliceosome assembly are indicated.
(Sharp, 1994).

There are numerous ways to alternatively splice a multi-intronic pre-mRNA (Fig. 1.4) that varies from a simple splice-don't splice choice to more complex pattern of alternative 5′ and 3′ choice, mutually exclusive exons and alternative 5′ or 3′ ends. Why an exon undergoes alternative splicing and how this regulation is achieved are still open questions. However in recent years a better understanding of the overall splicing mechanism and studies on a variety of alternative splicing models started to clarify these problems. Most of alternative splice sites do not have strong consensus 5′ or 3′ splice site sequences or are surrounded by a sub-optimal splicing environment (e.g. weak polypyrimidine or branch point signals or weak upstream and downstream splice sites). Due to their poor matching to the consensus several alternatively spliced exon should be, theoretically, not recognised by the splicing machinery however additive sequences in both, exons and introns, help their recognition. It seems likely that the splicing regulation is due to the activity of specific and general splicing factors only recently characterised.

1.2.1 Splice Sites Selection

To better understand the mechanism underlying the alternative splicing we have to know which are the structural and the sequence features required for pre-mRNA splicing.

1.2.1.1 5′ and 3′ splice sites requirements

Mutations of splice sites make a significant contribution to human genetic disease: approximately 15% of point mutations that cause genetic disease affect pre-mRNA splicing (Krawczak et al., 1992). Mutations are usually found in positions which are more than 80% conserved. In addition mutations causing aberrant splicing are more frequently found at the 5′ splice site than at the 3′ splice site. Exon skipping is the preferred in vivo phenotype (51%), although activation of cryptic splicing sites (32%), creation of a pseudo exon within an intron (11%) and intron retention (6%) are also found (Nakai and Sakamoto, 1994). In table 1.1

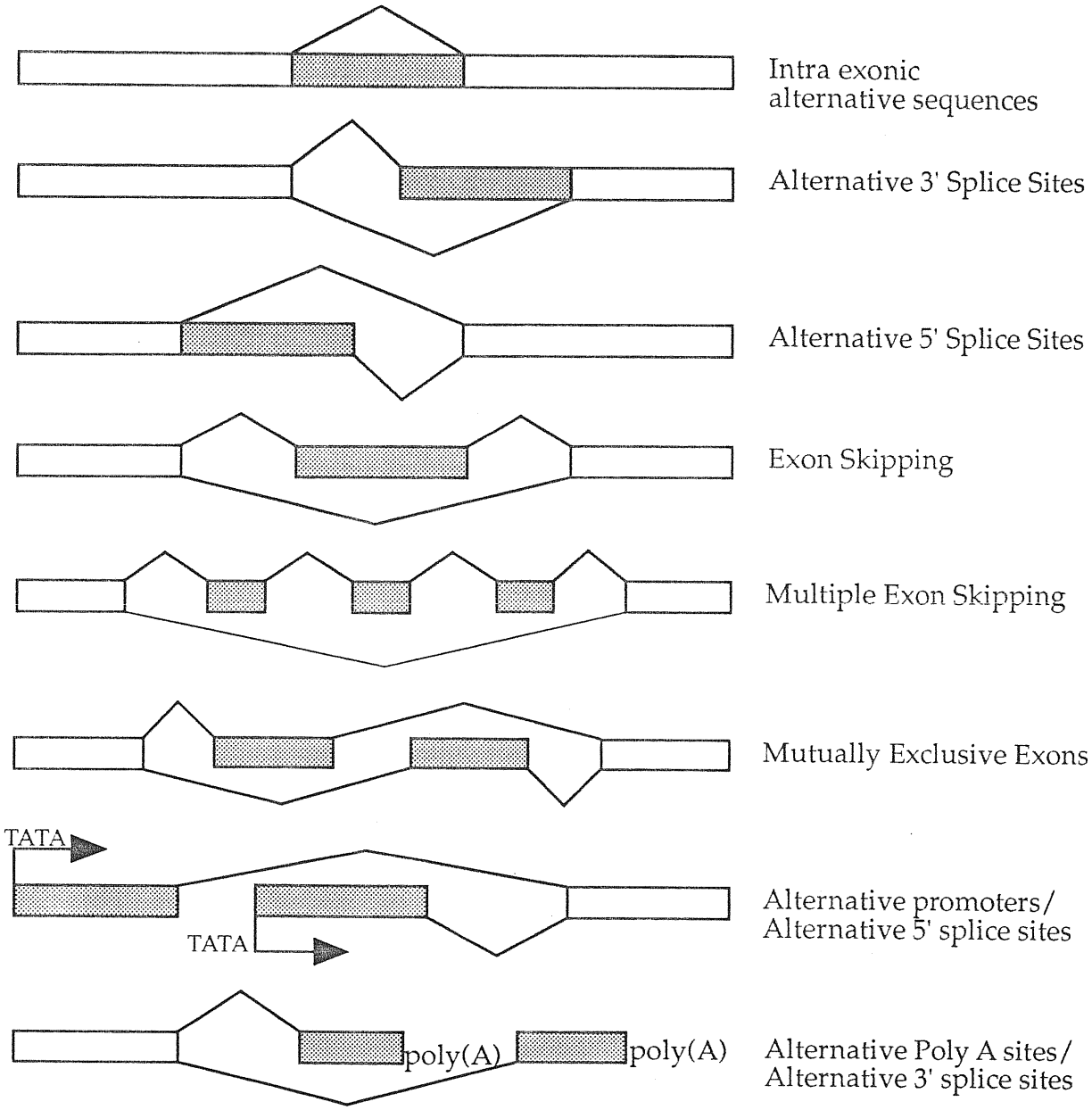


Figure 1.4. Various types of alternative splicing events.
Schematic representation of various alternative splicing events that occur in vivo. Open boxes represent constitutive exons, shadowed boxes represent alternatively spliced exons. The last two examples describes how splicing and other processes are connected. Alternative 5′ splice sites can give alternative 5′ ends and thus alternative initiation of translation. Alternative 3′ splice sites can give alternative polyadenilation sites.

some example of 5′ splice site mutations and relative phenotype are reported. Consensus sequences are limited at 4 bases at the 3′ intron/exon junction and for 7 bases at the 5′ exon/intron junction (Shapiro and Senapathy, 1987).

In alternative splicing models mutations of a weak splice sites to improve its consensus can led to constitutive recognition of alternatively skipped exon (Huh and Hynes, 1993; Del Gatto and Breathnach, 1995) or to the prevalent recognition of an alternative 5′ or 3′ splice site (Zhuang et al., 1989). The EIIIB alternatively skipped fibronectin exon has weak 5′ and 3′ splice site but mutations improving the 5′ splice site flanking the upstream intron also improve EIIIB inclusion (Huh and Hynes, 1993).

Splice site strength is thus an important determinant in splice site selection. However, the consensus sequences are not sufficient to account for the observed high specificity of splice site selection. Apparently strong sites are not always selected as splice sites, whereas some authentic sites seem to be weak (Brunak and Engelbrecht, 1991). Moreover, a synthetic splice site inserted into various regions of a pre-mRNA exhibited variable activity in a manner dependent on its relative location (Nelson and Green, 1988). These observations indicate that other elements are involved in the selection of splice sites.

1.2.1.2 The Branch point sequence and the polypyrimidine tract

A weakly conserved sequence element, the branch site, is required for lariat formation during the splicing reaction. In all genes examined except one (a branched C nucleotide in an intron of the human growth hormone gene (Hartmuth and Barta, 1988) the major site of lariat formation is at an A residue. The sequence is usually located within 40 nucleotides upstream the 3′ splice site. Although branch point strength is more loosely defined in pre-mRNA splicing of higher eukaryotes than it is in yeast its sequence composition (and its complementary to U2 snRNA) has been implicated in constitutive and alternative splicing as well (Fu et al., 1991; Noble et al., 1988; Zhuang et al., 1989). Recent evidence showed that up to 19 nucleotides upstream the branch point adenosine can be involved in proper functioning of the branch point itself (Dominski and

Gene	Mutation	Exon skipping	Cryptic use	Reference
Consensus	CAG/GURAGU			Shapiro et al (1987)
Adenine phosphoribosyltransferase (human)	Insertion GUU	Yes		Hidada et al (1987)
β-Globin (human)	.../A.....	Yes		Treisman et al (1982)
	.../....C.		Yes	Treisman et al (1982)
	.../A.....		Yes	Treisman et al (1982)
β-Globin (rabbit)	.../A.....		Yes	Wieringa et al (1983)
β-Hexosaminidase α chain (human)	.../C.....	Yes		Ohno et al (1988)
Chorionic somatomammotropin-like (human)	.../A.....		Yes	Chen et al (1989)
Dihydrofolate reductase (hamster)	.../C.....	Yes		Mitchell et al (1986)
Glycophorin B (human)	.../T.....	Yes		Kudo et al (1989)
γ2b heavy-chain immunoglobulin, human	Deletion	Yes		Brandt et al (1984)
μ heavy-chain immunoglobulin (human)	Deletion	Yes		Bakhshi et al (1986)
Hypoxanthine phosphoribosyltransferase (human)	.../....A	Yes		Gibbs et al (1990)
	.../A....	Yes		Steingrimsdottir et al (1992)
	.../....A.		Yes	Steingrimsdottir et al (1992)
	..A/.....	Yes		Steingrimsdottir et al (1992)
Kappa light-chain immunoglobulin (human)	Deletion	Yes		Seidman et al (1980)
MPC-11 kappa-chain fragment (human)	Deletion		Yes	Sikder et al (1985)
Ornithine transcarbamylase (human)	..A/.....	Yes		Hodges et al (1989)
Phenylalanine hydroxylase (human)	.../A.....	Yes		Marvitt et al (1987)
Pro α1(I) collagen (human)	..A/.....	Yes		Weil et al (1989)
Pro α2(I) collagen (human)	.../.C....	Yes		Weil et al (1988)
	..A/.....	Yes		Weil et al (1989)

Table 1.1
In vivo phenotype of mutation of the 5′ splice site of internal exons.
(modified from Talerico et al., 1990)

Kole, 1994). The distance between the branch point sequence and the pyrimidine tract is critical for efficient lariat formation (Gattoni et al., 1988)

The polypyrimidine tract is essential for efficient branch point utilisation and 3′ splice site recognition in metazoans in constitutive and alternative splicing. Several splicing factors such as U2AF, PTB and hnRNP C bind the polypyrimidine tract. It is known that polypyrimidine tract deletion and mutation decrease splicing efficiency (Roscigno et al., 1993) and the elongation of the number of pyrimidine in a row can led to improved efficiency (Dominski and Kole, 1991). However in the constitutively included exon 3 of the human apolipoprotein AII the flanking polypyrimidine tract is substituted by a $(GT)_{16}$ repeat (Shelley et al., 1985). Furthermore a common characteristic of the fibronectin EIIIB alternatively skipped exon, the β- and α-tropomyosin mutually exclusive exons and the E1A transcript are unusually long polypyrimidine tracts (Huh and Hynes, 1993; Gattoni et al., 1988; Goux-Pelletan et al., 1990; Smith and Nadal-Ginard, 1989). Interestingly these alternative 3′ splice sites have unusual lariat branchpoint located at greater distance (70-100 nucleotides) from the 3′ splice sites. It is likely that a balance of the relative strength and positions of branchpoint sequences and polypyrimidine tract will determine the efficiency of splice site usage.

1.2.1.3 Exon size requirements

Only less than 1% of the known internal exons in vertebrate are longer than 400 nucleotides. In in-vivo experiments expansion of internal exons to length above 300 nucleotides determine the activation of cryptic splice site inside the exon or exon skipping (Robberson et al., 1990; Berget, 1995).

A minimum length requirement seems also to be required to prevent steric hindrance between the complexes assembling at the 5′ and 3′ splice sites of the exon (Black, 1991). Experiments in the troponin T (Xu et al., 1993) and in the β-globin (Dominski and Kole, 1991) systems showed a minimum length requirement of about 50 nucleotides to obtain efficient splicing. Furthermore an increased matching to the consensus of the splicing sites lead to an alleviated requirement in exon length (Xu et al., 1993; Dominski et Kole 1992).

1.2.1.4 Pre-mRNA secondary structure

Secondary structure are very important for the splicing of type I and II self-splicing introns. It is becoming clear that RNA folding participates in the regulation of pre-mRNA splicing as well. In the adenovirus E1A transcription unit branch points are located unusually upstream of the 3′ splice site (50-60 nucleotides upstream). An hairpin structure necessary to bring the branch point in proximity of the 3′ splice site has been demonstrated using site directed mutagenesis (Chebli et al., 1989). An RNA secondary structure is also involved in the regulation of the mutually exclusive 6A and 6B exons in the β-tropomyosin gene (Helfman et al., 1988). A stem loop structure was shown to inhibit splicing of the 6B exon (Libri et al., 1991). However it seems likely that RNA sequence which have been defined as regulatory elements does not inhibit or enhance splicing by forming a particular secondary structure but interacting with some trans acting factors, as proposed for the β-tropomyosin (Guo et al., 1991). Most of the demonstration that higher order structure are involved in pre-mRNA splicing derives from in vitro splicing experiments and are not validated by in-vivo experiment or in vitro systems where transcription and splicing are coupled. Therefore most of the conclusion on the functional significance of these structures should be steel confirmed.

1.2.1.5 Intronic regulatory sequences

Intronic sequences are also involved in splice site selection and exon inclusion. The gag gene of Rous sarcoma virus (RSV) contain an intronic negative splicing regulator (NSR) that can inhibit splicing in heterologous context. NSR act upon binding the spliceosomal U1/U2 snRNPs and the non spliceosomal U11/U12 snRNPs whose functions are still unknown but probably regulate specific RNA processing events (Gontarek et al., 1993).

The mouse c-src gene has an unusually short (18 nucleotides) exon (N1) that is inserted into the src mRNA in neurones and skipped in all other tissues. A specific 123 nucleotides sequences at the 5′ end of the intron downstream the N1 exon is required for the correct splicing of the exon. In fact it has been proven that

factors bound to the intronic splicing enhancer are required for its effect (Black, 1992).

Fibroblast growth factor receptor 2 and the rat fibronectin gene also contain intronic regulatory regions that will be discussed later (see chapter 3).

1.2.1.6 Exonic regulatory sequences

In several pre-mRNAs specific exon sequences are involved in splice site recognition. Exon 2 of the episialin gene MUC1 either use one of two 3′ splice site. The splice site recognition is allele dependent and is based on a single A/G nucleotide difference in exon 2 8 nucleotides downstream of the second 3′ splice site (Ligtenberg et al., 1990). The pre-mRNA of the calcitonin (CT)/calcitonin gene-related peptide I (CGRP-I) gene is alternatively processed. The alternative inclusion of exon 4 in specific cell types led to the production of two different mRNAs, the CT mRNA and the CGRP-I mRNA. Two exonic sequences located in exon 4 are both essential for the inclusion of the exon in the spliced product. Deletion of one of these sequence result in production of both messengers, however deletion of both results in the complete switch from the production of CT RNA to CGRP RNA in CT specific cells (van Oers et al., 1994). One of the best characterised exonic regulatory elements is the sequence that control the somatic inhibition of the Drosophila P-element third intron splicing. P element somatic inhibitor (PSI) (Siebel et al., 1994), a non-snRNP splicing factor, bind to an exonic inhibitory element and stabilise U1 snRNP binding to the upstream inactive pseudo-5′ splice site. The multiprotein complex that form on the region that comprise the pseudo-5′ splice site and the inhibitory sequence interfere with the binding of U1 snRNP on the authentic splice site (Siebel et al., 1992).

1.2.1.7 Exon Splicing Enhancer (ESE) sequences

Many exons contain sequence elements essential to their recognition named ESE (Exon splicing enhancer). A number of these sequences are purine rich and many have a general consensus sequence of GARGARGAR (R is a purine) (Dirksen et al., 1994; Watakabe et al., 1993; Xu et al., 1993; Tsukahara et al., 1994; Steingrimsdottir et al., 1992; Katz and Skalka, 1990). ESE sequences have been

demonstrated to be binding sites for a novel class of non-snRNP splicing factors named SR proteins (Lavigueur et al., 1993; Sun et al., 1993; Ramchatesingh et al., 1995) (see section 1.4).

Purine rich sequences from various genes and synthetic (Pu)n repeat can facilitate splicing of heterologous introns in in-vivo (Xu et al., 1993) and in in-vitro experiments (Tanaka et al., 1994). However alternating polypurine sequences can function as splicing enhancer, while poly(A) and poly(G) sequence can not. The presence of U residues within the polypurine sequence greatly reduces the level of stimulation (Tanaka et al., 1994). ESE have been shown to increase inclusion of alternatively skipped exons (Tsukahara et al., 1994; Steingrimsdottir et al., 1992) or to increase in-vitro splicing of an upstream intron (Watakabe et al., 1993). The human caldesmon gene contain an unusually long (1090 nucleotides) exon that is spliced at two regulated 5′ splice sites. Four identical 32-nucleotide purine-rich element located between the two 5′ splice sites promote exon inclusion stimulating the use of the internal 5′ splice site (Humphrey et al., 1995). Other examples of long polypurine-rich repeats have also been found in database inside several large first and last exons.

Cardiac troponin T (cTNT) exon 5 is a developmentally regulated alternatively spliced exon. An internal polypurine sequence is necessary for the recognition of the 3′ splice site of the short (30 nucleotides) exon 5. Although necessary for the recognition of the exon by the general splicing machinery the ESE within cTNT exon 5 is not required for regulation of alternative splicing. Indeed the ESE is necessary for exon 5 recognition in different cell types and its deletion can be overcame by an increased exon length without lost of the tissue specific pattern of expression.

1.3 EXON VERSUS INTRON DEFINITION

How does the vertebrate splicing machinery recognise small exons, usually not more than 300 bases, amid introns long up to dozens of thousand of bases?

How does it discriminate among the correct splice sites, defined by short poorly conserved sequences, and frequent cryptic sites interspersed in both intronic and exonic sequences?

Early models of splicing and splice site selection suggested that the intron was the unit recognised by the splicing machinery and proposed a scanning mechanism in which the recognition of one site was followed by scanning through the intron to locate the second site (Lang and Spritz, 1983; Reed and Maniatis, 1986). This mechanism has been convalidated in systems were 5′ splice sites closer to the 3′ splice site are chosen in preference to those further away (Eperon et al., 1993; Reed and Maniatis, 1986; Cunningham et al., 1991). However the scanning mechanism is not universally used as demonstrated by experiments on duplicated splice sites (Kuhne et al., 1983) and trans-splicing (Konarska et al., 1985).

In 1990 Susan Berget's group proposed a model in which the exon was the unit recognised by the splicing machinery and the identification of splice sites was facilitated by interactions across the exon (Robberson et al. 1990). Basically the exon definition theory propose that in pre-mRNA with multiple short exons and long introns the splicing machinery search for two closely spaced splice sites. At that point U1 snRNP, U2 snRNP and the non-snRNP factors associated bind defining the exon. The next step involve the juxtaposition of the neighbouring exons probably aided by the same factors that recognise the exon.

Recently a growing number of reports are supporting both theories. Intron and exon could be thus both recognised by the splicing machinery and this could depend by a number of factors such as intron or exon length, splice sites sequences and other cis-acting sequences.

1.3.1 Evidence for an Exon Definition Model

a) Exon size requirement.

As exposed in section 1,2,1.3 exon length can affect splicing. Internal vertebrate exons have minimum and maximum length requirements. The length

of an exon is thus a factor involved in its recognition, while introns usually do not have particular length constrains.

b) Effect of splice site mutations.

In-vitro (Talerico and Berget, 1990; Kuo et al., 1991; Dominski and Kole 1991) and in-vivo (Xu et al., 1993; Dominski and Kole, 1991; Dominski and Kole 1994) experiments revealed that mutations in a splice site flanking an exon influence the splicing of both introns flanking the exon and not only of the intron bearing the mutated splice site as predicted from intron recognition oriented theories. Furthermore mutations at the 5′ splice site are suppressed by mutations that improve the consensus of the upstream 3′ splice site (Carothers et al., 1993).

c) A network of interaction span the exon.

Experiments carried out on the preprotachykinin pre-mRNA showed evidence for exon bridging interactions that occur between U1 snRNP bound at the 5′ splice site and U2AF bound at the 3′ splice site. U2AF is indeed recruited to the polypyrimidine tract by interactions with the downstream 5′ splice site and the U1 snRNP (Hoffman and Grabowski, 1992). Recently SR proteins have been implicated in these interactions in different experimental models (see section 1.4.1 and Fig 1.5).

d) Exon enhancer sequences.

As explained in sections 1.2.1.6 and 1.2.1.7 alternatively spliced exons with suboptimal splice sites can bear peculiar sequences that work as enhancer of splicing helping in the recruitment of splicing factors on an exon otherwise not recognised.

e) First and last exon recognition.

Since the lack of one functional splice site the first and the last exon should require a special mechanism for their recognition. To have an efficient first intron removal the capping and the proteins that bind the cap are essential (Iazaurraide et al., 1994). Therefore both, factors that recognise the cap and the 5′ splice site, are

necessary to define the first exon. The last exon is usually longer than the internal exon (average length 600 nucleotides versus 137). The removal of the last intron involves RNA splicing and polyadenilation factors (Gunderson et al. 1994).

1.3.2 Evidence for an Intron Definition Model

Experimental evidence generally suggest that in pre-mRNAs with small introns is the intron instead of the exon to be recognised by the splicing machinery.

a) The yeast model.

In yeast, were messengers often have unique introns and their length is usually below 100 nucleotides, exist direct support for the pairing of splice sites across introns during the first step of spliceosome assembly (Goguel and Rosbash, 1993).

b) Vertebrate intron/exon architecture.

In vertebrate most exon are small and are flanked by long introns. However some exceptionally long exon exists and when an expanded exon is placed in a gene in which the flanking introns are small, the exon is constitutively included. Experiments on a three exon minigene containing short intronic sequences showed that an expanded size of the middle exon up to 1400 nucleotides does not lead to exon skipping (Chen and Chasin, 1994).

c) Drosophila intron/exon architecture.

In drosophila most exons are 100 to 180 nucleotides in length; however 15% are more than 550 nucleotides. 50% of the introns are less than 100 nucleotides and are often flanked by large exons (Hawkins, 1989). Thus there are a number of gene in Drosophila that have an inverted intron/exon architecture respect the one found in vertebrate genes. At least in three drosophila genes 5′ splice site mutations of small introns cause intron retention instead exon skipping as

observed in vertebrate (Talerico and Berget, 1994; Berget, 1995)). As for vertebrate short exons Drosophila short introns seems to have size limitation. Expanding the size of the short introns cause repression of splicing or activation of cryptic site. Further evidence for a peculiar mechanism in the recognition of short introns derive from the observation that spliceosomial complex A formation require sequences at both end of the intron but not a polypyrimidine tract, while in vertebrate require only sequence at the 3′ splice site comprensive of the polypyrimidine tract (Talerico and Berget, 1994).

1.3.3 Coexistence and Co-Operation of Intron- and Exon-Based Models

Intron and exon definition mechanism are both validated from the experimental evidences. Depending on the kind of intron/exon architecture the intron or the exon and the flanking splice sites are the unit recognised by the splicing. In vertebrate exist a predominant recognition of short exons interspersed between longer introns. In yeast short and often unique introns are recognised. In Drosophila although most genes fit neatly into the two categories, genes with short introns and large exons and genes with short exons and large introns there are a considerable number of gene in which a mixed architecture is present and in which both mechanism could operate on the same pre-mRNA.

1.4 SR PROTEINS

In recent years a new class of non-snRNP associated proteins involved in pre-mRNA splicing has been characterised. This family called SR proteins for their serine- and arginine-rich carboxyterminal have originally identified since are recognised by a monoclonal antibody (mAb104) that recognise active sites of RNA polymerase II transcription (Roth et al., 1991). SR proteins are highly conserved throughout evolution, both human and Drosophila have a similar set of 6 proteins

that are recognised by mAb104 with similar structure and extensive sequence homologies (Zahler et al., 1992, Zahler et al., 1993a). Up to now at least 7 human SR proteins have been identified with different techniques by different groups. These have been named (with a multiple nomenclature according with the groups that independently identified the protein): SRp20, SRp30a (ASF/SF2), SRp30b (SC35/PR264), SRp40, SRp55 , SRp75 and 9G8. Each single SR proteins has the ability to complement a S100 extract, which contains all the splicing machinery except a limited number of factors, restoring splicing activity (Zahler et al., 1992).

SR proteins have been isolated in a variety of organisms (human, mouse, rat and drosophila) and structural comparison of the different SR protein sequences revealed peculiar structural feature (Zahler et al., 1992; Zahler et al., 1993a). All SR proteins contain an amino terminal RNA binding domain (RBD), also named RNA recognition motif (RRM), a structurally conserved motif found in hundreds of RNA binding proteins (Kenan et al., 1991). All SR proteins contains a serine-arginine rich domain at their carboxy-terminal region. This domain varies in length among the different SR proteins and many of the serine are phosphorylated by a cell-cycle regulated kinase specific for SR proteins, SR protein kinase 1 (Gui et al., 1994). However the significance of phosphorylation and dephosphorylation of these proteins remain unclear. SR proteins can also interact one to another and with other splicing factor that contains analogous arginine/serine rich regions (Tra and Tra2 regulatory element in Drosophila, U1 70K and $U2AF^{35}$) via the SR domain (Wu and Maniatis, 1993).

SRp30a (ASF/SF2) has been identified in two independent system for its ability to complement splicing deficient S100 extracts (Krainer et al., 1990) and for its influence on 5′ splice site selection, activating sites proximal to the 3′ splice site in substrate containing a distal competing 5′ substrate (Ge and Manley, 1990). The enhancement of the proximal 5′ splice site by SRp30a appear to be mediated by an increase in the strength of U1 snRNP binding to all 5′ splice sites indiscriminately (Eperon et al., 1993). In addition hnRNP A1 antagonise this effect by promoting the use of distal 5′ splice sites (Mayeda and Krainer, 1992; Mayeda et al., 1993). Different expression of the two factors influence alternative splicing in in-vivo

experiments (Caceres et al., 1994). SRp55 can substitute in in-vitro assay SRp30a in constitutive splicing and 5′ splice site selection (Mayeda et al., 1992).

SRp30b (SC35/PR264) share only 30% sequence homology with SRp30a but has equivalent activity in constitutive and alternative splicing activity (Fu et al., 1992). Recently has been demonstrated that in the SV40 pre-mRNA SRp30b influence splice site selection targeting the U1 snRNP to both distal (large T) and proximal (small t) 5′ splicing sites and in this case the proximal site is used. This is as predicted by the exon-intron definition theory in pre-mRNA with small introns (see section 1.3.2). Another SR protein (SRp40) promote production of large T messengers targeting U1 snRNP only to the distal 5′ splice site and this result in the prevalent utilisation of this site (Zahler and Roth, 1995).

1.4.1 SR Proteins Functions

1.4.1.1 Constitutive splicing

As exposed in section 1.1.2.4 pre-mRNA is committed to the spliceosomal assembly pathway at the formation of the prespliceosomal early complex. It has been demonstrated that two specific enhancer complex E5′ and E3′ assemble independently on RNAs containing only a 3′ or a 5′ splice site (Michaud and Reed, 1993). SR proteins promote initial recognition of both 3′ and 5′ splice sites interacting directly with the pre-mRNA. SRp30a, SRp30b and SRp20 UV cross-link to the RNA in the E5′ complex at the 5′ splice site sequence (Zuo and Manley, 1994; Eperon et al., 1993). SRp30a and SRp30b have been shown to promote U1 snRNP binding to the 5′ splice site via protein-protein interaction with the U1 snRNP component U1 70K (Kohtz et al., 1994; Staknis and Reed, 1994). SRp30a and SRp30b stimulate E3′ complex formation as well, binding specifically the 3′ splice site and via a direct protein-protein interaction with the 35 kDa subunit of the splicing factor U2AF (Staknis and Reed, 1994). The pre-binding of SR protein with pre-mRNA is sufficient for commitment to the splicing pathway (Fu, 1993). SR protein-RNA binding is thus a previous step in spliceosome assembly. These observations and the functional association detected between the 5′ and the 3′

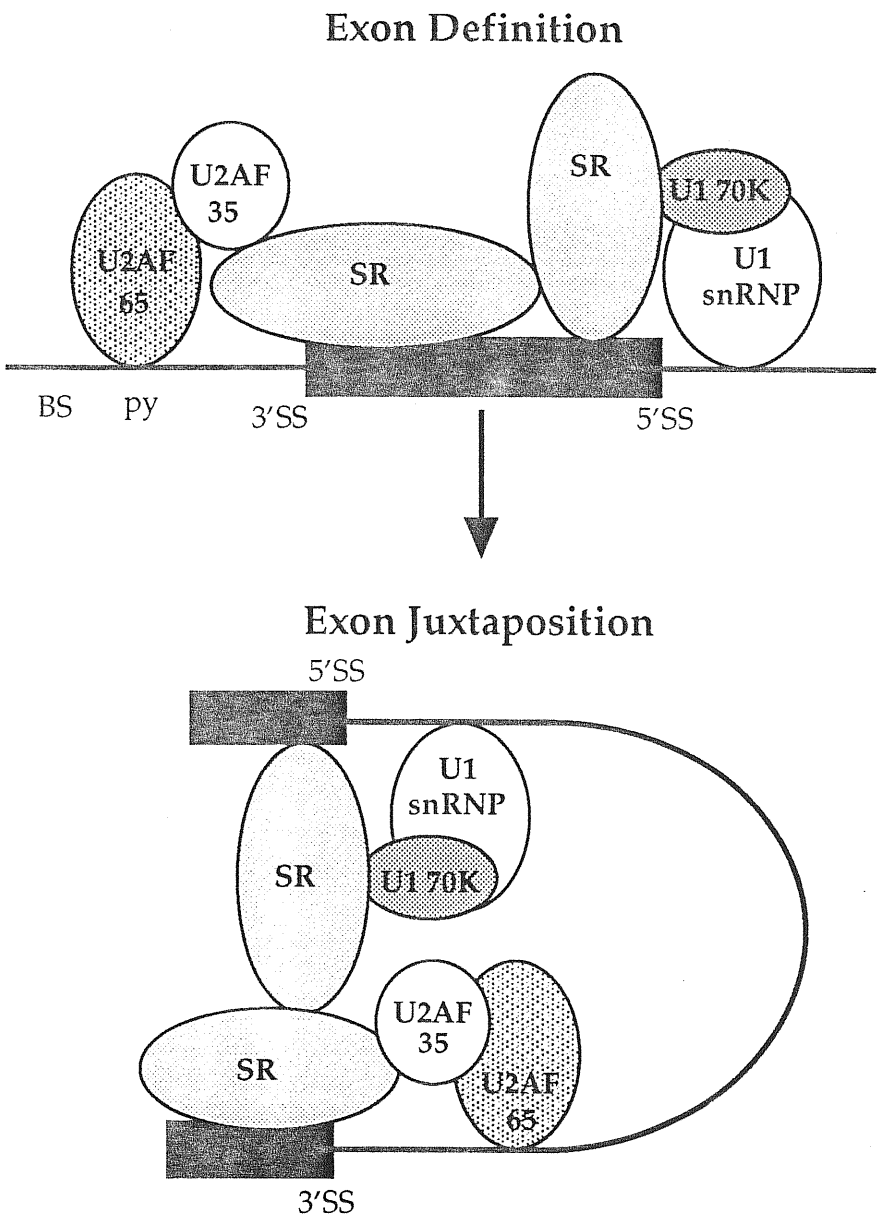


Figure 1.5. SR proteins role in exon recognition.
Schematic representation of the interaction involving SR proteins in the early step of splicing. SR proteins can help both, the recognition of the exon and the juxtaposition of neighbouring 5′ and 3′ splice sites as hypothesised by the exon recognition theory (see section 1.3).

splice sites in the early complex (Michaud and Reed, 1993) suggest that 5′ and 3′ splice sites are brought together by a network of interaction involving U1 snRNP, U2AF and SR proteins.

1.4.1.2 SR proteins, alternative splicing and the exon definition model

SR proteins seems to be essential splicing factors with redundant functions. Each individual SR protein can complement S100 splicing deficient extracts and this independently from its source since Drosophila SRp55 can efficiently complement HeLa extracts (Mayeda et al., 1992). However SR proteins are highly conserved in size and sequence during evolution, this apparent paradox suggest essential function for the single SR proteins in alternative splicing regulation. Furthermore western blot analysis showed that the relative levels of the single SR proteins vary in different tissues (Zahler et al., 1993b) giving consistence to the hypothesis of SR proteins as alternative splicing regulators.

In 1992 Staknis and Reed demonstrated that U1 snRNP bound to the 5′ splice site specifically target the U2AF 65 kDa subunit to the upstream polypyrimidine tract throughout a network of interaction spanning the exon. This finding and the SR protein bridging activity between U1 snRNP and U2AF strongly supported an exon definition model in which the exon and the surrounding splice sites are recognised as a whole and in a cooperative manner. SR proteins, thus, could recognise the exon and promote juxtaposition of adiacent 5′ and 3′ splice sites as proposed in figure 1.5. In this model the presence of a weak 3′ splice site that leads to exon skipping can be overcome by the presence of a strong downstream 5′ splice site by a mechanism in which U1 snRNP and the SR proteins associated to the 5′ splice site facilitate the binding of $U2AF^{65}$ through a network of RNA:protein and protein:protein interactions (Hoffman and Grabowski, 1992; 1995).

Direct interaction between SR proteins and ESE sequences has been demonstrated in the fibronectin EDA exon (Lavigueur et al., 1993), in the bovine growth hormone (Sun et al., 1993) in the troponin T alternative exon 5 (Ramchatesing et al., 1995) and in the avian sarcoma-leukosis virus env gene (Staknis and Reed, 1994). These studies suggest that different exonic sequences are

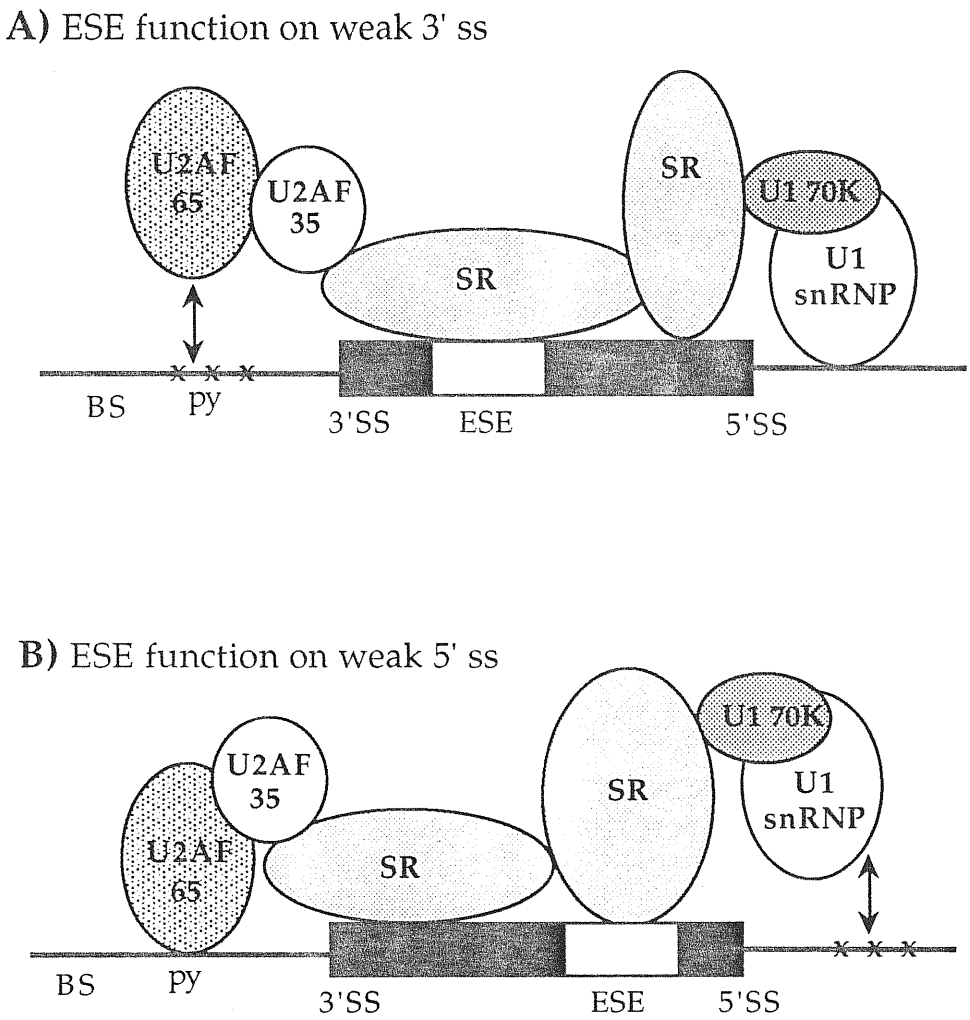


Figure 1.6 ESE functions at weak 3′ and 5′ splice sites.

Exonic Splicing Enhancer (ESE) can compensate the functions of constitutively weak 3′ (A) and 5′ (B) splice sites. ESE sequences bind SR proteins enhancing, thus, the recruitment of splicing factors and the recognition of the splice sites.

bound by different sub-set of SR proteins, however the exact binding sequences of the individual SR proteins has not been established. Furthermore a polypurine stretch is not always required for SR protein binding since the polypurine splicing enhancer contained in the ASLV env gene can be modified substituting 6 A residues with 6 U residues without losing enhancer activity. The substitution shift the binding ability of the enhancer that bind SRp40 when the purine are present while after the U substitution bind SRp30a and SRp30b loosing affinity for SRp40 (Staknis and Reed, 1994). Assembly of an early-like complex containing SR proteins and U1 snRNP has been evidenced on the bovine growth hormone and ASLV ESE sequences (Staknis and Reed, 1994). Direct proof of the ESE stimulation of U2AF binding to the polypyrimidine tract mediated by SR protein has been obtained using the rat-preprotachykinin pre-mRNA as a model. Insertion of heterologous ESE sequences in an exon which splices predominantly by exon skipping, due to the presence of suboptimal 5′ and 3′ splicing sites flanking the exon, promote U2AF binding to the polypyrimidine tract via SR protein:ESE interaction (Wang et al., 1995). Has been hypotised that Exonic Enhancer Sequences (see sction 1.2.1.7) act via this aided-binding model, assembling an early-like complex containing SR proteins and U1 snRNP that promote the interaction between U2AF and the polypyrimidine tract substituting in this way the functions of weak splicing sites (Fig. 1.6) (Staknis and Reed 1994; Wang et al., 1995).

1.4.2 The Drosophila Sex Determination Pathway

The sex determination pathway in Drosophila melanogaster is the best characterised example of alternative pre-mRNA splicing regulation (Reviewed in Moore et al., 1993). The ratio of X chromosome to autosome determines the activation of the early promoter of the Sex-lethal (sxl) gene in female (Keyes et al., 1992). In male a non functional truncated form of sxl protein is produced, while in female sxl autoregulate its own expression activating a late promoter and binding to the polypyrimidine tract of exon 3. In this way it compete with U2AF

"blocking" the expression of this splice site and promoting, thus, the utilisation of the downstream 3′ splice site flanking exon 4 (Valcarcel et al., 1993). Exon 3 contains several stop codons and its skipping allow the synthesis of functional sxl protein. With the same mechanism the sxl protein mediate exclusion of exon 2 of the transformer (tra) mRNA (Boggs et al., 1987). Tra pre-mRNA exon 2 also contain stop codons and its exclusion result in the production of the female specific tra protein. Tra protein together with transformer 2 (tra2), a protein expressed in both male and female, controls the splicing of the alternative splicing of doublesex (dsx) pre-mRNA. The male (default) splicing of dsx pre-mRNA produces an mRNA consisting of exons 1, 2, 3, 5 and 6 encoding a factor necessary for male sexual differentiation. The female splicing produces a mRNA consisting of exons 1, 2, 3 and 4, which encodes a factor necessary for female differentiation (Burtis and Baker, 1989). Tra and tra2 activate the 3′ splice site upstream exon 4 which has a low consensus sequence and is not recognised by the splicing machinery. Tra2 binds directly to a regulatory element located 300 nucleotides downstream the female-specific 3′ splice site characterised by the presence of 6 repeated 13-nucleotides repeated elements and a polypurine sequence, named the dsxRE element (Hedley and Maniatis, 1991; Tian and Maniatis, 1993; Lynch and Maniatis, 1995). Tra, tra2 and SR proteins assemble in a multiprotein complex on the dsxRE element and are essential to commit dsx pre-mRNA to the female pathway of splicing (Tian and Maniatis, 1992; 1993). Similar to the SR proteins tra2 contains both a RRM and a SR domain while tra contains only a SR domain. Tra and Tra2 interact with each other, with SR proteins and with other general splicing factor via their SR domain (Wu and Maniatis, 1993; Amrein et al., 1994). A multiprotein complex assembled on the dsxRE facilitate the assembly of spliceosomal components on the weak female-specific 3′ splice site through a network of protein-protein interactions similar to the one proposed for pre-mRNAs containing weak splicing sites and ESE sequences (see section 1.4.1.2). Dsx is thought to be the final regulatory gene in the pathway. It binds directly to gene enhancer sequences and has been proposed to be a transcriptional repressor (Burtis et al., 1991).

1.5 THE FIBRONECTIN AS AN ALTERNATIVE SPLICING MODEL

In order to look in more detail at the mechanism which lead to and regulate alternative splicing, expression of the human fibronectin (FN) gene has been studied in the work presented in this thesis. Alternative splicing of FN was discovered in 1983 almost simultaneously by the groups of F.E. Baralle (Kornblihtt et al., 1983), in Oxford and R. O. Hynes (Schwarzbauer et al., 1983), at the MIT. At that time alternative splicing was found only in a dozen of mammalian genes. More than a decade afterwards, hundreds of gene (approximately 1 of every 20) are known to express diversity by alternative mRNA splicing. Nevertheless, the FN gene remains paradigmatic for its complexity. In fact, it has been chosen by P. Sharp in its Nobel lecture (Sharp, 1994) to illustrate the adaptative value of split genes through exon shuffling and alternative splicing.

1.5.1 Fibronectin Structure

FNs are a family of large, adhesive glycoproteins which are found in blood plasma, extracellular matrices (ECMs), basement membranes and on the surface of a variety of cell types. They have been studied in a wide range of organisms, including mammals, birds, amphibians and invertebrates such as sea urchins and Drosophila. Structure and functions of FN proteins have been the subject of various reviews (Kornblihtt and Gutman, 1988; Paolella et al., 1993; Rouslahti, 1988; Mosher, 1989; Hynes, 1990).

Two major form of FN exist: a soluble dimeric form that is synthesised by hepatocytes and secreted into the bloodstream (plasma FN), and a dimeric or cross-linked multimeric form, made by fibroblast, epithelial and other cell types, which is deposited as fibrils in the ECM (cellular FN). Both forms are made up of similar polypeptides of about 250 kDa, which share a common modular organisation consisting of repeating homology units of 40, 60 and 90 amino acids, termed type I, II and III repeats respectively (Fig. 1.7). The three modules are also found in a wide range of vertebrate proteins. Particularly type III repeats appear in

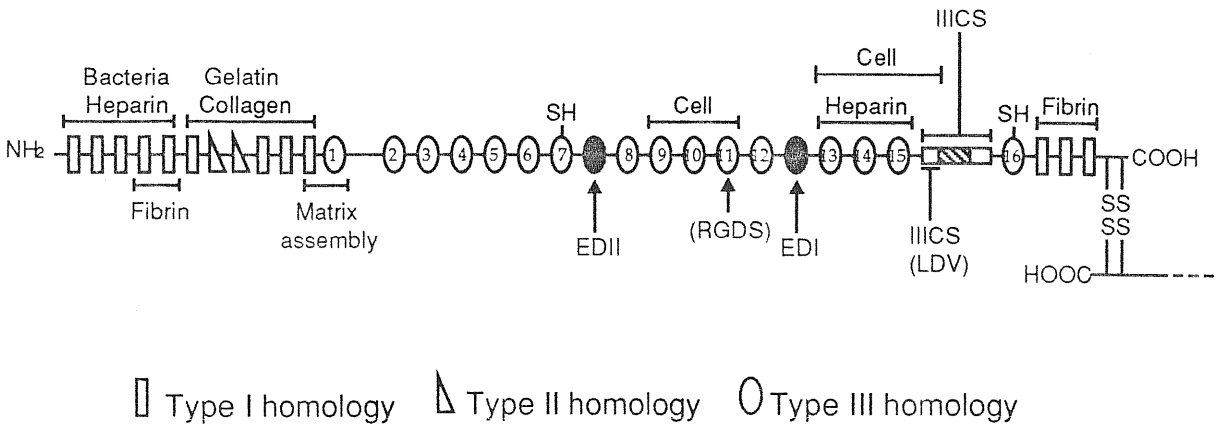


Figure 1.7. Schematic representation of the primary structure of the fibronectin molecule.

Type I, II and III repeats and binding sites are indicated. Alternatively spliced exons are highlighted. Position of the disuphide bridge is indicated.

a number of adhesive proteins including tenascin, undulin, certain collagens, contactin and N-CAM as well as in some intracellular domains of membrane receptors (Bork and Doolittle, 1992).

The human and rat genes are both single-copy, about 75 Kb long, contain about 50 exons and are transcribed by a single promoter into a single primary transcript (Schwarzbauer et al., 1987; Kornblihtt et al., 1984). All but one of the 12 type I repeats and the two type II repeats are encoded by a single exon each, while 14 of the 17 type III are encoded by two exons each. The other type III repeat (EDA, EDB and EIII9) are encoded by a single exon each.

1.5.2 Binding Properties

1.5.2.1 Cell-binding sites

The tetrapeptide Arg-Gly-Asp-Ser (RGDS) present in the tenth constitutive type III homology was originally described as the minimum sequence needed for interaction with the cellular receptor (Piershbacher and Rouslahti, 1984). RGDS or closely related peptides are also present in several ECM molecules other than FN (e.g. vitronectin, laminin, type I collagen) and in proteins involved in platelet adhesion and other functions. It is now clear that a whole family of RGD peptides interact with different affinities with many of the components of the membrane bound receptors, integrins. Particularly the FN RGDS cell-binding site interact mainly with the $\alpha_5\beta_1$ integrin (reviewed in Hynes, 1992). Although the RGDS motif has been identified as a key adhesive motif its activity is only a small fraction of the activity of larger FN fragments. Synergistic sites in the eight and ninth type III domain for the $\alpha_5\beta_1$ integrin-dependent fibronectin functions have been identified (Aota et al., 1991). Studies to fibronectin binding to $\alpha_{IIb}\beta_3$ integrin identified an 11-residue peptide from the ninth type III element that contains a novel integrin-binding site (Bowditch et al., 1994). The alternatively spliced IIICS domain (see section 1.5.6.1) and the adjacent heparin-binding domain provide a second major cell-binding region of fibronectin (CS1), which recognises the integrins $\alpha_4\beta_1$ and $\alpha_4\beta_7$ (Mould et al., 1990; Guan and Hynes, 1990). A key

minimal sequence, Leu-Gly-Asp (LDV), has been identified in the IIICS domain, but the tripeptide exhibit only a fraction of the activity of the full length spliced sequence, suggesting that synergistic sites are involved in this fibronectin/integrin interaction (Komoriya et al., 1991).

1.5.2.2 Collagen binding sites

Fibronectin ability to bind collagen is a key factor in the anchorage of cells to the extracellular matrix. The binding affinity is higher for denatured type I collagen (gelatin), but also other types of collagen can bind. The essential binding site has been located finding an essential sequence that includes the amminoacids placed immediately downstream from the second type II homology and extends farther to include the connecting region and the beginning of the following type I homology (Owens et al., 1986).

1.5.2.3 Bacterial-, heparin-, and fibrin-binding sites

A bacterial binding site is located in the amino-terminal domain of fibronectin, which recognises a surface protein determinant of staphylococci. Both a protective role, as a site for opsonization, and a negative role, as a specific receptor for the bacterium, have been assigned to this interaction (Hook and Svitalski, 1989). Heparin and fibrin both bind fibronectin in two distinct sites. Heparin binding sites interact with proteoglycans and are probably capable of modulating the attachment of fibronectin to the cell surface. Fibrin binding is important in the organisation of the clot and, in wound healing. However, no detailed mapping has been carried out, other than the initial assignment to proteolytic fragments as in figure 1.7.

1.5.3 Fibronectin Functions

The various binding sites of FN have been involved in many different functions. FN molecules interact with the cell surface and promote attachment and spreading of cultured cells. Fn is also able to promote cell migration, therefore

different activities are promoted by FN. It is not surprising that many fibronectin variants are involved in these functions.

1.5.3.1 Cellular fibronectin

The mechanism by which FN is involved in the formation of extracellular matrix is not completely clear; however, a model is now commonly accepted in which the protein acts as a bridge between the cell surface, to which it is attached by the cell binding sites (see section 1.5.2.1), and other component of the extracellular matrix such as collagen fibres. Moreover, other sites, such as heparin-binding sites may intervene and regulate the interaction. Heparin-binding sites bind to proteoglycans of the cell membrane and increase the strength. However, they may also interact with soluble proteoglycans, thereby weakening the FN-cell interactions. The binding of FN to cells is also mediated by other proteins: tenascin inhibits binding by competing with fibronectin (Paolella et al., 1993). In this complex set of interactions the presence of many fibronectin variants offer further possibilities for the anchorage system of the adult and developing organism. Cellular fibronectin is, in fact, a mixture of related protein isoforms that differ in both primary structure and posttranslational modification. Ascribing a specific biological mining to the various fibronectin isoforms has not yet been possible. However, the production of FN variants has been tested, at both the RNA and the protein level, in cell types and tissues. The cellular FN contains, at least in some extent the alternatively spliced EDA and the EDB domains (see section 1.5.6). The EDA and EDB amount is accurately regulated and their ratios vary in different cell types and during the stages of development (chapters 4 and 5).

1.5.3.2 Plasma fibronectin

In contrast with cellular FN, which is found as a insoluble multimer, plasma FN is a soluble dimer composed of slightly different subunits. Plasma FN is synthesised in the liver and its subunits are of a better-defined composition than those of cellular FN. They lack both the EDA and EDB alternatively spliced domains. A single major function has not be demonstrated for plasma FN, but it is implicated in various phenomena, such as clot formation and clearance, by the

reticuloendothelial system, of particles covered by FN (Paolella et al., 1993). This opsonization function is related to the FN capacity to bind fibrin, tissue debris and bacteria and could be important in the defence from pathogens, in coagulation and in response to injury. A role is also proposed for FN in tissue repair and wound healing (Knowlton et al., 1992; ffrench-Constant et al., 1989). FN concentrate immediately in the injured areas and seems to be involved in the first reactions to the trauma forming a matrix that favours the assembly of granulation tissue. FN extravased from the circulating blood is present in the early phases, while in the healing wound it is synthesised in situ (ffrench-Constant et al., 1989).

1.5.4 Control of Fibronectin Gene Expression

FN is synthesised in apparently all mammalian tissues, however, in vivo data showing FN mRNA regulation under physiological condition are scarce. Changes in mRNA levels occur in vivo in several pathological conditions including skin wound healing process (ffrench-Constant et al., 1989), hepatic fibrosis (Jarnagin et al., 1994), inflammation and malignant transformation (Mosher, 1989). FN gene expression is modulated by a number of stimuli such as forskolin, an activator of adenylate ciclase that influence steady state mRNA levels (Dean et al., 1989), physiological treatments including glucocorticoids as dexamethasone (Dean et al., 1988) and growth factors as transforming growth factor β1 (Ignotz et al., 1987).

The human (Dean et al., 1987), rat (Patel et al., 1986) and mouse (Polly and Nicholson, 1993) promoter have been cloned and shown to share extensive homology. The human promoter contains sequence conferring response to TGF-β, serum and cAMP, mimicking the regulation of the endogenous gene (Dean et al., 1990). Expression of the FN gene has provoked particular interest since it decrease in most transformed cell types, obtained by spontaneous, chemical or viral transformation. The loss of FN is described as one of the major cause of the reduced adhesion and altered morphology of transformed cells. In addition, several studies show that tumoral cells with reduced FN mRNA levels have an increased metastatic ability and tumorigenicity (Mosher, 1989). However, we still

do not have precise hints on the mechanism by which this is achieved, except for some recent preliminary works on viral transformation. Adenovirus E1A oncoprotein transforms, in fact, rat fibroblasts and reduces FN transcription through the stimulation of a repressor activity termed G10bp (G10 binding protein) (Nakamura et al., 1992).

1.5.6 Alternative Splicing of Fibronectin pre-mRNA

Alternative splicing is responsible for all the primary structure variants of fibronectin. Three sites have now been described: from 3′ to 5′, IIICS (type III homology Connecting Segment), EDI (Extra Domain I) and EDII (Extra Domain II) (Fig. 1.7). EDI and EDII are single exons coding for single type III repeats, that are included or skipped during splicing. The IIICS region undergoes a more complex pattern of splicing originating 2 variant in chicken (Norton and Hynes, 1987), 3 in rat (Schwarzbauer et al., 1983) and 5 in human (Gutman and Kornblihtt, 1987). As a consequence of all possible combinations, the maximum number of resulting FN variants is 8 in chicken, 12 in rat, 20 in human (Fig. 1.8).

1.5.6.1 The type III connecting segment (IIICS)

This region, located towards the C-terminus, correspond to a segment with no internal homology, connecting the last two type III units (III-14 and III-15). It is also referred to as the V (variable) region in the rat FN molecule. Its alternative splicing involves a complex system of alternative 5′ and 3′ splice sites which is not conserved between human, rat and chicken. Variable lengths of amino acid sequence are inserted into the peptide, in the absence of frameshifts (Fig. 1.8). The different splicing pattern between human and rat is probably due to the mutation of the internal 5′ splice site that varies from the consensus GT in human to AT in rat. Given the degree of variation in splicing of IIICS, this raises the question of the significance of the individual forms. The 5 human variants in this region are named IIICS-0, IIICS-120, IIICS-95, IIICS-89, IIICS-64 according to their lengths in amino acid residues. only the first three are present in rat.

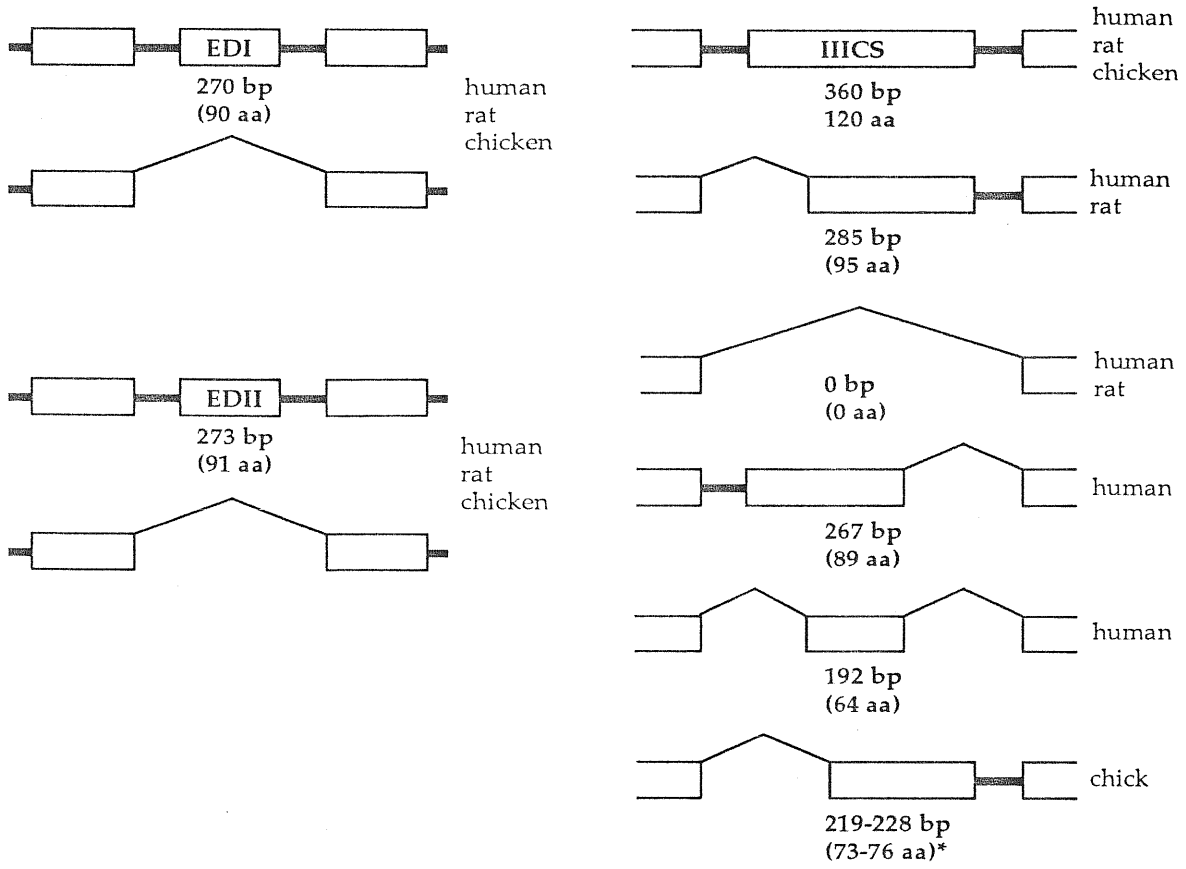


Figure 1.8. Representation of the alternatively spliced regions EDI, EDII and IIICS of the fibronectin messenger.

The splicing variants, their lenght in nucleotides and the number of amino acids predicted to be present in the translated polypeptide is indicated. Species were the different variants can be found are also indicated. In chicken IIICS marked * the value is based on S1 mapping of mRNA, therefore the precise sequence of the splicing variant is not known.

Alternative splicing of IIICS seems to be both cell type and developmentally regulated. Plasma FN made in hepatocytes contains up to 50% of IIICS-0 subunits, while fibroblast synthesising cellular FN make very few IIICS-0 subunits (Paul et al., 1986). Inclusion of the IIICS, and in particular of its subsegment CS-1 consisting of the amino terminal 25 amino acids, is significantly higher in all fetal than in adult tissues in both rat (Pagani et al., 1991) and human (Mardon and Sebastio, 1992). Furthermore the IIICS region is required for the secretion of FN dimers during the biosynthesis, few homodimers lacking this region are, indeed, secreted. The IIICS, probably, play a key role in intracellular protein-protein interactions that regulate dimer secretion (Schwarzbauer et al., 1989). Plasma FN is mostly IIICS-/IIIICS+, whereas cellular FN is IIICS+/IIICS+. This asymmetry of IIICS in FN dimers seems to be the molecular basis for the ability of plasma FN, but not cellular, to be cross-linked to fibrin by coagulation factor XIIIa during blood clot formation (Wilson and Schwarzbauer, 1992).

As exposed in section 1.5.2.1 the IIICS contain a cell binding site whose major determinant is a LDV sequence within CS-1 that is completely conserved among human, rat, bovine and avian FN (Komoriya et al., 1991). The membrane receptor for CS1 is the integrin $\alpha_4\beta_1$, that is mainly localised on lymphoid cells (Guan and Hynes, 1990). Certain cell types like melanoma, lymphocytes, monocytes and neural crest cells bind and spread on FN through their $\alpha_4\beta_1$ integrin, which recognise the alternatively spliced LDV sequence. Other cell types, such as fibroblast, use mainly the $\alpha_5\beta_1$ integrin for binding the RGDS site. These studies provide a key role for the alternatively spliced IIICS region in the selective adhesion of different cell types to specific isoforms of FN. However, the precise function of IIICS in cell allocation or specific signal transduction is still obscure.

1.5.6.2 Extra domain A (EDA)

The extra domain A (also named EDI and EIIIA in rat) is a type III repeat that can be alternatively skipped in the final FN messenger. Its inclusion is distinctive of cellular FN. In fact, the EDA sequence was identified in mRNA from fibroblast and a number of cell types but not from liver mRNA (Gutman and Kornblihtt, 1987). Furthermore antibodies against the EDA sequence recognise cellular FN but

not plasma FN. In vivo, EDA+ FN is poorly represented in the ECM of adult normal tissues. However, this variant is overexpressed in proliferating tissues such as developing embryos (Vartio et al., 1987, Pagani et al., 1991), and malignant liver tumours (Oyama et al., 1989). EDA inclusion has been reported to increase 8 fold in senescent fibroblast by in vitro passage (Burke and Danner, 1991) but no inclusion occur in senescent endothelial cells (Pagani et al., 1993). Higher level of inclusion of this region in tumoural versus normal cells is controversial (Hynes, 1990). What seems to be clear is that EDA expression in adult tissues increases in specific pathological circumstances such as wound healing (ffrench-Constant et al., 1989; Brown et al., 1993), epithelial fibrosis (Barnes et al., 1994), vascular intima proliferation (Glukhova et al., 1989) and peripheral nerve injury (Mathews and ffrench-constant, 1995). It has been speculated that this change in the splicing pattern provides an extracellular matrix that facilitates wound repair, perhaps by promoting cell migration. Most interestingly, EDA+ FN, which is absent from tunica media of normal adult human arteries and rat aorta, reappears in the human artherosclerotic plaque and in experimentally induced thickening of rat aorta, in a characteristic population of vascular smooth muscle cells which have undergone conversion from a contractile to a fibroblast like phenotype (Glukhova et al., 1989). A valid approach to determine the specific functions of the alternatively spliced EDA region comes from experiment performed recently by Jarnagin et al. (Jarnagin et al., 1994) on the FN expression in fibrotic liver. As in cutaneous wound healing, fibrogenesis that follow liver injury is characterised by the appearance of myofibroblast. While the myofibroblast in cutaneous wound are of uncertain origin, in liver injury they arise from lypocites (fat storing cells also called ito cells). In normal liver, lypocites are synthetically quiescent, while in fibrosis injury they pass to an "activated" state. Upon activation cytokine receptors and ECM production are upregulated, features of smooth muscle cells appear and the cell acquire the phenotype of myofibroblast. Among the earliest detectable changes in the ECM of the injured liver is an increase in total FN. The most conspicuous event in FN expression is a sharp increase in the synthesis of EDA+ FN by sinusoid endothelial cells , to an extent that more than 80% of the FN matrix made by these cells has the EDA sequence. The FN matrix deposited by sinusoidal

endothelial cells coming from fibrotic liver is able to induce the conversion of lypocites into myofibroblast in vitro. Moreover, conversion is also activated by a recombinant EDA polypeptide, which suggest that EDA is biologically active per se. This finding strongly argues against the hypothesis of a significant synergetic interaction with the nearby RGDS cell binding site. So far no cell binding sites have been mapped on the extra domain and it was postulated to alter the conformation of the known binding sites rather than acting via a specific receptor. Results on lypocites activation constitute the first evidence of a specific function for EDA and suggest new hypothesis that should be tested. Namely, the existence of a specific receptor, whether other kind of cells regulate FN expression as sinusoidal endothelial do, contributing to conversion of other cell types in wounds or artherosclerotic plaques and the nature of the signal transduction pathway in the activation events.

Though the cis and trans acting elements that regulate alternative mRNA splicing in general are beginning to be deciphered, very little is known about the molecules that trigger these events. Changes in the alternative splicing of EDA that occur in healing and proliferation must be controlled by external signal. Among these, TGF-β, released by platelets, is known to initiate the cascade of events resulting in the formation of granulation tissue in wound healing (Borsi et al., 1990). Cells that do not synthesise EDA+ FN in vivo, like hepatocytes and vascular smooth muscle cells, begin to synthesise EDA+ FN mRNA when placed into cultures, suggesting that this change in alternative splicing is triggered by the exposure to growth factors.

1.5.6.3 Extra domain B (EDB)

EDB (also named EDAI and EIIIB in rat) was the last alternative splicing site discovered in FN mRNA (Gutman and Kornblihtt, 1987; Zardi et al., 1987; Schwarzbauer et al., 1987). Hepatocyte mRNA, and consequently plasma FN, also lack EDB. EDB+ FN is more abundant in early embryos than in adult tissues (Pagani et al., 1991), in cultured cells than in the organ from where they were obtained (Magnuson et al., 1991), in cutaneous wounds (ffrench-Constant et al., 1989) and injured nerves (Mathews and ffrench-Constant, 1995). Its expression

seems also to be stimulated by TGF-β (Borsi et al., 1990). It is evident that in most cases the appearance of EDB correlates well with that of EDA, however in the rat hepatic fibrogenesis system (see section 1.5.6.2) only hepatic lypocites synthesise EDB+ FN forms while endothelial cells do not (Jarnagin et al., 1994). A distinct expression of EDA and EDB is also observed in hepatoma cell lines Hep G2 and Hep 3B. EDA is expressed in both cell lines, while EDB+ FN is expressed only in Hep G2 cells (Barone et al., 1989). These results suggest that regulation of EDA processing would involve different mechanism from those controlling EDB splicing and this turned out to be true as exposed in chapter 3. EDB amino acid sequence is more conserved in evolution than any other type III repeat (100% between human and rat and 96 between human and chicken), which strongly suggests a biological function for this segment. However, up to now no specific role has been assigned to EDB, except for the observation that alternative splicing of this region produce a change in the conformation of the preceding type III element that can be detected by a specific monoclonal antibody (Carnemolla et al., 1992).

Chapter 2

Materials and Methods

2.1 MATERIALS

2.1.1 Chemical reagents

Chemical were purchased from Sigma Chemical Ltd. and Carlo Erba Pharmaceutical and were of Technical or Analar grade.

2.1.2 Enzymes

Restriction endonucleases were supplied by Pharmacia, New England Biolabs Inc. and Boehringer Mannheim and were used according to the manufacturer's instructions.

T4 DNA ligase, DNA Polymerase I large fragment (klenow enzyme), Calf intestinal Phosphatase, T4 Polynucleotide Kinase were purchased from New England Biolabs Inc. and used according to manufacturers instruction.

Taq DNA polymerase was purchased by Promega and Cetus and Boehringer Mannheim and were used as indicated in each reaction.

RNAase A was purchased from Sigma Chemicals Ltd. A 10 mg/ml solution of RNAase A was prepared in 5 mM EDTA and boiled for 10 minutes to destroy trace amounts of DNAase activity.

Proteinase K was purchased from Sigma Chemicals Ltd. in a 10 mg/ml solution.

2.1.3 Radioactive isotopes

Radioactive α-^{32}P dCTP, γ-^{32}P dATP and ^{35}S were supplied by Amersham U.K. Ltd.

2.1.4 Bacterial strains

Strains of the K12 E. Coli family were maintained in short term as single colonies on agar plates at 4°C and for longer term storage were kept on glycerol stocks, made by adding sterile glycerol to a final 15% v/v t an overnight culture of bacteria grown in LB. Glycerol stocks were stored at -20°C.

The DH5, JM 101 and JM 103 strains were used for transformation by plasmid and growth of the plasmid.

2.1.4 Liquid media

LB: Luria-Bertani medium per litre:

Difco Bactotryptone 10 g, Oxoid yeast extract 5 g, NaCl 10 g, pH 7.5.

Bacterial growth media were sterilised before using by autoclaving. Where appropriate antibiotics, ampicillin and tetracycline, were added to media to a final concentration of 75 µg/ml (ampicillin) and 50 µg/ml (tetracycline).

DMEM (Flow) culture media was used supplemented with 10% Fetal Calf Serum, glutammin 2 mM and gentamicin.

2.1.5 Standard solutions

All solutions have been defined in the text where first used except for:

TE: 10 mM tris-HCl pH 7.4, 1 mM EDTA pH 7.4

PBS: NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 10 mM, KH_2PO_4 1.8 mM, pH 7.4

10x TBE: 108 g/l tris, 55g/l boric acid, 9.3 g/l EDTA

5x Ficoll loading buffer: 0.25% w/v bromophenol blue, 15% w/v Ficoll type 400 in H_2O.

10x MOPS: 200 mM MOPS, 50 mM CH_3COONa, 10 mM EDTA pH 7.0

Formamide dye: 0.25% bromophenol blue, 0.25% xylene cianol, 20 mM EDTA in deionised Formamide

2.2 METHODS

2.2.1 Nucleic Acids Preparations

2.2.1.1 General techniques for nucleic acid purification and concentration

RNA and DNA were precipitated from aqueous solution containing 0.3M sodium acetate by adding ethanol, as described in Sambrook et al. (1989). The nucleic acid pellet was then resuspended in dH_2O at the required concentration. Protein contaminants were removed from aqueous DNA solutions by phenol-chloroform extraction (Sambrook et al., 1989).

2.2.1.2 Small scale preparation of plasmid DNA from bacterial cultures

Rapid purification of small amounts of recombinant plasmid DNA was basically performed the method described in Sambrook et al. (1989) based on alkaline lysis of recombinant bacteria. The final pellet was resuspended in 50 µl of dH_2O and 5 µl of such preparation were routinely taken for analysis by restriction enzyme digests.

2.2.1.3 Large scale preparations of plasmid DNA from bacterial cultures

Large scale preparations of plasmid DNA were carried out as described in Sambrook et al. (1989).

A 0.5 ml aliquot of a fresh overnight bacterial culture was used to inoculate 500 ml LB in a 2 l volume flask, which was then incubated with vigorous shaking at 37°C overnight. Bacteria were harvested by centrifugation for 5 minutes at 5000 rpm, 4°C.

DNA was then prepared from the harvested bacteria as described and centrifuged to equilibrium using caesium chloride the samples were centrifuged at 60,000 rpm for 16 hours at 20 °C in a Beckman Vti65.2 vertical rotor. Under these conditions separation of covalently closed plasmid DNA occurs on the basis of buoyant density. All the following steps were performed according to the protocol above. DNA was then recovered by ethanol precipitation and was resuspended in TE and stored at -20°C.

This method yielded 200 μg-5 mg depending on the plasmid grown.

2.2.1.4 Preparation of RNA from cells in culture

Preparation of total RNA was carried out according to the method of Chomczynski and Sacchi (1987), from one 60 mm diameter dish containing a confluent monolayer of cells. The medium was removed and cells were washed twice with PBS. 1 ml of D solution (Guanidine tiocianate 4M, β-mercaptoethanol 100 mM, Na-citrate 25 mM, Lauroylsarcosine 0,5%) was added directly to cells. Cells not lysing instantly were scrapped from the plate and the whole mixture was transferred to a sterile 2 ml tube. All the following steps were performed according to the protocol above. The final pellet was then resuspended in 50 μl of ddH_2O and frozen at -80°C.

2.2.1.5 Animals and tissue preparation

6 and 24 months old Wistar rats were housed in a specific pathogen-free facility and fed on a standard rat chow diet. Animals of different age (0.4-0.6 kg) were used as indicate in each experiment. Rats were sacrificed and the tissues immediately frozen on liquid nitrogen and stored at -80°C.

2.2.1.6 Preparation of RNA from tissues

To prepare RNA from tissues the basic protocol of Chomczynski and Sacchi (1987) was followed. Frozen tissue (0.5-1 g) was homogenised in a correspondent volume (0.5-1 ml) of D solution using a UltraTurrax homogenizer. A further phenol-chloroform step was added to the basic protocol to improve the sample purification from proteins and fatty acids. The final pellet was then resuspended in 100 μl of ddH_2O and frozen at -80°C.

2.2.2 Estimation Of Nucleic Acid Concentration

2.2.2.1 Spectrophotometric

An optical density of 1.0 at 260 nm is roughly equivalent to a concentration of 50 μg/μl for double stranded DNA, 33 μg/μl for single stranded DNA and 40 μg/μl for RNA samples. the ratio of values for optical densities measured at 260 nm and 280 nm is 1.8 for pure sample of DNA and 2 for RNA, these are reduced by protein contaminants and therefore the values were used to asses both the concentration and the purity of the samples.

2.2.2.2 UV fluorescence of intercalated ethidium bromide

Nucleic acids size-fractionated on agarose or polyacrylamide gel electrophoresis can be visualised after staining with ethidium bromide (0.5 mg/ml) since fluorescence of ethidium bromide is enhanced by intercalation between bases in the nucleic acid helix. It is therefore possible to estimate the quantity of nucleic acid on a gel by comparing the intensity of UV-induced fluorescence of the sample with that of a standard sample of known quantity.

2.2.3 Electrophoresis Of Nucleic Acids

2.2.3.1 Agarose gels

DNA or DNA restriction fragments were size fractionated by electrophoresis in 0.8-1.8% w/v agarose gels containing ethidium bromide and 1x TBE. Horizontal 10x8 cm mini-gels were used routinely for fast analysis of DNA restriction enzyme digests, estimation of amount of DNA, or DNA fragment separation prior to elution from the gel. Sample of up to 20 μl containing 1x Ficoll dye were loaded into submerged wells. These were electrophoresed at 50-80 mA for a time depending on the fragment length expected and the gel concentration, in a running buffer of 1x TBE. DNA was then visualised by UV transillumination (the lower limit for this was about 10 ng DNA) and the result could then be recorded by Polaroid photography. When better resolution was required, for

example separation of PCR products with 5% of difference in length, samples were electrophoreses in vertical "Mighty Small" protein electrophoretic apparatus supplied by Hoefer. The contained width of 1.5 mm allowed agarose gels, used to separate radio labelled sample, to be easily dried and exposed overnight.

RNA was also visualised in 1.5% agarose gels containing EtBr, 1x MOPS buffer, and 2% formamide. Samples containing 1x Formamide dye were loaded into submerged wells. These were electrophoresed at 80 mA for 20 minutes in a running buffer 1x MOPS. RNA was then visualised by UV transillumination. The total RNA quality and quantity was deduced from the ribosomal 60s and 40s RNA bands that are clearly visible while the mRNA is not.

2.2.3.2 Polyacrylamide gels

DNA fragments of less than 500 nucleotides were fractionated according to size on gels containing 6% w/v acrylamide, 0.2% w/v bis-acrylamide, 1x TBE. to this mixture to catalyse cross-linking for each 50 ml gel mix, 400 µl TEMED were added. The gel was formed between two 380x200x0.3 mm glass plates separated by teflon spacers.

Radiolabelled samples to be loaded contained 1x formamide dye mix. These were subsequently electrophoresed vertically in a running buffer of 1xTBE at 50 W for a time depending on the length of the fragments analysed.

DNA fragments were detected by autoradiografy using X-O-Mat "S" film (Kodak) at room temperature or at -80°C with an enhancer screen for the required time.

2.2.3.3 Elution and purification of DNA fragments from agarose gels

This protocol was used to purify small amounts (less than 1 µg) of DNA for subcloning or radiolabelling. The DNA samples were loaded onto an agarose minigel and electrophoresed as described above.

The DNA was visualised with UV light and the required DNA fragment band was excised from the gel. This slab was put into the minimum length of dialysis tubing with as little gel buffer as possible and was electrophoresed in a minigel apparatus at 150 V for the time required for the DNA sample to migrate

out of the gel slice. At this stage the current was reversed for 30 seconds to elute any bound DNA from the dialysis tubing.

Buffer containing DNA was removed to a microcentrifuge tube and the tubing was washed twice with 100 µl of gel buffer. Aliquots were pooled and DNA recovered by ethanol precipitation. To remove small gel fragments remaining the aqueous solution was centrifuged for 10 minutes at 4°C and the supernatant was transferred to a fresh tube.

2.2.4 Identification and Analysis of Specific DNA Sequences

2.2.4.1 Detection of Specific DNA Sequences by Southern Blotting

Identification of specific DNA sequences cloned or specific products of PCR amplification was performed using the method originally described by Southern (1975). The samples were electrophoresed as described previously. The gel was then washed in denaturing and neutralising solutions and the DNA transferred on a nitrocellulose membrane as described (Sambrook et al., 1989).

a) Hybridisation and prehybridisation of filter bound nucleic acids

Filters were enclosed in plastic envelopes and were incubated for 1-2 hours with moderate agitation at 65°C (DNA probes above 100 nucleotides) or at 45°C (oligonucleotide probes), in approximately 10 ml of hybridisation buffer (SET 5x, lyophilised milk 2%. For hybridisation most of the solution was removed from the filters and denatured radioactively labelled probe was added and incubated, shaking, overnight at 65°C or 45°C. Probes were denatured by boiling for 5 minutes and were quickly cooled on ice.

b) Removal of non-specifically hybridised probe from Southern blot

Non-specific hybridisation of probes to the filters was reduced by multiple washes at decreasing salt concentrations as required. The following protocol was used:

a) 20 minutes at 65°C or 45°C in 3x SET, 0.5% SDS.

b) 20 minutes at 65°C or 45°C in 2x SET, 0.3% SDS.

c) 20 minutes at 65°C or 45°C in 1x SET, 0.1% SDS.

The filter was then dried and exposed using X-O-Mat "S" film (Kodak) at room temperature or at -80°C with an enhancer screen for the required time.

2.2.4.2 Sequence determination by chain termination

Plasmidic DNA has been sequenced using the "T7 Pharmacia Sequencing Kit" supplied by Pharmacia Inc. based on "dideoxy chain termination" method previously published by Sanger et al. (1980). As template we used 10 µl from the mini preparation of plasmidic DNA. As primer we used the pUC18 polylinker universal and reverse primer and specific primers when required. Samples were then denatured by heating 5 minutes at 95°C, electrophoresed in a 8% w/v polyacrylamide gel and autoradiographed.

2.2.5 Radiolabelling of Nucleic Acids

2.2.5.1 5′ end labelling of single stranded oligonucleotides with T4 Polynucleotide Kinase

The reaction for 50 ng oligonucleotide (RNA or DNA oligonucleotides) was carried out in a total volume of 50 µl. The reaction mix contained 1xKinase buffer (50 mM tris-Cl pH 7.6, 0.1 M $MgCl_2$, 5 mM dithiotreitol, 0.1 mM spermidine, 0.1 mM EDTA), 10 U T4 Polynucleotide Kinase and 20 µCi γ^{32}P-dATP. Samples were incubated at 37°C for 30 minutes before stopping reaction by addition of 2 µl 0.5 M EDTA pH 8.0. Radiolabelled oligonucleotides were purified from unincorporated nucleotides by ethanol precipitation.

2.2.5.2 Nick translation of DNA for use as hybridisation probes

Hybridisation probes were generated by nick tanslation of 20-100 ng DNA by the random priming method using the "Oligolabelling Kit" supplied by Pharmacia. Reaction were carried out in 50 µl for 30 minutes, 20 µCi α^{32}P-dCTP

were used. Radiolabelled DNA probes were purified from unincorporated nucleotides by ethanol precipitation.

2.2.6 Enzymatic Modification of DNA

2.2.6.1 Restriction enzymes

Restriction endonucleases recognise and cut within specific sequences of double stranded DNA leaving blunt ends, 5′ and 3′ protruding ends. These were used in the construction and analysis of recombinant plasmids. Each restriction enzyme functions optimally in a buffer of specific ionic strength. All buffers were supplied by the same company that supplied the enzymes and were used according with the manufacturer's instructions.

For analytical digests 100-500 ng DNA were digested in a volume of 10-20 μl containing 5 U of the appropriate restriction enzyme per μg DNA. The digest was incubated for 3-6 hours at the optimal temperature required by the enzyme used.

Preparative digest were made of 1-20 μg DNA using the above conditions but a larger reaction volume. Enzymatic activity was then removed either by incubation at 70°C for 20 minutes or phenol-chloroform extraction.

2.2.6.2 Large fragment of E. Coli Polymerase I

The large fragment of DNA Polymerase I (Klenow) lacks the 5′ to 3′ exonuclease activity of the intact enzyme, but retains the 5′ to 3′ exonuclease activity and 3′ to 5′ exonuclease activities. It is used to catalyse the polymerisation of deoxyribonucleotide triphosphates in a 5′ to 3′ direction on a template of double stranded DNA with a recessed 3′ hydroxyl and protruding 5′ phosphoryl terminus to generate a flush-ended DNA molecule. This was useful for creating compatible ends for ligation during construction of recombinant plasmids, to digest aspecific A residues added by Taq DNA polymerase at the 5′ terminus and for radiolabelling of DNA fragments by random priming.

Klenow fragment was used with the proper buffer supplied by Biolabs Inc. at a final concentration of 5U per μg DNA. When a "fill-in" was required (DNA

fragments with protruding 3′ ends) dNTP for a final concentration of 0.5 mM were added. The mixture was incubate at 37°C for 20 minutes.

2.2.6.3 Dephosphorylation of DNA 5′ termini

Calf intestinal phosphatase catalyses the removal of 5′ terminal phosphate groups from linear DNA molecules. This is used to reduce self ligation of vector DNA during generation of recombinant clones, thus increasing the proportion of resulting recombinants containing the required DNA inserts.

This reaction was carried out in a final volume of 50-100 µl using 1U of enzyme per 0.5 µg DNA incubating for 1 hour at 37°C. The enzyme has then be inactivated by adding 1 mM EDTA and incubating for 20 minutes at 75°C.

2.2.6.4 T4 DNA ligase

T4 DNA ligase catalyses the formation of a phosphodiester bond between adjacent 3′ hydroxyl and 5′ phosphoryl termini in DNA, requiring ATP as a cofactor in this reaction. This enzyme was used to join double stranded DNA fragments with compatible sticky or blunt ends, during generation of recombinant plasmid DNAs.

20 ng of linearised vector were ligated with a 5-10 fold molar excess of insert in a total volume of 20 µl containing 1x ligase buffer and 5 U T4 DNA ligase. Reaction was carried out at 16°C for 6-12 hours.

In some reactions synthetic oligonucleotide were included in the reaction. In these cases amounts added to each reaction to obtain inclusion of oligonucleotides in the resulting plasmid were about 100 fold molar excess over the DNA vector.

2.2.7 Transformation of Bacteria

2.2.7.1 Preparation of competent cells

E. Coli strains were grown overnight in a 200 ml volume of LB at 37°C. Afterwards cells have been incubated in ice for 15 minutes, centrifuged and the pellet resuspended in 20 ml of TPB1 solution (K-acetate 30 mM, KCl 100 mM,

$CaCl_2$ 10 mM, $MnCl_2$ 50 mM, glycerol 15%, pH 5.8. The solution has been incubated in ice for 5 minutes, centrifuged and the pellet resuspended in 2 ml of TPB2 solution (MOPS 10 mM, $CaCl_2$ 75 mM, KCl 10 mM, glycerol 15%, pH 6.5. The solution has incubated in ice for 15 minutes aliquoted and stored at -80°C.

2.2.7.2 Transformation

60 µl of competent cells have been transformed with 1-5 ml of the ligation product and incubated for 30 minutes in ice. The cells have then been incubated for 90 seconds at 42°C. 60 µl of LB have then been added and the cells incubated for 10 minutes at 37°C. The cells have then been plated on agarose plates containing the appropriate antibiotic. When DNA inserts have been cloned in the polylinker of the commercial pUC18 plasmid destroying the β-galactosidase activity of the plasmid 30 µl of IPTG 100 mM and 20 µl of x-Gal 3% have been added on the surface of the plates to test the growing colony for galactosidase activity. The plates have then been incubated for 12-15 hours.

2.2.8 Maintenance and Analysis of Cells in Culture

HeLa cells obtained from the tissue culture facility of the ICGEB have been used and all tissue culture work was performed using sterile techniques in a laminar flow tissue culture hood.

2.2.8.1 Maintenance of cells in culture

Cells were propagated in DMEM media and were maintained in 100x20 mm Falcon tissue culture plates, incubated at 37°C and with 5% carbon dioxide.

Plates containing a confluent monolayer of cells were passaged 1 in 5 with trypsin as follows. Cells were washed with 5 ml PBS solution, then were incubated at room temperature with 3 ml PBS/EDTA/trypsin solution (PBS containing 0.02% w/v EDTA and 5% v/v trypsin solution) for 2 minutes or until cells were dislodged. After adding 10 ml DMEM cells were pelleted by centrifugation in polypropylene tubes and were resuspended in 5 ml prewarmed medium. 1 ml of

this cell suspension was added to 10 ml medium in fresh plate and was gently mixed before incubation. This procedure was required on average once every three days.

2.2.8.2 Transfection of recombinant DNA into cells maintained in culture

Cells were passaged as above into 60x15 mm tissue culture plates and grown to a confluency of 40-70%. Transfections were performed with the DEAE dextran method as described (Sambrook et al., 1989). Cells were then incubated for 12-24 hours before RNA extraction. Transfection of HeLa cells used 3 μg of test plasmid together with an equal amount of plasmid DNA pβ5'SVBglII (Grosveld et al., 1982) added as a source of SV40 T antigen required for stimulation of replication of the test plasmids which all contained an SV40 origin of replication.

2.2.9 mRNA Analysis by Polymerase Chain Reaction

2.2.9.1 cDNA synthesis

The cDNA were obtained using the "Pharmacia cDNA Synthesis Kit" following manufacturer instruction. 1 μg of total RNA from cells or 5 μg of total RNA from tissues were used in a 20 μl reaction. The appropriate specific primer (as indicated in tab. 2.1) was used in each reaction. The reaction was incubated for 1 hour at 37°C.

2.2.9.2 Polymerase Chain Reaction

The polymerase chain reaction was performed following the basic protocols of the Boehringer and Promega Taq DNA Polymerases. The volume of the reaction was 50-100 μl with 1x Taq buffer, dNTP mix 200 μM each, oligonucleotide primers 2 μM each, Taq DNA Polymerase 2.5-5 U. When appropriate DMSO to improve the specificity of the reaction or α^{32}P-dCTP to obtain radiolabelled products were added. As template were used 1.5-3 μl of the cDNA reaction mix. If plasmidic or genomic DNA were amplified for cloning porpoises 0.1 ng or 100 ng of DNA were

respectively used. The amplifications were performed on a Perkin Elmer Cetus DNA Thermo Cycler.

PCR reactions performed to analyse the RNA samples were carried out as follows:

a) RNA extracted from transfected cells (see chapter 3).

Each cDNA was analysed by PCR using the primers PSV5'J and PSV3'J (Tab. 2.1) in a final volume of 50 µl. 1.5 µl of DMSO (Fluka) were added. Each amplification cycle was carried out as follows: denaturation step 45 seconds at 93°C, annealing step 1 minute at 60°C, extension step 40 seconds at 72°C. The number of rounds of amplification was 28.

b) RNA extracted from tissues for linkage analysis and splicing isoforms determination(see chapter 4 and 5).

PCR reactions were performed in a final volume of 100 µl adding 0.185 MBq of (α-^{32}P)-dCTP and 3% of dimethyl-sulfoxide. Each cycle was carried out as follows: denaturation step (1 minute at 93°C), annealing step (1 minute at 56°C), extension step (1 minute, 30 seconds at 72°C). The number of amplification cycles ranged from 21 to 24. Locations of the primers on the FN molecule, primers sequence and the predicted segment length of each amplification product are shown in figure 4.1 and 5.2 and table 2.1 and 4.1.

Control reactions for cDNA synthesis and following PCR amplification were set up including all the reagents except the RNA sample and were routinely performed for each series of RT-PCR assay.

2.2.9.3 Statistical analysis of the amplified products

Three independent RT-PCR assays have been carried out for each sample. The data analysis has been performed with the StatView (Abacus Concept, Inc.) program. The unpaired Student t test was used to determine significance. A p value of $< 0,05$ was considered significant.

2.2.9.4 Oligonucleotides primers

Synthetic DNA and RNA oligonucleotides were purchased by Primm s.r.l. (Milano). Their sequences and utilisation are outlined in table 2.1.

2.2.10 Analysis of the PCR-Amplified Products

2.2.10.1 Not radiolabelled products

10 µl of each amplification reaction were then analysed by electrophoresis on a 1.2% agarose gel, which was stained with ethidium bromide and subject to densitometric analysis. The optical density assigned to each band was then normalised on the basis of the length of the product.

2.2.10.2 Radiolabelled products

5 µl of each reaction product were analysed by vertical electrophoresis on a 1.2% agarose (SeaKem) gel, stained with ethidium bromide. Verification of the fragment size was done by visualisation of their migration relative to the molecular weight standard (1 kb DNA Ladder, Bio-England Research Laboratories) and by sequence analyses of the clones obtained from the purified PCR products. The gel was then dried under vacuum and exposed on Kodak X-Omat film. The intensity of each band was quantified by densitometric analysis. The value assigned to each band was then integrated taking into account the cytosine content and the amount of (α-^{32}P)-dCTP incorporated in each fragment.

2.2.12 Plasmid Construction

2.2.12.1 pSVEDA construction

The two complementary synthetic oligonucleotides EDA99dir and EDA99rev were annealed generating a synthetic fragment of 99 bp, with the sequence of the central part of the EDA exon and Sal I/Bam HI sticky ends. The fragment was then cloned into Sal I/Bam HI digested M13 mp18 obtaining the construct mp ED-A-99. The fragment -1/ED-A (1235 bp) comprising the last 7 bases of the exon -1, the 5' end of the exon EDA and the -1 intron was obtained after digestion of the vector pSVEDα1W/FN (Mardon et al., 1987) with Pst I and Bst EII. The synthetic fragment E/-1 (77 bp) carrying rest of the 3' part of the exon -1 and two sticky ends complementary to digested Sal I and Pst I sites was obtained after annealing the

two synthetic oligonucleotides E-13'dir and E-13'rev. The -1/ED-A and the E/-1 fragments were then contemporary cloned into the Sal I/Bst EII mp ED-A-99 digested construct, originating the mp-1 ED-A .

Finally the ED-A/+1 fragment (1676 bp) that contain the 3' end of the EDA exon, intron +1 and the 5' end of exon +1 was obtained from the Stu I/BstE II pSVα1W/FN digested vector, blunt ended and cloned into the Stu I site of mp-1 ED-A obtaining the final construct M13 mp18 EDA Tot. From this construct the EDA Tot (3078 bp) fragment was excised with Sal I and Bam HI, blunt ended and cloned into the Bst EII site in the third exon of the α1 globin of the pSVα1W vector (Higgs et al., 1983) obtaining the final construct pSVED-A Tot.

The pSVED-A Tot Sac vector was obtained cloning a mutated ED-A fragment carrying the Sac I site.

2.2.12.2 pSVEDA variants

The construct pSVED-A ΔXho-Stu was obtained after Sal I/Xho I and Sal I/Stu I digestion of the pSVED-A Tot Sac, blunt ending of the Xho I site and ligation of the Sal I/Xho I and the Sal I/Stu I fragments. The vector pSVED-A ΔXho-Sac was obtained excising the Xho I/Sac I fragment from the pSVED-A Tot Sac vector and religating it after blunt ending. The pSVED-A ΔSac-Stu construct was obtained after Sal I/Sac I and Sal I/Stu I digestion of the pSVED-A Tot Sac, blunt ending of the Sac I site and ligation of the two fragments. The constructs pSVED-A Δ1, Δ2c, Δ2e, Δ3, Δ4 were obtained cloning the Sal I/Sac I fragment, obtained from the vector pSVED-A Tot Sac, into the pUC18 vector obtaining the pUC-SS3.9 construct. Synthetic fragments Δ1, Δ2c, Δ2e, Δ3, Δ4 were obtained anealing the primers ED-Δ1-/ED-Δ1+, ED-Δ2c-/ED-Δ2c+, ED-Δ2e/ED-Δ2e+, ED-Δ3-, ED-Δ3+/ED-Δ4-/ED-Δ4+. Synthetic fragments Δ4B1, Δ4B2, Δ4B3, Δ4B4 were obtained annealing the degenerated primers SSrevM2/SSdirM2. After Sac I/Stu I digestion of pUC-SS3.9 the synthetic fragments Δ1, Δ2c, Δ2e, Δ3, Δ4, Δ4B1, Δ4B2, Δ4B3, Δ4B4 were ligated obtaining the final constructs were the Sal I/Sac I fragment was excised and ligated into the previously Sal I/Sac I pSVED-A Tot Sac digested vector.

2.2.12.4 pSVαΔ4 variants

pSVαBstINS has been originated after pSVα1W digestion with BstE II, blunt ending and insertion of the synthetic fragment SpeINS. The SpeINS was obtained after annealing of the primers SpeINS5′ and SpeINS3′. The pSVαΔ4+S, Δ4+AS, Δ4+2cS, Δ4BS, Δ4BAS variants have been originated after insertion of the fragments Δ4+ and Δ4B inside the Pml I digested pSVαBstINS vector in different orientation (Sense and Antisense) and in tandem (+2cS). The Δ4+ and Δ4B fragments have been obtained after annealing of the primers αgloΔ4+S/αgloΔ4+AS and αgloΔ4BS/αgloΔ4BAS.

2.2.12.5 pRFout260 and pGAPDH260

Total mRNA extracted from rat liver tissues was reverse transcribed using the primers FNcDNA and GAPDHcDNA. FN and GAPDH cDNAs were then amplified using the primers 5'OUT, 3'OUT and 5'GAPDH, 3'GAPDH respectively. Both amplifications gave a product of 598 bp. In order to produce a DNA standard for the competitive PCR assays the amplified fragments were cloned into the Sma I digested PUC19 vector, giving the two constructs pRF5'3'out and pGAPDH5'3'. The plasmids were digested with Apa I and Nco I respectively and blunt ended by fill in. Afterwards a stretch of 260 bp derived by PCR from the exon 10 and flanking regions of the human Eosinophil Peroxidase (Sakamaki et al., 1989) was cloned in both constructs giving the two plasmids pRFout260 and pGAPDH260.

2.2.11 Gel Retardation Assay

HeLa total nuclear extracts were obtained by the lysolecithin method as described by Zerevitz and Akusjarvi (1989). Nuclear extracts (up to 30 μg) were incubated with 1 ng of the appropiate probe at 30°C in the presence of Mg/Spermidine 4 mM, NaPi (Na_2HPO4) 1 mM, EDTA 0.05 mM, glycerol 10% v/v, dithiothreitol 1 mM in a total volume of 20 μl for 5-20 minutes. Complexes were separated on a 6% polyacrylamide native gel in 0.25x TBE for 2 hours at 180 volts. The gel were

than fixed, dryed and exposed at -80°C with an enhancer screen for the required period of time.

Primer	Sequence	Use
EIIIB5' (Sense)	AACCGAGTAGTGACACCGCT	Amplification EIIIB
EIIIB3' (Antisense)	ACACTGACTAGGTACTCAGT	Amplification EIIIB
PREIIIB- (Sense)	ATACCGTCATCCCAGC	Amplification EIIIB-EIIIA
EIIIA3' (Antisense)	ACTCGGTAGCCAGTGAGCTT	Amplification EIIIA
PREIIIB+ (Sense)	GACTATGACATCAGCGTT	Amplification EIIIB-EIIIA
EIIIA5' (Sense)	CCTGGTTCAGACTGCAGTGA	Amplification EIIIA
PREIIIA- (Sense)	TCAGACTGCAGTGACCAC	Amplification EIIIA-V
V3' (Antisense)	TGTCTGAGAGAGAGCTTCTT	Amplification EIIIA-V
PREIIIA+ (Sense)	TCAGACTGCAGTGACCAA	Amplification EIIIA-V
EIII95'I (Sense)	CTCTTGGTGCGCTACTCA	Amplification EIII9
EIII93'I (Antisense)	TGTCTCTCCATAGGTGAT	Amplification EIII9
EIII95'II (Sense)	GTGAAGAACGAGGAGGAT	Amplification EIII9 nested
EIII93'II (Antisense)	TCTGTAATAGCGCACAGA	Amplification EIII9 nested
cDNAFN (Antisense)	CTCAGAACTCTCCTGGAAT	cDNA FN synthesis
GAPDHcDNA	TGTTAAGGTAGGGTCTGG	cDNA GAPDH synthesis
5'GAPDH (Sense)	ACATGTTCCAGTATGACTCT	GAPDH quantitative PCR pGAPDH260 constuction
3'GAPDH (Antisense)	ACGGAAGGCCATGCCAGTGA	GAPDH quantitative PCR pGAPDH260 constuction
5'OUT (Sense)	TATGCTCTCAAGGACACA	FN quantitative PCR pRFout260 construction
3'OUT (Antisense)	CTGTCTTTTTCCTCCCAA	FN quantitative PCR pRFout260 construction
pSVcDNA	GGTATTTGGAGGTCAGCA	cDNA pSVEDA synthesis
PSV5'J (Sense)	CACTGCCTGCTGGTGACTCGA	pSVEDA amplification
TESTEDA5' (Sense)	ATGACTATTGAAGGCTTGC	Amplification EDA in mRNA from cells
TESTEDA3' (Antisense)	TCATTGGTCCGGTCTTCTC	Amplification EDA in mRNA from cells
PSV3'J (Antisense)	GCGGCCAGGGGTCACGAT	pSVEDA amplification

EDA99dir (Sense)	TCGACGTACAGGGTGACCTAC TCGAGCCCTGAGGATGGAATC CATGAGCTATTCCCTGCACCT GATGGTGAAGAAGACACTGC AGAGCTGCAAGGCC	pSVEDA construction
EDA99rev (Antisense)	GATCCAGGCCTTGCAGCTCTG CAGTGTCTTCTTCACCATCAG GTGCAGGGAATAGCTCATGG ATTCCATCCTCAGGGCTCGAG TAGGTCACCCTGTACG	pSVEDA construction
E-13′dir (Sense)	TCGACAGTGGAGTATGTGGTT AGTGTCTATGCTCAGAATCCA AGCGAGAGAGTCAGCCTCTG GTTCAGACTGCA	pSVEDA construction
E-13′rev (Antisense)	GTCTGAACCAGAGGCTGACTC TCTCCGCTTGGATTCTGAGCA TAGACACTAACCACATACTCC ACT/G	pSVEDA construction
SSdirM2 (Sense)	CTTCCCTGCACCTGATGGTGA AGAAGACACTGCAGAGCTGC A(AC)(TG)G	pSVEDA mutations (B element constructs)
SSrevM2 (Antisense)	C(AC)(GT)TGCAGCTCTGCAGT GTCTTCTTCACCATCAGGTGC AGGGAAGAGCT	pSVEDA mutations (B element constructs)
ED-Δ1+ (Sense)	CTGGTGAAGAAGACACTGCA GAGCTGCAAGG	pSVEDA mutations
ED-Δ1- (Antisense)	CCTTGCAGCTCTGCAGTGTCTT CTTCACCAGAGCT	pSVEDA mutations
ED-Δ2+ (Sense)	CTTCCCTGCACCTGATGGCAC TGCAGAGCTGCAAGG	pSVEDA mutations
ED-Δ2- (Antisense)	CCTTGCAGCTCTGCAGTGCCA TCAGGTGCAGGGAAGAGCT	pSVEDA mutations
ED-Δ3+ (Sense)	CTTCCCTGCACCTGATGGTGA AGAAGAAGCTGCAAGG	pSVEDA mutations
ED-Δ3- (Antisense)	CCTTGCAGCTTCTTCTTCACCA TCAGGTGCAGGGAAGAGCT	pSVEDA mutations

ED-Δ4+ (Sense)	CTTCCCTGCACCTGATGGTGA AGAAGACACTGCAGAGCTG	pSVEDA mutations (B element constructs)
ED-Δ4- (Antisense)	CAGCTCTGCAGTGTCTTCTTC ACCATCAGGTGCAGGGAAGA	pSVEDA mutations (B element constructs)
αglobE15′	GGTCACGCTGGCGAGTATGGT	α globin and pSVα1W constructs amplification
SpeINS5′ (Sense)	CTAGTAGTGCGGCCGCTCCAT ACAAGATGCGGCCGA	pSVα1W INS construction
SpeINS3′ (Antisense)	CTAGTCGGCCGCATCTTGTAT GGAGCGGCCGCACTA	pSVα1W INS construction and amplification
αgloΔ4+S	AGCTGCAAGGCCTCAGACCG G	pSVα1W INS constructs
αgloΔ4+A	CCGGTCTGAGGCCTTGCAGCT	pSVα1W INS constructs
αgloΔ4BS	AGCTGCAAGTCCTCAGACCG G	pSVα1W INS constructs
αgloΔ4BA	CCGGTCTGAGGACTTGCAGCT	pSVα1W INS constructs
RNAΔ4+	CAGAGCUGCAAGGCCUCAGA C	Gel retardation assay
RNAΔ4-	CAGAGCUGCCUCAGAC	Gel retardation assay

Table 2.1. Sequence and employment of the oligonucleotide primers used in this research.

Chapter 3

EDA exon alternative splicing mechanism

The mechanism for alternative splicing of the extra domains A and B have been studied in the past years by several groups using a common approach. Transient expression of a three exon minigene containing the alternatively spliced exon the flanking introns and part of the flanking exons has proven to be useful to characterise cis acting elements involved in the alternative splicing regulation of both regions. Preliminary observation carried out by Vibe Pedersen et al. (1984) showed that the alternative splicing of the human EDA exon can be accurately reproduced in HeLa cells by transient expression of a hybrid minigene in which a 3 Kb fragment of human FN gene, containing the EDA exon, its flanking introns and part of the neighbouring exons, was inserted in the third exon of the α-1 globin gene. This was the first report showing that alternative splicing could be duplicated in transfection experiments. Using the same experimental approach Mardon et al. (1987) obtained the first experimental evidence of the involvement of exonic sequences, not related to the splice sites, in the mechanism of exon recognition. In that work a 81 nucleotides sequence, located within the central region of the exon, was found to be essential for the EDA inclusion. The fact that the 81 nucleotides sequence did not stimulate EDA inclusion when reinserted in the opposite orientation in template DNA strongly suggested that there was at the RNA level a sequence specific effect. A recent work carried out by Lavigueur et al. (1993) showed that the 81 nucleotide region is also able to enhance splicing when assayed in an in vitro splicing system using HeLa cells nuclear extracts in an heterologoues context.

Alternative splicing of the human EDB exon was reproduced by transient expression of an α-globin-FN minigene in several cell lines (Barone et al., 1989). This study evidenced the existence of different splicing mechanisms involved in the regulation of the EDA and EDB exons. Indeed, in Hep3B cell line only a messenger lacking the EDB exon is produced after transfection with the EDB minigene and the same is observed in vivo in the liver. In contrast, the minigene containing the EDA area failed to reproduce the liver specific splicing pattern in Hep3B. As predicted by Barone et al. two recent works carried out by Huh and Hynes (1993; 1994) on transfection studies showed that EDB itself does not play a relevant role in its inclusion. On the contrary, the intron downstream of EDB contains elements that promote EDB inclusion in a cell type specific manner. These elements are a series of 7 non-tandemly repeated nucleotide examers, scattered in a region of 500 bases located in the intron downstream EDB. These repeated sequences seems to stimulate the 5′ splice usage of the FN EDB exon and are also active when placed downstream of unrelated alternatively spliced exons.

Notwithstanding the earlier observation of Mardon et al. there has not been a follow up study on the structural elements within the 81 nucleotides exonic deletion. We have now analysed in detail the cis-acting elements contained in this region and involved in the alternative splicing of the human EDA exon.

3.1 ANALYSIS OF SEQUENCE ELEMENTS KNOWN TO BE REQUIRED FOR RNA SPLICING OF THE EDA EDB AND EIII9 REGIONS

Of the three type III homology that are encoded by a single exon only two (EDA and EDB) are alternatively spliced, while EIII9 is constitutively included (see section 4.1.5). To understand the differences between the expression of these exons one approach is to analyse the DNA primary structure of the fragments encoding the information for the accurate splicing of the exon (EDA, EDB or EIII9 exons+flanking introns+neighbouring exons) and compare with sequence features

known to affect splicing pattern in other genes. Such elements include consensus sequences at 5′ and 3′ splice sites and the weak consensus at the branch point.

In figure 3.1 are evidenced the matches with the consensus sequences of the splice sites flanking the human EDA, EDB and EIII9 exons together with the 5′ splice site of the upstream exon and the 3′ splice site of the downstream exon. Previous work on the primary structure of these regions evidence the existence of additional predicted splice sites within intron and exon structures (Henchcliffe, 1988). Some of these showed greater homology with the consensus than the experimentally determined splice sites. This emphasise the importance of other factors (also probably sequence specific) in defining splice sites in pre-mRNA.

3.1.1 EDA splice sites analysis

The 5' and 3' splice sites of the EDA exon have a poor homology with the known consensus sequences. The 3' splice site upstream of the EDA presents a polypyrimidine stretch interrupted by purines. As exposed in section 1.2.1.2 this can decrease the splicing efficiency perturbing the binding of splicing factors as U2AF. In a similar way the downstream 5' splice site departs from the consensus sequence at the positions +4 and +5. The 5' and 3' splice sites flanking the EDA may then make exon definition dubious in the early spliceosome complex. This hypothesis was confirmed modifying the 5′ and 3′ splice sites of the EDA exon and matching them to the consensus. The modification of both or, either, one of the two splice sites flanking the exon gave complete inclusion in the final messenger of the EDA exon (Muro,A. personal communication).

3.1.2 EDB and EIII9 splice sites analysis

The human EDB exon is flanked by weak splice sites. The polypyrimidine tract is not well conserved bearing 7 pyrimidine out of 14 and only three in a row. Furthermore the 5′ splice site differ from the consensus in position +5 as for the EDA exon. Huh and Hynes (1993) demonstrated that a substitution of one or both splice sites improved the recognition of the rat EIIIB exon (EIIIB is the rat homologue of the EDB in human). Furthermore they suggested that a regulated state of the EIIIB splicing is maintained in part by a balanced competition between

```
EXON -1                                INTRON -1                                                EXON (EDA)                INTRON +1
     36 13 11  100100 60  9 82 53                11 12 53 51 49 11 38 38 47  8  25 72100100   26                36 56 78  100100 60 11  7  53
5'-...C C A / G T A C G T......1086 nt.....A A T T T G C C T A A C A G / A......1439 nt......C A G / G T A T A T...-3'
5'-...N A G / G T R A G T..................Y Y Y Y Y Y Y Y Y Y N C A G / R..................N A G / G T R A G T...-3'
        56 78  100 100 95 69 82 53          76 76 80 80 79 78 82 82 85 84   72 100 100  76                56 78  100 100 95 69 82 53

EXON -1                                INTRON -1                                                EXON (EDB)                INTRON +1
     36 56 78  100100 60 69  5  14               46  4411 51 11  1138 44 38  8  25 23100100   26                36 14 78  100100 60 69  7  53
5'-...C A G / G T A A T A.......1224 nt....T T G T A A C T C A A T A G / A.......1066 nt.....C G G / G T G A A T...-3'
5'-...N A G / G T R A G T..................Y Y Y Y Y Y Y Y Y Y N C A G / R..................N A G / G T R A G T...-3'
        56 78  100 100 95 69 82 53          76 76 80 80 79 78 82 82 85 84   72 100 100  76                56 78  100 100 95 69 82 53

EXON -1                                INTRON -1                                                EXON (EIII9)              INTRON +1
     36 56 78  100100 35 69  82 14               13  1211 51 10  11 10 10  8  8  25 72100100   26                36 56 78  100100 60 59  6  53
5'-...C A G / G T G A G C.......±1.1 kb....G G G T G G A A A A A C A G / A.......±1.9 kb.....C A G / G T A A C T...-3'
5'-...N A G / G T R A G T..................Y Y Y Y Y Y Y Y Y Y N C A G / R..................N A G / G T R A G T...-3'
        56 78  100 100 95 69 82 53          76 76 80 80 79 78 82 82 85 84   72 100 100  76                56 78  100 100 95 69 82 53
```

Figure 3.1. Comparison of splice sites from the human fibronectin gene with consensus sequences for 5′ and 3′ splice sites.

Intron exon boundaries are shown for EDA, EDB and EIII9, together with those of upstream (-1) and downstream (+1) exons. Sequences from the human gene (upper line) and consensus (lower line) are given. Points indicate homology between the two. Respective base frequencies are also indicated for both, consensus sequences and fibronectin gene. N indicates any nucleotide, Y pyrimidine and R purine. Length of each intron is given.

the 5′ splice sites of the upstream exon and of the EIIIB for splicing to the 3′ splice site of the downstream exon. This is because improving the match to the consensus of the upstream 5′ splice site there is lost of regulation with almost complete exclusion of the EIIIB exon.

Particularly interesting is the analysis of the splice sites flanking the EIII9 exon. The 3′ splice site absolutely lack a polypyrimidine tract and the 5′ splice site has a G/C divergence with the consensus in position +5 as observed for the EDA and EDB exons. The presence of particularly weak splice sites, especially the 3′ splice site that completely lacks the polypyrimidine tract, strongly suggest its alternative processing. This hypothesis was tested in section 4.1.5 but surprisingly a complete EIII9 inclusion was detected in a various extent of tissues and cell lines, emphasising the incomplete nature of our knowledge and the impossibility of making predictions based on RNA sequence alone.

3.2 pSVED-A CONSTRUCTS ANALYSIS

A hybrid minigene α1globin/fibronectin was constructed as described in methods (section 2.2.12.1), preserving the fibronectin gene structure in this region. A synthetic approach was chosen to insert the -1, EDA and +1 exons, including a minimum of linker sequences (Fig. 3.2 and section 2.2.12.1). Site directed mutagenesis was carried out to obtain a new Sac I site that allowed the manipulation of the central region of the EDA exon. The sequence of the region extending between the Xho I and Stu I sites in the centre of the EDA exon is shown in figure 3.3. In order to obtain the relative amount of ED-A+ and ED-A- isoforms by transfection of different constructs into a small number of cells (10^6) a PCR-based method has been set up. Specific fibronectin cDNA was synthesized from total RNA extracted from transfected cells and then amplified using the primers described in section 2.2.9.2 and table 2.1. These primers match specifically the sequences derived from the transfected DNA, that is their 3′ ends lie in the linker sequences as indicated in figure 3.2, thus they are able to amplify only the cDNAs

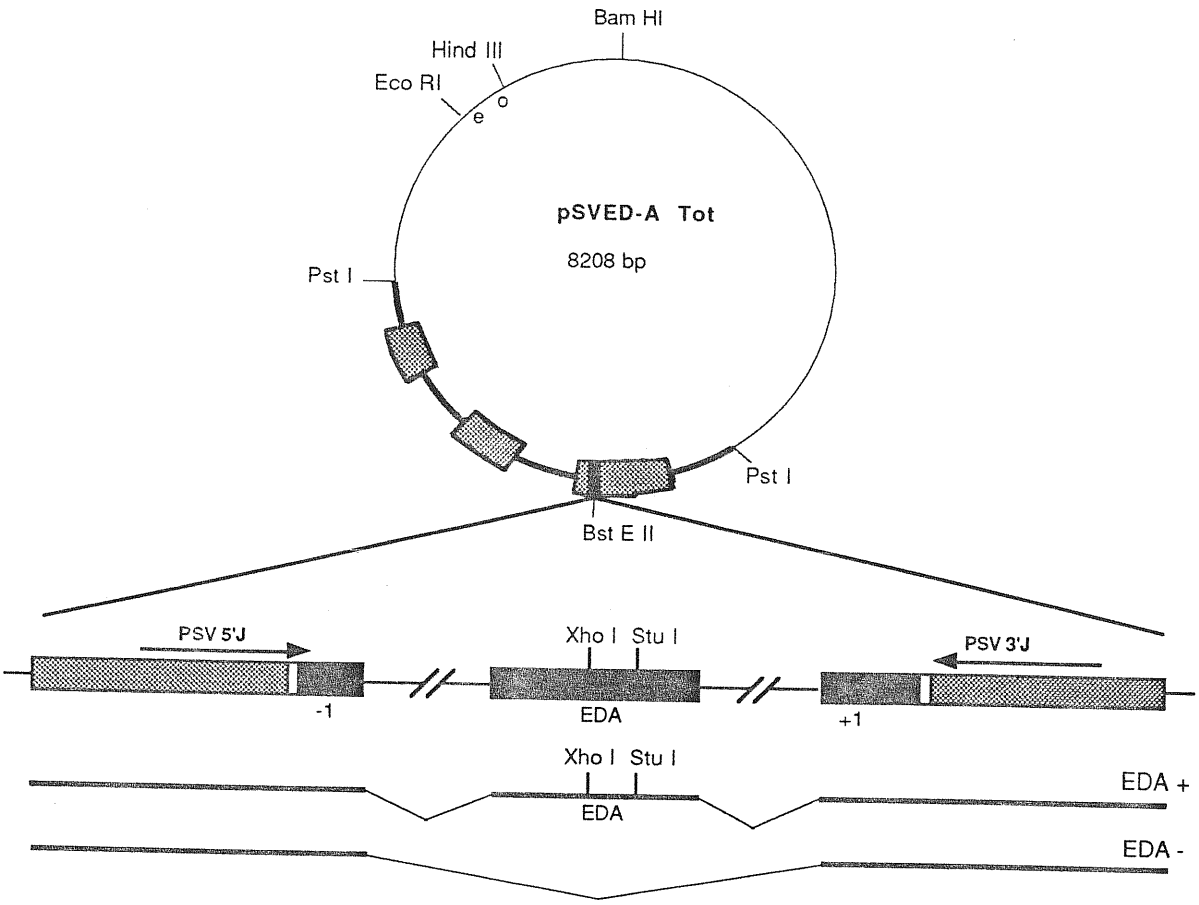
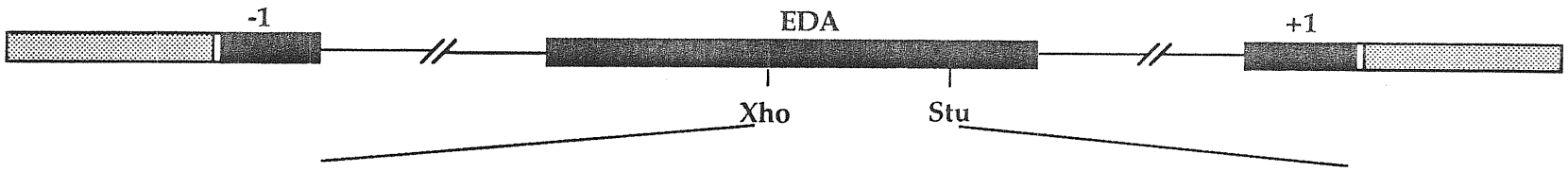


Figure 3.2. Schematic representation of the pSVED-A Tot construct.
The EDA human fibronectin exon with its -1 and +1 introns, plus 81 bp at the 3' end of the -1 exon and 112 bp of the 5' end of the +1 exon have been inserted in the BstE II site of the third α globin exon of the pSVα1W vector. Black boxes indicate fibronectin sequences, shadowed boxes α globin sequences, empty boxes polylinker sequences. The position of the primers used in the PCR analysis is shown together with the alternative splicing pattern of the EDA exon. The SV40 enhancer and origin of replication (e, o) are also indicated.



Splicing Pattern

Xho I Sac I Stu I

pSVED-A Tot .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUAUUCCCUGCACCUGAUGGUGAAGAAGACACUGCAGAGCUGCAAGGCCUCAGA.-3' **WT**

pSVED-A Tot Sac ..5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGUGAAGAAGACACUGCAGAGCUGCAAGGCCUCAGA.-3' **WT**

pSVED-A ΔXho-Stu .5'-.UGACCUACUCGA_____CCUCAGA.-3' -

Figure 3.3. Representation of the central part of the EDA exon in the pSVED-A constructs.

In the upper part a schematic representation of the fibronectin insert in the pSVED-A construct indicates the localisation of the studied fibronectin fragment. The sequence of the mRNA region transcribed from the plasmid that comprise the 81 nucleotides Xho I/Stu I fragment is reported. The A underlined in the pSVED-A Tot has been mutated in C in the pSVED-A Tot Sac construct to create the Sac I site. The extent of the original Xho I/Stu I deletion has been reported for the pSVED-A ΔXho-Stu construct. The splicing pattern resulting from the expression of the constructs is reported at the far right (WT=wild type pattern, -=complete EDA exclusion).

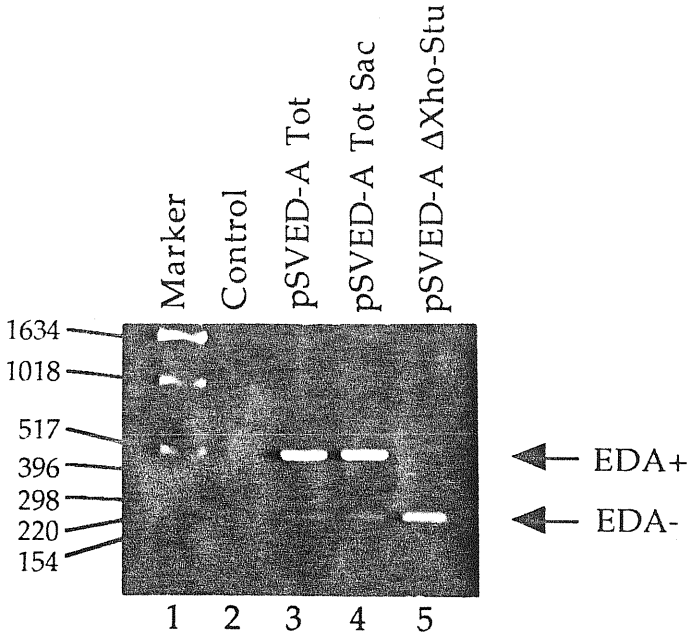


Figure 3.4. RT-PCR analysis of mRNA from HeLa cells transfected with the pSVED-A Tot, Tot Sac and ΔXho-Stu constructs.

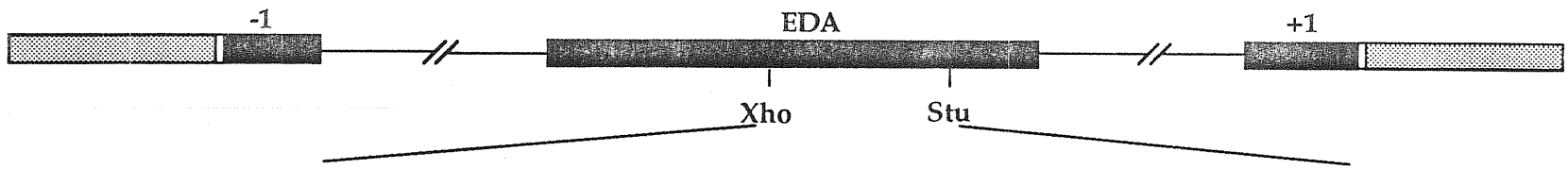
The bands corresponding to the EDA+ (527 bp) and the EDA- (257 bp) products are indicated by arrows. Lane 2 shows a mock transfection control. Lanes 3-5 show analysis of the transfection of the three pSVED-A constructs.

synthesized from messengers originated from the constructs and not from the endogenous gene. The amplification give two different bands, one of 527 bp that corresponds to the EDA+ (EDA included) form, and one of 257 bp that corresponds to the EDA- (EDA excluded) form (Fig. 3.4).

Two versions of the complete sequence, pSVED-A Tot and pSVED-A Tot Sac (Fig. 3.3), produced identical relative proportions of EDA+ and EDA- mRNAs when transfected into HeLa cells (Fig. 3.4, lanes 3,4). This demonstrates that the A to C mutation introduced to create the Sac site was irrelevant regarding the splicing specificity. A Xho I/Stu I fragment was deleted in the centre of the EDA exon (Fig. 3.3). The deleted sequence corresponds to the 81 nucleotides original deletion described by Mardon et al. (1987). The deletion completely abolished alternative splicing confirming thus the previous result (Fig. 3.4 lane 5).

3.2.1 Deletion mutants analysis of sequences involved in EDA splicing

The 81 nucleotides fragment, necessary for the EDA inclusion in the spliced messenger, was dissected with a series of deletion mutants to characterise the sequences involved in this process. The deletion Xho I/Sac I (Fig. 3.5) affects the relative proportion of EDA+ and EDA- mRNA decreasing the relative amount of EDA+ messenger (Fig. 3.6 lane 3). The Sac I/Stu I deletion reproduce the complete abolition of the alternative splicing pattern given by the Xho I/Stu I deletion (Fig. 3.5 and Fig. 3.6 lane 4). The analysis of the effect of deletions obtained with restriction enzymes cannot give a precise picture of the critical sequences. Hence the effect of short 5-10 nucleotide deletions obtained cloning suitable oligonucleotides between the Sac I and Stu I sites was analysed (Fig. 3.5).The Δ1-Sac deletion gives a pattern similar to the wild type (Fig. 3.6 lane 5), while the shorter Δ1 deletion cause a decrease in the EDA inclusion (Fig. 3.6 lane 6). Of particular interest is the Δ2e construct (Fig. 3.6 lane 8) which has the same effect as the extensive Xho I/Stu I deletion causing complete EDA exclusion. The Δ2e deletion eliminates a polypurine sequence (named A element) that, as discussed in section 1.2.1.7, are characterised splicing enhancer. The Δ2c construct is identical to Δ2e but a C has been deleted, so forming, again, a continuous stretch of 5 purines



		Splicing Pattern
	Xho I / Sac I / Element A / Stu I	
pSVED-A Tot Sac	..5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCAAGGCCUCAGA.-3'	**WT**
pSVED-A ΔXho-Sac	.5'-.UGACCUACUCGA_____CUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCAAGGCCUCAGA.-3'	**WT**
pSVED-A ΔSac-Stu	.5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUG_____CCUCAGA.-3'	**-**
pSVED-A Δ1-Sac	...5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUG_____UGGU**GAAGAAGA**CACUGCAGAGCUGCAAGGCCUCAGA.-3'	**WT**
pSVED-A Δ1	.....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUC_____UGGU**GAAGAAGA**CACUGCAGAGCUGCAAGGCCUCAGA.-3'	**WT**
pSVED-A Δ2e	.....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGG_____CACUGCAGAGCUGCAAGGCCUCAGA.-3'	**-**
pSVED-A Δ2C	.....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGG_____CACU**G**_**AGAG**CUGCAAGGCCUCAGA.-3'	**WT**
pSVED-A Δ3	.....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**_____AGCUGCAAGGCCUCAGA.-3'	**WT**

Figure 3.5. pSVED-A deletion mutants.

The extent of the deletions carried out on the pSVED-A constructs to characterise the cis-acting elements involved in EDA splicing are indicated in the transcribed mRNA sequence. The polypurine sequence necessary for EDA inclusion, named A element, is in bold characters. On the far right are reported the splicing pattern resulting from the expression of the constructs. For wild type pattern (WT) is intended each expression pattern in which both forms, EDA+ and EDA-, are present even in different ratios and are distinct from a complete EDA exclusion pattern (-). A putative polypurine enhancer sequence is underlined in the pSVED-A Tot sequence.

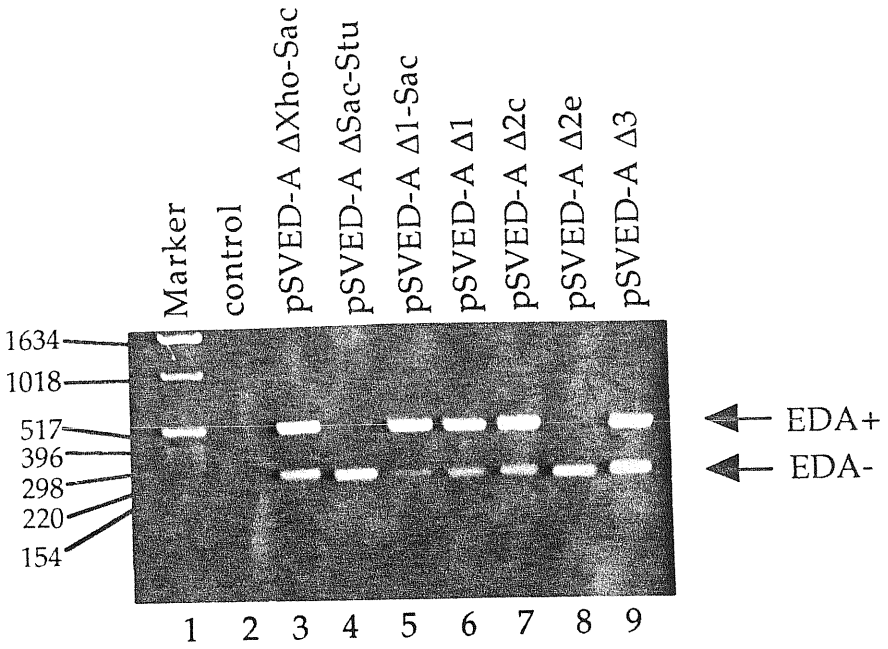


Figure 3.6. RT-PCR analysis of mRNA from HeLa cells transfected with the pSVED-A deletion mutants.

The products of the RT-PCR analysis of the pSVED-A deletion mutants transient expression are shown in lanes 3-9. The bands corresponding to the EDA+ (527 bp) and the EDA- (257 bp) products are indicated by arrows. Lane 2 shows a mock transfection control.

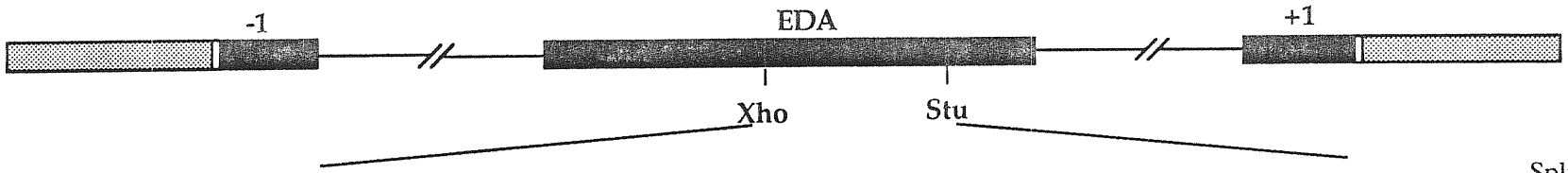
in the same position of the deleted sequence. This results in the restoration of EDA inclusion with a pattern close to the wild type (Fig. 3.6 lane 7). It is clear that a polypurine sequence at that position is essential for the recognition of the EDA exon, while a similar polypurine contained in the Xho I/Sac I fragment (GAGGA) does not seem to have such a striking effect and its deletion causes only an alteration in the EDA recognition but not its complete exclusion. The Δ3 deletion involves the region including the C deleted in Δ2c and produces a decreased EDA inclusion but not abolition of the alternative pattern (Fig. 3.6 lane 9).

3.2.2 Analysis of a non polypurinic sequence involved in EDA recognition

A further pSVED-A deletion mutant Δ4 has a 5 nucleotides deletion (CAAGG) located 13 nucleotides downstream the A element (Fig. 3.7) also abolish the alternative splicing pattern, but, surprisingly, it cause complete EDA inclusion (Fig. 3.8 lane 3). The 5 nucleotides sequence (named B element) is thus essential for EDA alternative splicing regulation and its deletion cause constitutive EDA inclusion. To further characterise this sequence and to investigate on its mechanism of action a fine mutation mapping was carried out with the constructs Δ4B1, Δ4B2, Δ4B3 and Δ4B4 (Fig. 3.7). A complete EDA inclusion was obtained mutating the last guanosine residue of the B element in uracil as in the Δ4B1 construct (Fig. 3.8 lane 4) or with this mutation combined with other nucleotide changes in the B element as in the constructs Δ4B2, Δ4B3 and Δ4B4 (Fig. 3.8 lanes 5,6,7). Unfortunately, since a random mutation approach was follow no other single base mutants are yet available. These results evidence a second regulatory region in the EDA exon sequence that act as a splicing inhibitor. Furthermore the complete abolition of the alternative pattern, given by a single nucleotide change in the messenger sequence, strongly suggest that the B element acts through interaction with trans acting elements and do not alter directly the pre-mRNA secondary structure.

3.2.3 pSVED-A constructs expression in mouse cells

To test if the variations in the human EDA exon splicing pattern caused by the mutations tested were species specific the pSVED-A variants have been



Splicing Pattern

Xho I Sac I Stu I

Element A Element B

pSVED-A Tot Sac ..5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUG**CAAGG**CCUCAGA.-3' **WT**

pSVED-A Δ4 .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUG_____CCUCAGA.-3' **+**

pSVED-A Δ4B1 .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCAAG**U**CCUCAGA.-3' **+**

pSVED-A Δ4B2 .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCAA**UU**CCUCAGA.-3' **+**

pSVED-A Δ4B3 .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCA**C**G**U**CCUCAGA.-3' **+**

pSVED-A Δ4B4 .....5'-.UGACCUACUCGAGCCCUGAGGAUGGAAUCCAUGAGCUCUUCCCUGCACCUGAUGGU**GAAGAAGA**CACUGCAGAGCUGCA**CUU**CCUCAGA.-3' **+**

Figure 3.7. pSVED-A B element mutants.

The extent of the deletions in the pSVED-A Δ4 construct and the point mutations carried out to characterise a second cis-regulatory sequence involved in EDA splicing are indicated in the transcribed mRNA sequence. The mutated bases are underlined in bold characters. The polypurine sequence necessary for EDA inclusion (A element) and the second regulatory sequence (B element) are in bold characters. On the far right are reported the splicing pattern resulting from the expression of the constructs as wild type pattern (WT) or complete EDA inclusion pattern (+).

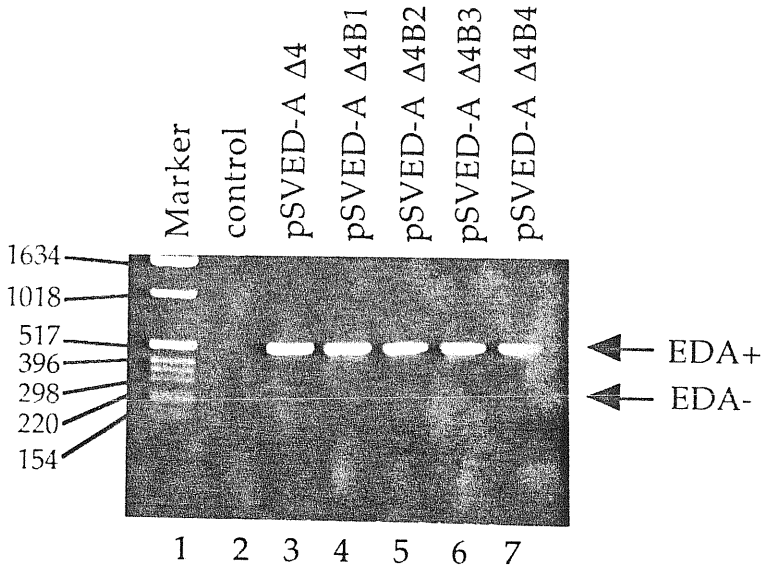


Figure 3.8. RT-PCR analysis of mRNA from HeLa cell transfected with the B element pSVED-A mutants.

The products of the RT-PCR analysis of the pSVED-A B element mutants transient expression are shown in lanes 3-7. The bands corresponding to the EDA+ (527 bp) and the EDA- (257 bp) products are indicated by arrows. Lane 2 shows a mock transfection control.

transfected in mouse NIH3T3 cells. The pSVED-A Tot, that gives the pattern of expression of the endogenous EDA exon, gives the same ratio between EDA+ and EDA- forms in HeLa and NIH3T3 cells (Fig. 3.9 lanes 3,7). The pSVED-A Δ2e and Δ4B1 constructs give respectively complete exclusion and complete inclusion of the EDA exon in both HeLa (Fig. 3.9 lanes 4,5) and NIH3T3 (Fig. 3.9 lanes 8,9). From these results the mechanisms and the factors involved in the EDA regulation do not appear to be species specific.

3.2.4 Compared analysis of the EDA primary sequence in human, rat and mouse

The degree of homology of the primary sequence of the EDA exon among human, rat and mouse is above 90%. As indicated in figure 3.10 12 out of 15 of the not conserved nucleotides are located in the same half of the EDA exon were the A and B elements are also found in the human sequence. Surprisingly neither the sequence of the A element nor the one of the B element are conserved. Furthermore there are only two nucleotides that are not conserved between at least two of the three species analysed. Those nucleotides and are located one in the middle of the A element and one in the middle of the B element. In all the three EDA variants a polypurine stretch of at least five nucleotides is present in the positions of the human A element (GAAGAAGA), GAGGA in rat and GAAGA in mouse. Although the presence of such a polypurine element is essential for the function of a cis acting splicing enhancer its sequence could extend beyond the polypurine and has still to be completely defined. Similarly the precise cis acting sequence involved in the B element is not yet characterised in detail, except for the guanosine nucleotide whose mutation abolish alternative splicing as seen in the construct Δ4B1 expression. The G in this position is conserved among the species analysed.

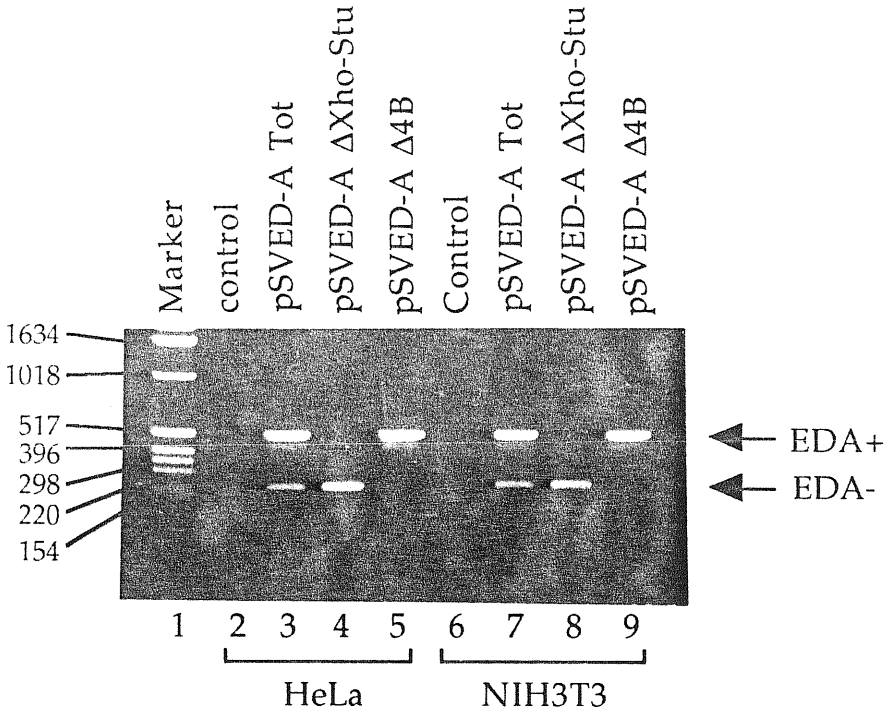


Figure 3.9. pSVED-A constructs expression in mammalian cells.

The pSVED-A constructs containing the wild type and mutated human fibronectin sequence have been transfected in human HeLa cells and mouse NIH3T3 fibroblast. The products of the RT-PCR analysis of the transient expression in HeLa cells is shown in lanes 2-5, the products of the transient expression in NIH3T3 cells in lanes 6-9. Lanes 2 and 6 show mock control transfections. The bands corresponding to the EDA+ (527 bp) and the EDA- (257 bp) products are indicated by arrows.

```
Rat EDA      1..CCATTAATTTGCCTAACAGACATTGAcCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATTCCATCAAAATTGCcTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTACAGGGT
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Mouse EDA    1..CCATTAATTTGCCTAACAGACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATTCCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTACAGGGT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Human EDA    1..CCATTAATTTGCCTAACAGACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATTCCATCAAAATTGCTTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTACAGGGT

consensus    ...CCATTAATTTGCCTAACAGACATTGAtCGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATTCCATCAAAATTGCtTGGGAAAGCCCACAGGGGCAAGTTTCCAGGTACAGGGT

Rat EDA      123 GACCTACTCaAGCCCTGAGGATGGAATCCatGAGCTTTTCCCTGCgCCTGATGGTGAcGAgGACACgGCAGAGCTGCAcGGCCTCAGGCCGGGtTCTGAGTACACAGTCAGTGTGGTT
             ||||||||| |||||||||||||||||||  |||||||||||||| |||||||||||| || ||||| ||||||||||| |||||||||||||| |||||||||||||||||||||||||
Mouse EDA    123 GACCTACTCGAGCCCTGAGGATGGAATCCggGAGCTTTTCCCTGCACCTGATGGTGAAGAcGACACTGCAGAGCTGCAgGGCCTCAGGCCGGGgTCTGAGTACACAGTCAGTGTGGTT
             |||||||||||||||||||||||||||||  ||||| |||||||||||||||||||||||| |||||||||||||||| |||||||| ||||| |||||||||||||||||||||||||
Human EDA    123 GACCTACTCGAGCCCTGAGGATGGAATCCatGAGCTaTTCCCTGCACCTGATGGTGAAGAaGACACTGCAGAGCTGCAaGGCCTCAGaCCGGGtTCTGAGTACACAGTCAGTGTGGTT

consensus        GACCTACTCgAGCCCTGAGGATGGAATCCatGAGCTtTTCCCTGCaCCTGATGGTGAaGA-GACACtGCAGAGCTGCA-GGCCTCAGgCCGGGtTCTGAGTACACAGTCAGTGTGGTT
                                                                          Element A          Element B

Rat EDA      245 GCCTTGCACGgTGgcATGGAGAGCCAGCCCCTGATTGGAgTCCAGTCCACAGGTATATCGgT
             |||||||||| ||  |||||||||||||||||||||||| |||||||||||||||||||| |
Mouse EDA    245 GCCTTGCACGATGATATGGAGAGCCAGCCCCTGATTGGAATCCAGTCCACAGGTATATCGTT
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Human EDA    245 GCCTTGCACGATGATATGGAGAGCCAGCCCCTGATTGGAAcCCAGTCCACAGGTATATgGTT

consensus        GCCTTGCACGaTGatATGGAGAGCCAGCCCCTGATTGGAatCCAGTCCACAGGTATATcGtT
```

Figure 3.10. Comparison of human, mouse and rat EDA exon sequences.
Sequences for the human, rat and mouse EDA exon plus boundaries are given. Exonic sequences are underlined while intronic sequences are not. The A and B element are indicated in bold characters. Upper case and bars indicate nucleotide homology. Consensus sequences is given in the lower line.

3.3 INSERTION OF THE B ELEMENT SEQUENCE IN A CONSTITUTIVE α GLOBIN EXON

To test the possibility that a FN fragment that comprise the B element acts as a general splicing inhibitor we inserted a sequence of 21 bp centred around the B element in the second exon of the α globin gene contained in the vector pSVαBstINS (as described in section 2.2.12.4). The α globin second exon is constitutively included in both endogenous and pSVαBstINS plasmid messengers. To detect the messengers transcribed from the mutated vectors after transfection in HeLa cells and to discriminate from the endogenous α globin messengers a PCR based method has been set up. cDNA reverse transcribed from the transfected cells was amplified using the primers SpeINS3′ and αglobE15′ (Table 2.1 and Fig. 3.11). Those primers can amplify only the messengers transcribed from the pSVαBstINS plasmid since, as indicated in figure 3.11, the αglobE15′ primer anneal with the α globin sequence and the SpeINS3′ primer with the insert in the third α globin exon.

3.3.1 Expression of the pSVαΔ4 insertion mutants

A series of mutant plasmids was created inserting in the Pml I restriction site of the second exon of the α globin gene the FN 21 bp fragment indicated in figure 3.11 in both orientation. A FN fragment carrying the G to T mutation in the B element that cause complete exon skipping of the FN EDA exon (pSVED-AΔ4B1 construct, Fig. 3.8 lane 4) was also inserted in the Δ4B plasmids in the right (pSVαΔ4BS) and reverse orientation (pSVαΔ4BAS) as control. The expression of these constructs (Fig. 3.12 lanes 2,3) gave complete inclusion of the second α globin exon. In the Δ4+ constructs the inserted fragment bear the wild type FN sequence in single copy in the right orientation (pSVαΔ4+S), repeated in tandem (pSVαΔ4+2cS) and in reverse orientation (pSVαΔ4+AS). The expression of all the three plasmids gave complete α globin exon 2 inclusion (Fig. 3.12 lanes 4,5,6). These results indicates that a 21 nucleotides sequence centred around the B element is not able per se to produce alternative skipping in constitutively included exons.

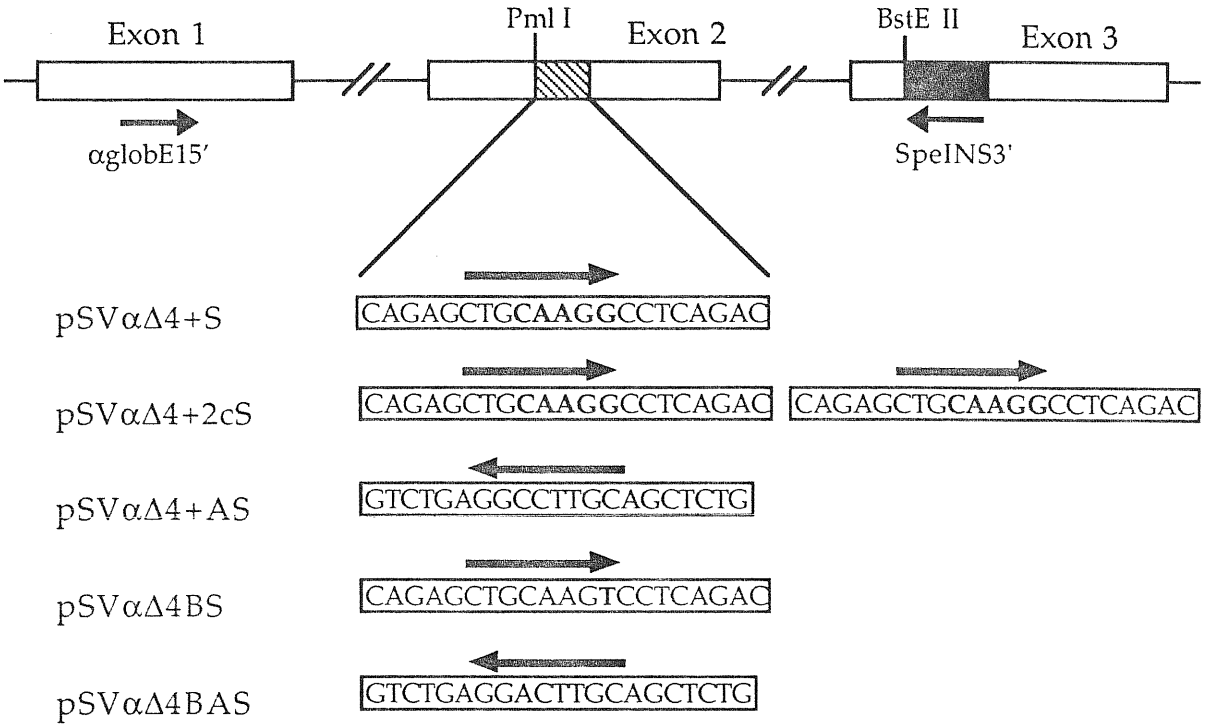


Figure 3.11. Graphic representation of the pSVαΔ4 constructs.

In the upper part is represented the α globin region of the pSVαΔ4 constructs. Black box indicates the sequence inserted in the BstE II site. Dashed box indicates the fibronectin sequences inserted in the second exon. Locations of the primers SpeINS3′ and αglobE15′ are indicated. The sequences of the insertions are indicated together with their orientation respect the original fibronectin sequence. The CAAGG sequence that characterise the B element is in bold character. The mutated T in the pSVαΔ4BS construct is also in bold.

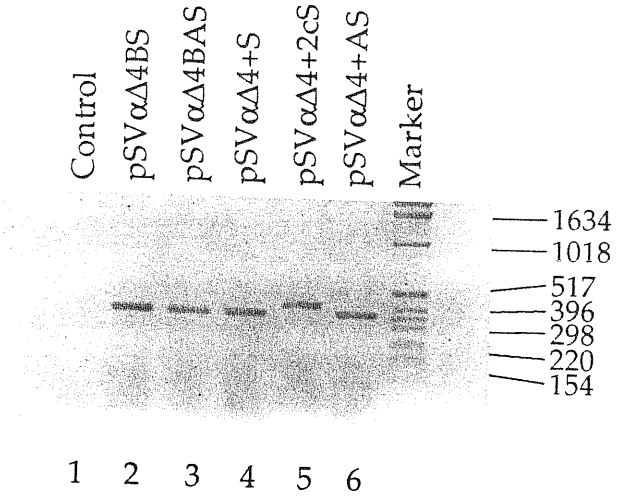


Figure 3.12. RT-PCR analysis of the pSVαΔ4 constructs expression.

The products of the RT-PCR analysis of the pSVα Δ4 constructs transient expression in HeLa cells are shown in lanes 2-6. Lane 1 shows a mock transfection control. Only one product is visible in each lane this indicates exon 2 constitutive inclusion.

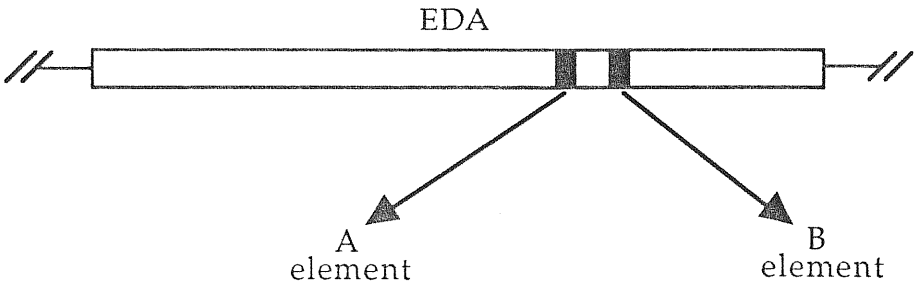
3.4 GEL RETARDATION ASSAYS

To address the possibility that the B element and surrounding sequence might interact with trans-acting factors to mediate exclusion of the EDA exon, gel retardation assays were carried out using HeLa total nuclear extract and 5′ labelled oligoribonucleotides as probes. The oligoribonucleotide Δ4+ is a 21 nucleotides of wild type FN sequence that include the B element and the surrounding sequence (Fig. 3.13), it was used to detect specific protein binding and as a cold competitor. Oligoribonucleotide Δ4- contains the same sequence of Δ4+ except for the deletion of the B element (Fig. 3.13) and was used as control for specific protein binding and as aspecific competitor. The reactions were performed with different concentration of cold competitors and in presence or absence of ATP.

3.4.1 Trans-acting factors bind the B element

Several retarded complexes were observed when the Δ4+ probe was incubated with total HeLa nuclear extracts in absence or presence of ATP (Fig. 3.14 lanes 2,3). Retarded complexes were also obtained after incubation of the Δ4- probe with HeLa extracts in presence or absence of ATP (Fig. 3.14 lanes 5,6). However a low mobility complex, named A complex, was obtained only with the Δ4+ probe (Fig. 3.14 lanes 2,3). Interestingly in presence of ATP no particular differences are detected in the retarded complexes given by both probes except an overall increase in the RNA:protein binding efficiency.

To test the specificity of the A complex competition experiments, using both, the Δ4+ and the Δ4- oligoribonucleotides, as cold competitor, was performed. Only the Δ4+ competitor was able to completely abolish A complex formation (Fig. 3.14 lanes 7,8). Conversely the Δ4- competitor was not able to consistently abolish A complex formation at a molar excess up to 300 fold respect the Δ4+ probe (Fig. 3.14 lanes 9,10,11). These results indicates that nuclear factors bind specifically an EDA sequence that contains the B element, those factors could thus be implicated in the alternative splicing regulation of the exon.



CUGCACCTGAUGGU**GAAGAAGA**CACUGCAGAGCUG**CAAGG**CCUCAGACCGGG

probe Δ4+ CAGAGCUG**CAAGG**CCUCAGAC

probe Δ4- CAGAGCUG____CCUCAGAC

Figure 3.13. Oligoribonucleotide probes used in the gel retardation assays.
Sequence and localisation of the oligoribonucleotides used in the gel retardation assays as probes and cold competitor are reported. The B element sequence is indicated in bold characters in the Δ4+ oligoribonucleotide. The Δ4- oligoribonucleotide is missing the B element sequence as indicated.

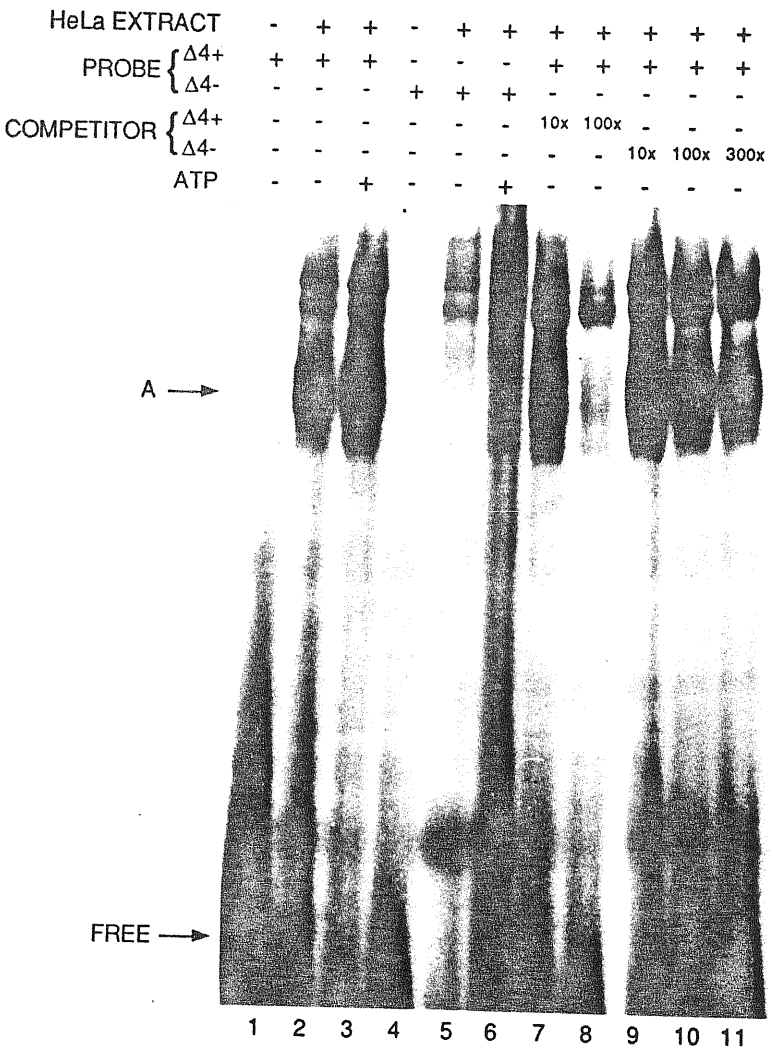


Figure 3.14. Gel retardation assays of binding of the Δ4+ and Δ4- oligoribonucleotides.

Protein extracts were incubated with the ^{32}P-labelled Δ4+ and Δ4- oligoribonucleotides and subjected to electrophoresis on non denaturing gels. Competition were carried out with the indicated molar excess of cold Δ4+ and Δ4- oligoribonucleotides. The A complex is indicated. ATP is added at a final concentration of 0.5 mM.

3.5 DISCUSSION

The FN alternatively spliced exon EDA was originally reported to contain sequence that participate in splice site selection. In fact a deletion of an internal 81 nucleotides fragment abolished EDA exon inclusion in HeLa cells (Mardon et al., 1987). This was the first of a family of exonic sequences necessary for exon recognition and correct splicing to occur (see sections 1.2.1.6 and 1.2.1.7). In this study this observation was confirmed and extended considerably in in-vivo systems. First a short polypurine stretch (GAAGAAGA) contained in the original 81 nucleotides fragment is responsible for EDA exon inclusion, this sequence was named A element. In addition, deletion of a further short sequence (CAAGG) contained in the same 81 nucleotides fragment altered the original alternative splicing pattern causing constitutive EDA inclusion, this was named B element. Gel retardation assays showed specific complex formation when ^{32}P labelled probes containing the B element were incubated with HeLa nuclear extracts.

3.5.1 Poor splice sites do not allow EDA exon constitutive splicing

The EDA exon is one of the three FN exons that encode for a full type III homology element (see section 1.5.1). Two of these exons, EDB and EDA itself, are alternatively spliced in a variety of tissues and physiological situations (see chapters 4 and 5), while the third exon, named EIII9, is constitutively included. The analysis of the primary sequences of those exons showed that all of them have poor consensus 5′ and 3′ splice sites or are placed in a weak splicing contest with, as for EDB, a weak upstream 5′ splice site. Previous results indicated that the rat EIIIB exon could be constitutively spliced if its 5′ and 3′ splice sites and the 5′ splice site of the upstream exon were matching the consensus sequence. The comparison of the splice sites flanking the EDA exon with the consensus sequences showed that the 3' splice site upstream of the human EDA exon presents a polypyrimidine stretch interrupted by purines, that probably lower its affinity for U2AF. Analysis of the 5′ splice site flanking the EDA exon also show an extremely poor match to the consensus, particularly substitution of the consensus G in position +5 is known to produce exon skipping if it occasionally occurs in

constitutively spliced exons (Talerico et al., 1990; Nelson and Green, 1990). Moreover, the 5' splice junction downstream of the EDA exon has a structure similar to some disease-causing mutations that produce exon skipping in β thalassemia (Krainer et al., 1984) and Lesch-Nyhan syndrome (Gibbs et al., 1990). Recent experiments, carried out in our laboratory, confirm these observations showing that mutated EDA sequences matching the 3′ or the 5′ splice site to the consensus lead to constitutive EDA inclusion (Muro A., personal communication). The EDA exon is, thus, flanked by extremely poor splice sites that should not be recognised by the splicing machinery and needs additional cis-acting signals that enhance its recognition.

3.5.2 The exonic enhancer is a polypurine

Analysis of expression of the pSVED-A deletion mutants shows that a polypurine stretch (GAAGAAGA), named element A, is essential for EDA exon recognition and its removal produces omission of the EDA exon. The A element is a member of the exon splicing enhancer sequences (ESE) discussed in section X. It is interesting to note that the splicing enhancing activity exerted by this ESE could not only be a consequence of nucleotide sequence but a positional effect may also be operative. In fact, the polypurine sequence GAGGA, located 35 nucleotides upstream the A element (Fig. 3.5) does not have the effect of the polypurine stretch of element A, as can be seen when comparing the patterns produced by the ΔXho/Sac and ΔSac/Stu constructs (Fig. 3.6 lanes 3,4). Consistent with this interpretation, the Δ2c deletion reconstitutes a polypurine stretch with a different sequence but in the same position of element A and restores a transcription pattern similar to that of the wild type (Fig. 3.6 lane 7). Constructs ΔXho/Sac, Δ2c, Δ3, Δ1 and Δ1Sac do not completely abolish the EDA inclusion but alter, with various extent, the original splicing pattern. This clearly indicates that the complete sequence needed for the EDA exon alternative splicing is far more complex than the polypurine stretch. Additive nucleotides located upstream and downstream from the characterised A element sequence could be involved together with their relative position, as suggested by the Δ2c construct, in the complex mechanism underlying EDA regulation.

3.5.3 Trans-acting factors regulating EDA inclusion

Lavigueur et al. (1993) have recentky shown, using an in vitro approach, that the EDA polypurine GAAGAAGA sequence was able to enhance intron removal in heterologous constructs stimulating the usage of the 3′ splice site flanking the exon in which the enhancer sequence was placed. The ability to enhance the use of the upstream 3′ splice site was associated with more efficient A complex formation and increased U2 snRNP binding at the branch site. The same workers also show that SR proteins interact specifically with the EDA ESE sequence, mediating the stimulation of splicing. This probably happens, as described in section 1.4.1.2, throughout a network of RNA:protein and protein:protein interactions, involving direct ESE:SR protein interaction that, in turn, promotes assembly of an E-like complex on the ESE sequence containing: SR proteins, U2AF and U1 snRNP. This complex promotes the U2AF:polypyrimidine tract interaction substituting in this way the presence of consensus splice sites.

Lavigueur et al. (1993) reported other two observations that are confirmed by the results obtained in this work. First a positional effect was found to influence the EDA ESE sequence splicing stimulation. Increased distance between the ESE sequence and the 3′ splice site lowered the rate of 3′ splice site usage. Second a weak downstream 5′ splice site is necessary for ESE function. In vitro experiments showed that EDA ESE requirement could be overcome by a consensus downstream 5′ splice site. These evidences confirms that the EDA splicing pattern is obtained through a combination of weak splice sites, enhancer sequences and the relative positions of these elements.

Other nucleotides not constituting a polypurine stretch can be involved in the sequence recognised by the SR proteins as discussed in section 1.4.1.2 and confirmed by deletion mutants that influence the EDA splicing pattern without altering the A element (Fig. 3.5 and 3.6). Since the EDA exon is highly regulated in a variety of physiological situations it is likely that different SR proteins are required for different responses. It will be, thus, of particular interest the complete characterisation of the sequence involved in SR protein binding. This could also led to the identification of different sequences bound from different subset of SR proteins. In this view the A element can be considered as a general SR protein

binding site feature with other sequences conferring the binding specificity. This hypothesis is confirmed by the results obtained with the pSVED-A constructs expression in mouse NIH3T3 cells and the comparative analysis of the human, rat and mouse EDA exon sequences. Expression in mouse NIH3T3 of the pSVED-A constructs bearing the wild type human EDA sequence and the deletion of the A element give the endogenous mouse EDA splicing pattern and complete EDA exclusion, respectively, as observed in human HeLa cells. This indicates that the mechanism that regulates EDA inclusion is not species specific and the trans acting factors involved are highly conserved among species and should recognise the same binding sequences. SR proteins are known to be highly conserved during evolution but little is known on their binding specificity. Interestingly a comparative analysis of the primary sequences of the human, rat and mouse EDA exon show that despite a degree of conservation above 90% among the three species the A element sequence is not well conserved being reduced to a stretch of 5 purines both in rat (GAGGA) and mouse (GAAGA). Therefore, in mice, the same splicing factors that regulate with accurate precision the inclusion of the human EDA exon can recognise endogenous sequences that differ from the human in the essential polypurine tract. It is likely that a general polypurine sequence placed in the A element position can enhance SR protein binding while the precise extent and specificity of this binding is given by other sequences.

3.5.4 Identification of a negative regulator of splicing: the B element

In this study was shown that the EDA exon system has an added complexity that contributes to its regulation. Weak splice sites and polypurine enhancer are not the only cis-acting elements regulating EDA exon splicing. Removal of element B a CAAGG sequence placed 13 nucleotides downstream of element A, results in 100% inclusion of the EDA exon (Fig. 3.8). Thus a repressor sequence, named B element, seems to be also involved in EDA splicing. This is not surprising, since the control of EDA alternative splicing and its tissue specific and developmental modulation would be difficult to achieve by a one way system.

A bipartite element constituted by a enhancer and a silencer of splicing separated by short spacer sequence seem to regulate the complex FN EDA exon

splicing. Similar polypurine-spacer-B element have been recently observed in alternatively spliced exons of the bovine growth hormone (Sun et al., 1995) Rous sarcoma virus (McNally et al., 1995) and bovine papilloma virus (Zheng et al., 1995). Except for the bovine growth hormone, were the last guanosine is substituted by an adenosine, the CAAGG sequence appear to be a common determinant in the silencer sequence.

Recently negative and positive splicing regulatory elements have been identified within the terminal tat-rev exon of the human immunodeficiency virus type 1 (Amendt et al., 1995; Staffa and Cochrane, 1995). The positive acting sequence is a polypurine sequence while the negative sequence, AGATCC, has not apparent homology with the FN B element sequence and it is also present in the tat exon 2 with the same inhibitory activity. The alternative splicing of the K-SAM alternative exon of the fibroblast growth factor receptor 2 is also regulated by a combination of positive and negative splicing regulators (Del Gatto and Breathnach, 1995). The short K-SAM exon sequence TAGGGCAGGC inhibits splicing of the exon while two separate sequence elements in the intron immediately downstream are both needed for exon inclusion. As for the FN EDA three elements are necessary in controlling splicing of the tat-rev and K-SAM alternative exons, suboptimal 5′ and 3′ splice sites, exonic inhibitory sequences, a non polypurinic intronic activating sequence, for the K-SAM exon, and a polypurinic exonic enhancer for the tat-rev exon. Thus the arrangement of contiguous positive and negative elements may be common to a number of alternatively spliced genes.

3.5.5 Mechanism of action of the bipartite splicing enhancer

While much recent work has described exonic splicing enhancer, relatively few examples of exonic splicing-inhibitory sequences are known. One of the best characterised example of the latter is the exon inhibitory sequence close to the 3′ splice site of the β-tropomyosin alternative exon 6B, which functions by forming a secondary structure that "hides" the 3′ splice site. A similar mechanism is not likely to be involved in the B element action. The EDA exon inhibitory sequence is at about 170 nucleotides from the 3′ splice site. Furthermore the G to T mutation

carried out in the construct Δ4B1 (Fig. 3.7) is sufficient to obtain complete exon inclusion. It is therefore unlikely that this splicing suppressor acts directly throughout the pre-mRNA secondary structure but rather by interaction with trans-acting factors.

Few examples of sequences within exons that bind factors which negatively regulate splicing have been reported. In section 1.2.1.6 has been exposed the example of regulation of the drosophila P-element transposase, in which somatic factors bind to the exon upstream of the 5′ splice. These factors act to block U1 snRNP binding by stabilising U1 snRNP binding to an inactive pseudo-5′ splice site. The drosophila P-element is an example of simple regulation of the alternative splice mechanism since it acts with an on-off mechanism that allow correct splicing in germinal cells and inhibit splicing in somatic cells. This is not the case of the EDA and other highly regulated exons that may require more elements that act together as seems to be the case of the terminal tat-rev exon of the human immunodeficiency virus type 1. In this case both enhancer and suppressor sequences share similar properties, both elements require suboptimal splice sites, both types of elements can be transferred to heterologous exons and both elements bind cellular factors. The enhancer is a polypurine ESE that is supposed to bind SR proteins while the silencer sequence is specifically bound by at least one, not yet identified, nuclear protein.

The preliminary data obtained by gel retardation assays reported in this thesis show that cellular factors selectively bind a 21 nucleotides RNA fragment containing the B element and the flanking bases. Weather the isolated complex is a single protein or a multiprotein complex is not known, however it is likely that it corresponds to the inhibitory activity exerted by the B element. This indicates that nuclear factors act on the A and B element to specifically regulate the EDA expression. The precise identity and the modality of action of these factors are still obscure, except for the interaction observed between members of the SR protein family and the polypurine enhancer. However some experimental observations could help elucidate this mechanism. The element B seems to be recessive to the element A since simultaneous deletion of both (pSVED-A ΔSac-Stu construct) results in complete EDA exclusion. This strongly indicates that the function of the

B element is subordinated to the functions of the enhancer sequence. Thus it is unlikely that the silencer element acts per se as a general splicing repressor, rather its activity is connected with the ESE ability to stimulate exon recognition and spliceosome assembly at weak splice sites. This hypothesis is confirmed by the results given by the insertion, in the second exon of the α-1 globin gene, of the 21 nucleotides FN sequence used as oligoribonucleotide probe in the band shift assay. The same sequence, centred around the B element, that showed specific binding properties for nuclear factors was not able to repress, in any extent, the recognition and inclusion of a constitutive exon that do not bear known splicing enhancer sequences and is flanked by consensus splicing sites.

The complete sequence of the splicing suppressor needs to be further characterised. Transfection of the pSVED-A Δ4B constructs, bearing the human EDA sequence mutated at the B element, in mouse NIH3T3 cells showed complete inclusion of the EDA exon as observed in human HeLa cells. This indicates that the factors involved in this mechanism are conserved throughout evolution and recognise sequences that are maintained similar, at least in their secondary structure, in the different species. However the compared analysis of the primary sequences in rat, mouse and human showed not complete homology in the human CAAGG sequence with the second adenosine residue mutated in cytosine in rat and guanosine in mouse. Furthermore the last guanosine residue seems to be essential but it may be just neede a purine as in the negative splicing element localised in the bovine growth hormone alternatively spliced exon has been characterised the sequence CAAGA. As for the polypurine enhancer, that characterise the A element, the B element CAAGG sequence is probably only part of a complex sequence that is necessary to specifically modulate the interactions with trans acting factors. The complete silencer sequence could overlap or be part of the nearby enhancer element.

These observations suggest the working model proposed in figure 3.15. In this model weak splice sites do not allow an efficient recognition of the EDA exon, this can be overcame by the presence of a splicing enhancer that acts binding SR proteins, that in turn facilitate the assembly of the early spliceosomal complex at the 3′ splice site and the recognition of the exon by the splicing machinery. The

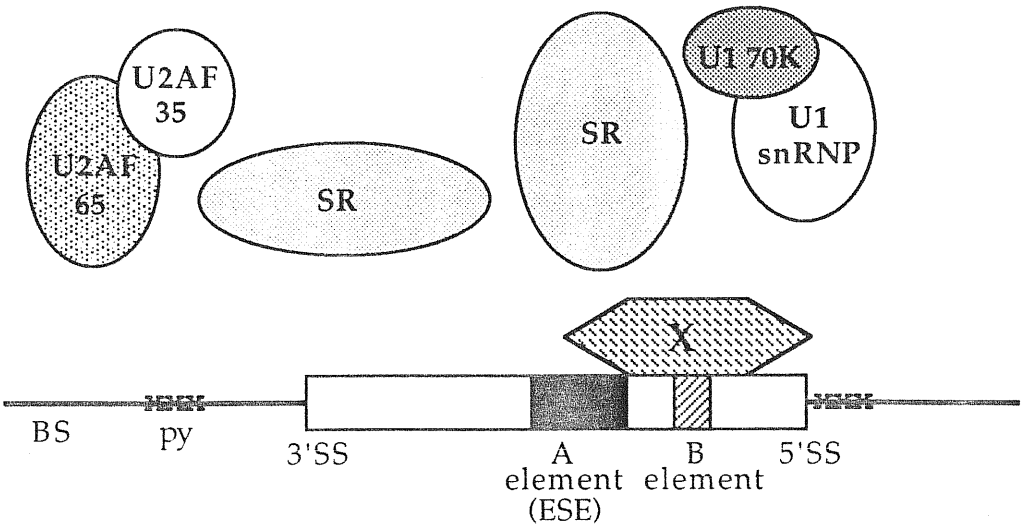


Figure 3.15. Working model for the EDA exon bipartite splicing regulator.
The A element polypurine sequence (black box) and nearby sequences are recognised by SR proteins, that facilitate recognition of the exon and assembly of the early spliceosomal complex interacting with $U2AF^{35}$ and U1 70K. This overcome the pour recognition of weak splice sites. Trans-acting factors (termed X) interact with sequences that comprise the B element (dashed box). This inhibit SR protein binding and, in turn, efficient exon recognition and spliceosome assembly.

interaction between the enhancer and the SR proteins can be inhibited by the presence of trans-acting factors binding the silencer sequence. This inhibition could be due to steric hindrance or to direct competition for common binding sequences.

Since the proximity of the B element with the polypurine enhancer and the scarce information available on the binding sequences recognised by the single SR proteins it is plausible that both, factors enhancing and factors inhibiting EDA exon recognition, are member of the SR proteins family. A mechanism similar to the one controlling the large T and small t SV40 alternative processing can be evoked (see section 1.4). SRp30b promote selection of a proximal 5′ splice site and thus production of the small t messenger. SRp40 promote selection of a distal 5′ splice site and large T messenger production. In a similar way two different SR proteins can bind the central region of the EDA exon with different sequence specificity. Only one of the two is able to promote correct EDA splicing while the other, that has a major binding determinant on the B element, is not active in promoting EDA inclusion and compete with the former for common binding sequences (e.g. the A element polypurine).

The proposed mechanism of actions involving trans-acting factors that act on cis-acting elements organised in nearby stimulatory and inhibitory sequences resemble the organisation and regulation of many transcriptional promoter. The relative level of expression or activation (e.g. given by phosphorylation or dephosphorylation) of the positive and negative splicing factors involved may regulate the expression of the EDA exon in response to many different physiological stimuli. This is similar to the final events of many transduction signal pathway that cause stimulation and inhibition of transcription factors. Splicing is an additive mechanism of control of gene expression and its regulation is probably controlled by complex signal pathway overlapping, perhaps, the ones regulating transcription. The next two chapters will give examples of the complex regulation reached by the splicing machinery in the processing of the FN messenger in response of different physiological situations.

Chapter 4

Fibronectin pre-mRNA Processing in Rat Tissues

The pattern of FN pre mRNA splicing of the EIIIA, EIIIB and V regions in different rat tissues during development and ageing have recently be analysed using an RNAase protection technique to discriminate the variants (Pagani et al., 1991). A similar study was carried out afterwards by Magnuson et al (1991) using PCR methodology. The data collected in these studies were not in complete agreement since different ratios among the isoforms of a given tissue were found by both groups. Furthermore none of these studies investigated on the coordination of splicing between the three different spliced sites of the molecule. Except for occasional information obtained by sequencing different isolated cDNAs, covering in one molecule the EIIIA, EIIIB and V region no systematic study has been carried out on the linkage of the three variants. This problem was addressed in this study and the quantitative data was reanalysed to explain the disagreement. We have focused our research on brain and kidney tissues. Furthermore we have tested the possibility that the type III homology encoded by the EIII9 exon could undergo alternative splicing. The EIII9 is the only other type III homology encoded by a single exon as the alternatively spliced type III domains EIIIA and EIIIB. All the other type III elements are encoded by two exons.

4.1 ALTERNATIVE SPLICING PATTERNS AND LINKAGE BETWEEN THE SPLICED SITES

To perform the analysis of the alternative splicing pattern of the FN messenger and to analyse the linkage between the different splicing sites a "long range RT-PCR" (1000-1500 bp) was developed in 1991. Total RNA was extracted from the tissues of 6 and 24 months old rats and reverse transcribed (RT). PCR reactions were performed using the cDNA synthesised by the RT reaction as described (see 2.2.9.2). Locations of the primers on the FN molecule, primers sequence and the predicted segment length of each amplification product are shown in figure 4.1 and table 4.1. The number of PCR amplification cycles was kept below 24 in order to maintain the level of amplification within the linear range of the technique, and to avoid unspecific products as already described (Wang et al., 1989). The main pitfall in quantitative PCR studies is the variation in the efficiencies of different primer pairs. To avoid this problem the ratios of alternative splicing products generated by the same primers within the same reaction was compared. The amplified fragments were of optimal length, in such a way that the difference between two products was not greater than 40% (e.g. 1000 bp and 1400 bp). Under these conditions quantitative estimation of 2 or 3 types of mRNA were possible.

4.1.2 Analysis of the EIIIA-V linkage

In order to investigate the correlation between the expression of the EIIIA site and the V region on each cDNA sample two distinct reactions were carried out, the first was done with the primers PREIIIA-, V3' and the second with the primers PREIIIA+, V3' (Fig. 4.1, Tab. 2.1 and 4.1). The first reaction gave the ratio among the messengers lacking the EIIIA exon (EIIIA-) and carrying the three different forms of the V site (V0, V95 and V120). The second reaction amplified the messengers including the EIIIA exon (EIIIA+) coupled with the three different V forms. Figures 4.2 and 4.3 shows the linkage between the EIIIA and V regions in rat brain and kidney in relation to ageing. No changes in the expression pattern of the three V variants between both the EIIIA+ and the EIIIA- messengers was

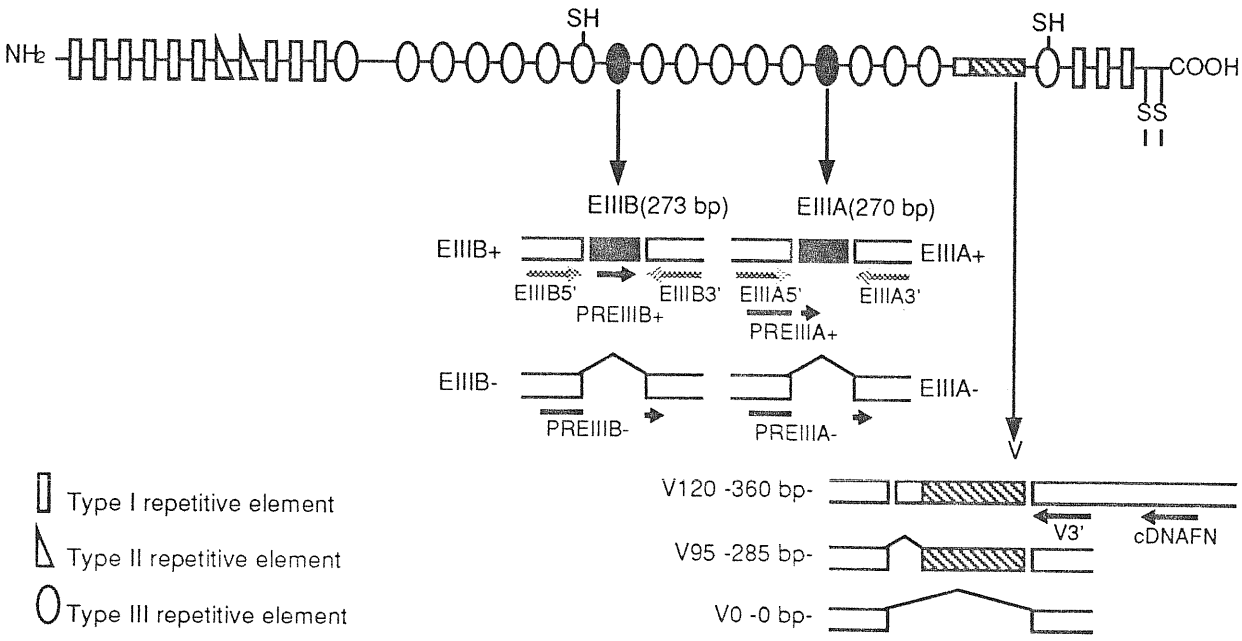


Figure 4.1. Schematic map of a subunit of rat FN molecule.
Type I, II and III repeats are represented as rectangles, triangles and ovals respectively. Alternatively spliced regions (EIIIA, EIIIB, and V) are indicated together with their pattern of splicing and the primer localisation.

Primer Pairs	Products	See Figure
EIIIB5' (Sense)	787 bp (EIIIB+)	4.5
EIIIB3' (Antisense)	617 bp (EIIIB-)	5.4.A
PREIIIB- (Sense)	1485 bp (EIIIB-, EIIIA+)	4.6
EIIIA3' (Antisense)	1212 bp (EIIIB-, EIIIA-)	
PREIIIB+ (Sense)	1540 bp (EIIIB+, EIIIA+)	4.6
EIIIA3' (Antisense)	1267 bp (EIIIB+, EIIIA-)	
EIIIA5' (Sense)	779 bp (EIIIA+)	4.4
EIIIA3' (Antisense)	506 bp (EIIIA-)	5.4.B
PREIIIA- (Sense)	1212 bp (EIIIA-, V120)	4.2
V3' (Antisense)	1137 bp (EIIIA-, V95)	4.3
	852 bp (EIIIA-, V0)	
PREIIIA+ (Sense)	1485 bp (EIIIA+, V120)	4.2
V3' (Antisense)	1410 bp (EIIIA+, V95)	4.3
	1125 bp (EIIIA+, V0)	

Table 4.1. Primers and products obtained. For each single reaction are indicated the pair of primers used and the products obtained (size and isoform represented).

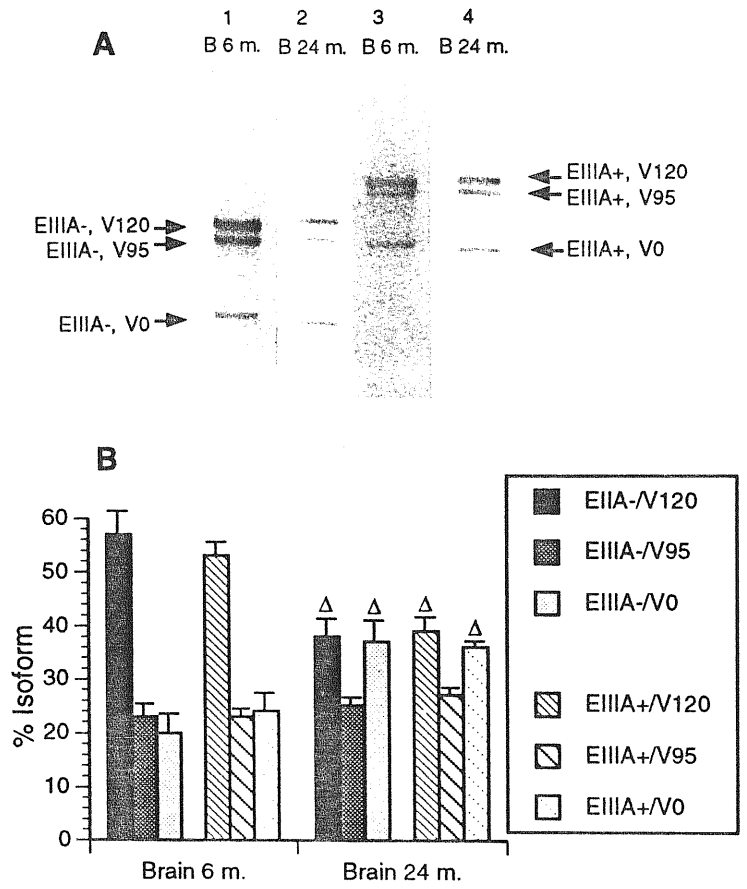


Figure 4.2. Linkage between EIIIA and V regions in brain.

A) In lanes 1 and 3 a 6 months old rat brain was analysed using respectively primers PREIIIA-, V3' and PREIIIA+, V3'. The same analysis was performed for the 24 months old rats (lanes 2,4).

B) The histogram summarise the results of the linkage. The relative ratios of the bands obtained have been normalised for their relative content in (α-^{32}P)-dCTP. Statistically significant differences between 6 and 24 months old animals are noted (Δ). There is a reduction of the V120 and an increase of the V0 isoforms in the 24 months old rat. No significant changes occurs between the expression patterns of the EIIIA- and the EIIIA+ messengers in the same age group.

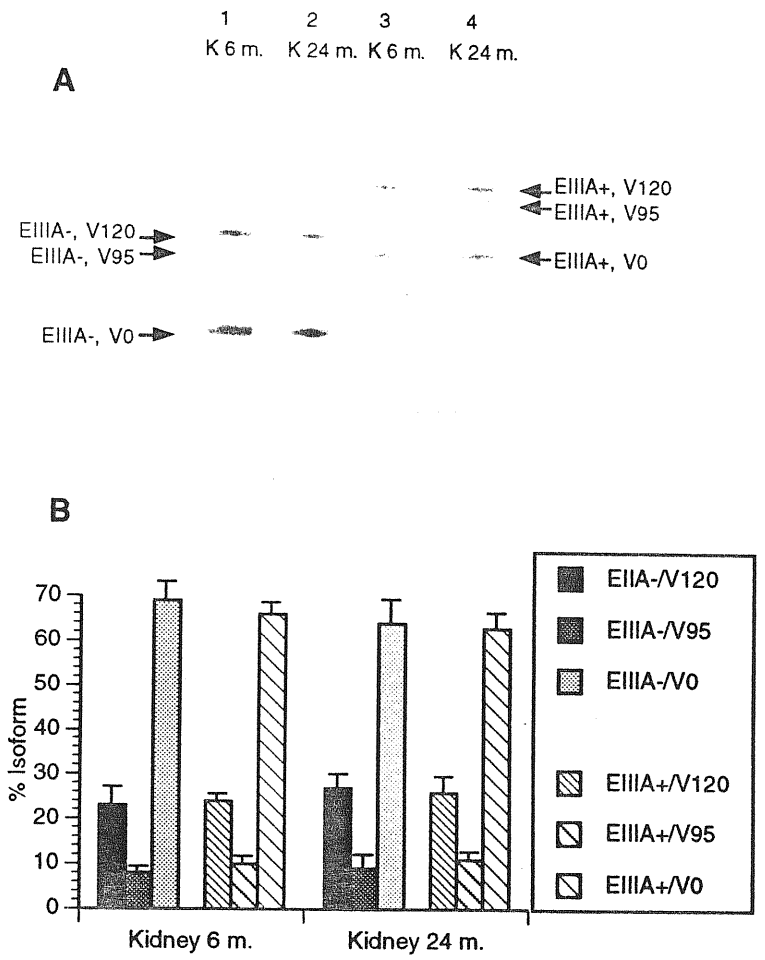


Figure 4.3. Linkage between the EIIIA and V sites in kidney.

A) Lanes 1 and 3 represent the results of the analysis carried out on the kidney of a 6 months old rat, whereas lines 2 and 4 represent the results of the analysis of the kidney of a 24 months old rat. Primers PREIIIA-, V3' were used to obtain the products of line 1 and 2 and primers PREIIIA+, V3' to obtain the products of line 3 and 4.

B) The histogram show the results of the linkage. In kidney, there's an inversion of the pattern with respect to the cerebral tissue concomitantly with a predominance of the V0 isoforms. No significant changes between the expression pattern of the EIIIA- and the EIIIA+ isoforms are detected in the same age group, nor when young and old animals are compared.

detected, this imply that, at least in these tissues, a co-ordinated expression of the V and EIIIA regions does not seem to occur.

4.1.3 Analysis of the EIIIA, EIIIB, V regions

The brain tissue shows a decrease of the V120 form in the 24 months old rats (39%) with respect to the 6 months old animals (57%) (Fig. 4.2). Conversely no change was detected in the ratio of the V forms in the kidney between the 6 and the 24 months old rats (Fig. 4.3). Furthermore consistent differences were found in the pattern of splicing between kidney and brain. In the former there is a prevalence of the V0 form in both 6 and 24 months old rats with the V95 form being the least represented (Fig. 4.2).

To quantify the ratio between the EIIIA+ and EIIIA- messengers a segment including this site was amplified using the primer EIIIA3' and EIIIA5'. The amplification gave two products, one including this region (1450 bp) and one lacking it (1180 bp). In kidney a predominance of the EIIIA- form (85%) was found in both 6 and 24 months old rats (Fig. 4.4.A). A different ratio was found in the rat brain, with a constant, but less enhanced predominance (65%) of the EIIIA- messengers form in the young rats (Fig. 4.4.B lanes 1,2)and with large individual variations among the 24 months old rats (Fig. 4.4.B lanes 3-5). The individual variations observed in the older animals ranged from a decrease (lane 3) to an increase (lane 5) of the EIIIA+ form with a standard deviation of ±22 on a mean EIIIA- value of 67%. These results do not allow us to make a reliable statistical prediction of the variation in the expression of this site during the ageing process in the brain.

For the analysis of the EIIIB region a similar strategy was used. The primer EIIIB5' and EIIIB3' gave two products one including (1253 bp) and one lacking (980 bp) the EIIIB site. The expression pattern of the EIIIB exon was found to be remarkably constant in brain of both old and young rats , showing a constant predominance of the EIIIB- form (90%) (Fig. 4.5, lanes 1-4). This was true for the same individual and mRNA samples where a marked variation of EIIIA was observed (see above). A similar result was seen in kidney with a constant and even

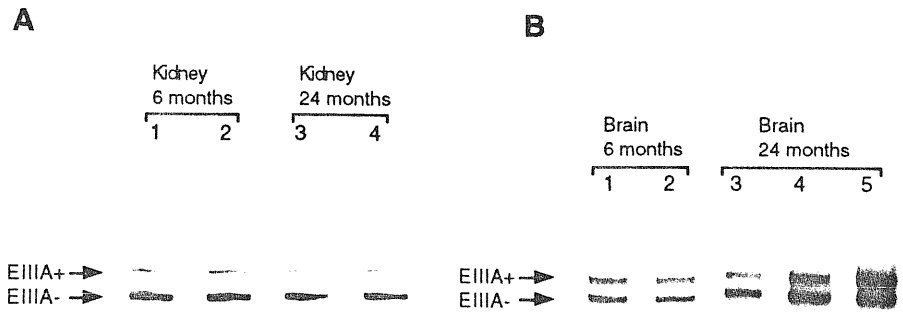


Figure 4.4. EIIIA region analysis.

A) Analysis of the EIIIA region in kidney. Lanes 1 and 2 show the results of the amplification of the EIIIA region of the kidney taken from 6 months old rats. Lanes 3 and 4 shows the results of the amplifications of the kidney taken from 24 months old rats.

B) Analysis of the EIIIA region in brain. In lanes 1 and 2 is shown the result of the amplification of the brain taken from 6 months old rats. Lanes 3, 4 and 5 represent amplification of three samples taken from 24 months old rats brain. A difference in the ratio among the samples 3, 4 and 5 is easily detectable (S.D. ±22 on a mean value 67% for EIIIA-).

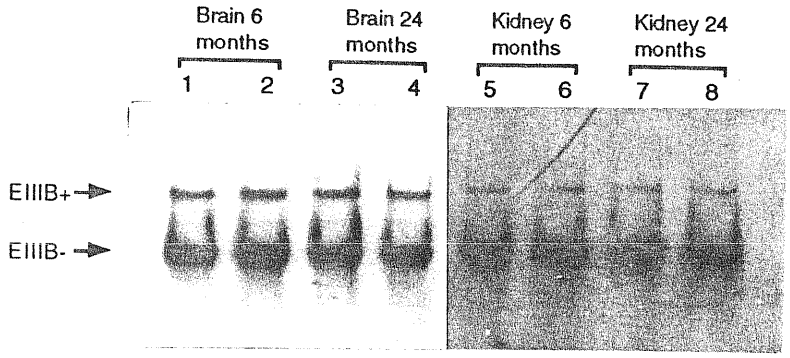


Figure 4.5. Analysis of the EIIIB region in brain and kidney.
Lanes 1, 2 and 3, 4 show the results of the amplification of the EIIIB region in samples taken from two 6 months old and two 24 months old rats' brains, respectively. There's no evident change in the expression pattern with a constant dominance of the EIIIB- form. Lanes 5, 6 and 7, 8 show the results of the analysis of two 6 months old and two 24 months old rats' kidneys, respectively.

more striking, predominance of the messengers lacking the EIIIB exon (95%) (Fig. 4.5., lanes 5-8).

4.1.4 The EIIIA-EIIIB linkage

The analysis of the co-ordinate expression of the EIIIA and EIIIB sites was carried out by amplifying a segment of cDNA containing both regions. For each sample two distinct reactions were performed: one using the primers PREIIIB-, EIIIA3' and one with the primers PREIIIB+, EIIIA3'. The first reaction gave two different products, both lacking the EIIIB, representing respectively the messengers carrying or not the EIIIA exon, whereas the second reaction gave two products including the EIIIB carrying or not the EIIIA exon (Fig. 4.1 and 4.6.A). The analysis of the renal tissue shows no evident correlation between the two sites, neither in the 6 nor in the 24 months old rats. Both EIIIB+ and EIIIB- messengers, were found to have a constant predominance on the EIIIA- form (Fig. 4.6.A, lines 5, 6). On the other hand, in brain, after taking into account the constant predominance of the EIIIA- form in the 6 months rats and the variance in the 24 months animals, there seems to be present a co-ordinated expression pattern of the two exons. There's a striking decrease in the presence of the EIIIA exon in the messengers carrying the EIIIB (Fig. 4.6.A, lanes 2,4) with respect to the ones lacking it (Fig. 4.6.A, lanes 1,3). Considering that in the absence of linkage it would be expected to observe the same EIIIA ratio between the EIIIB+ and EIIIB- forms, the results obtained seems to indicate that, in rat brain, there is a partial co-ordination of expression of the EIIIA and EIIIB exons.

4.1.5 Analysis of the EIII9 exon expression

Systematic searches have failed to find a fourth site of alternative splicing (Schwarzbauer et al., 1987). Like EIIIA and EIIIB, the ninth type III homology is encoded by one exon instead of two, furthermore, from sequence homology analysis it seems to have a weak 3′ splice site missing a true polypyrimidine tract (see Fig. 3.1.2). It is, therefore, a likely candidate for a fourth alternatively spliced region in the FN molecule. This possibility was investigated using a nested-PCR assay that permits the amplification of DNA starting from an extremely low

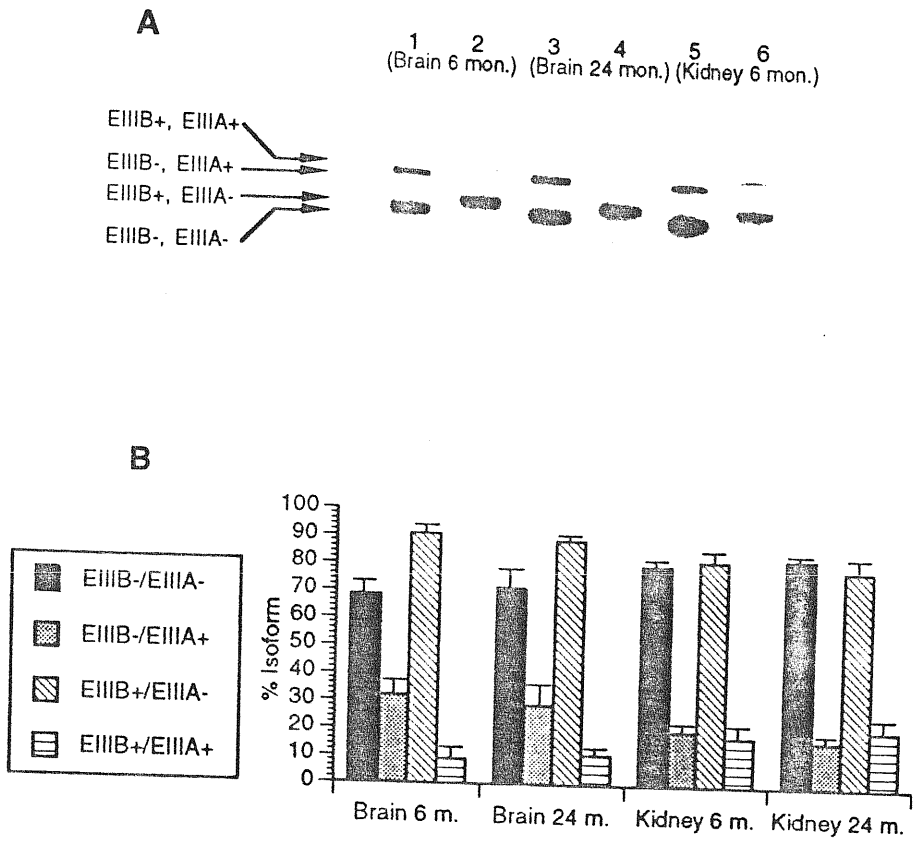


Figure 4.6. Linkage between the EIIIB and the EIIIA region in brain and kidney.
A) Lanes 1-4 show the results of the linkage analysis of a 6 months old (lanes 1, 2) and a 24 months old (lanes 3, 4) rat brains. Lanes 5, 6 show the results of the same analysis performed on a 6 months old rat kidney. Lanes 1, 3 and 5 show the results of the reaction carried out with the primers PREIIIB-, EIIIA3' that detect the messengers lacking the EIIIB exon. Lanes 2, 4 and 6 show the results of the reaction carried out with the primers PREIIIB+, EIIIA3' that detect the messengers carrying the EIIIB exon.
B) The histogram summarises the results of the linkage analysis. In brain there is an evident decrease of the EIIIA presence in the isoform carrying the EIIIB exon with respect to the ones lacking it, whereas in kidney the rapport is maintained constant.

amounts of template. The cDNA synthesized from total mRNA was amplified in a first PCR round for 20 cycles using the primers EIII95'I and EIII93'I (Tab. 2.1 and Fig. 4.7). The products were then digested with the restriction enzymes Apa I and Eco RI that cut inside the EIII9 exon in such a way to eliminate most of the products that include the exon. An aliquot of the digested product was then amplified in a second PCR round for other 25 cycles using the primers EIII95'II and EIII93'II that are internal respect to the ones used in the first round to improve the specificity of the reaction. The strategy chosen minimise the presence of aspecific products after a large number of PCR cycles. Furthermore, it reduces the presence of the product that is mainly represented (the EIII9+ form) and, thus, can enhance the amplification of a less represented population as are expected to be the EIII9- messengers.

With this methodology we analysed brain and kidney tissues of 6 and 24 months (Fig 4.8 lanes 1-4), liver tissues taken from a 6 months rat during a partial hepatectomy (methodology discussed in section 5.1) and 24 hours after the operation (lanes 5 and 6), and HeLa and HepG2 human cell lines (lanes 7 and 8). Figure 4.8 clearly shows that after a total of 45 cycles no products lacking the EIII9 exon appear in the tissues and cells analysed (expected length 256 bp). However a product of 526 bp is present. This product represent the fraction of amplified template containing the EIII9 exon that has not been digested by the "killer cut".

4.2 DISCUSSION

4.2.1 Are EIIIA, EIIIB and the V region subject to a coordinate expression?

The alternative splicing of FN pre mRNA and hence their possible FN isoform translations are the result of an interaction between well defined cis-acting elements and trans acting elements which can be either cell type specific or may be modulated by the cell environment (see chapter 3). Although several studies have already addressed the problem of the modulation of expression of single alternatively spliced exons (Pagani et al., 1991; Oyama et al., 1989; Magnuson et

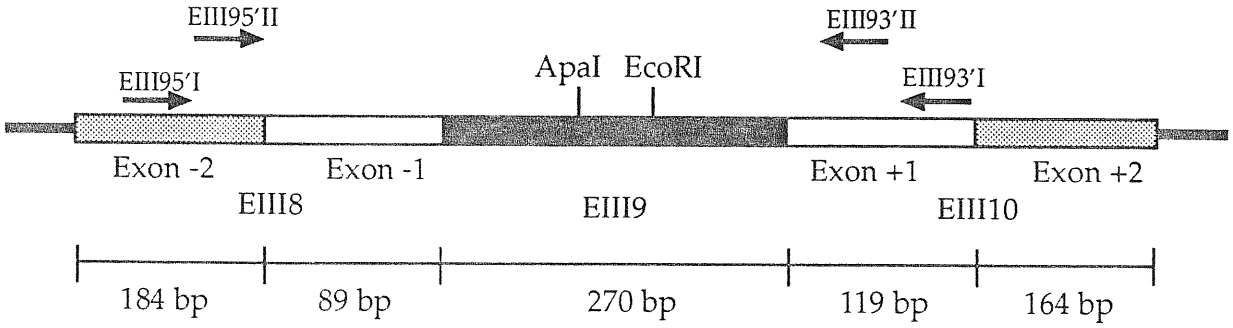


Figure 4.7. Schematic representation of the EIII9 region.
The location of the primers used in the two amplification rounds (EIII95'I, EIII93'I and EIII95'II, EIII93'II) is indicated. The restriction enzymes used in the killer cut (ApaI and EcoRI) are also indicated. Upstream and downstream type III homology are divided in two exons while the EIII9 homology is encoded by a single exon. At the bottom of the figure are reported the length of the exons.

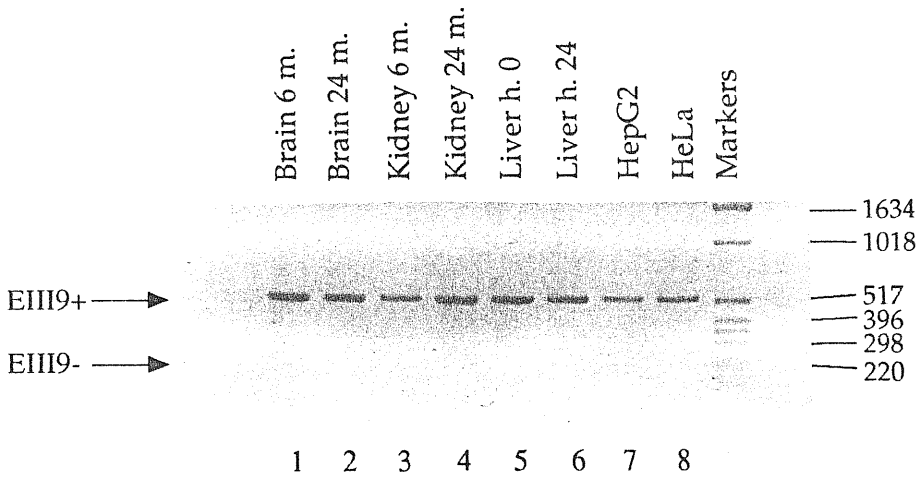


Figure 4.8. Products of PCR analysis of the EIII9 region.

Lanes 1-4 represent the result of the amplification of the EIII9 region in 6 and 24 months old brain and kidney tissues. In lanes 5 and 6 the products of the amplification of liver tissue taken during a partial hepatectomy and after 24 hours from the operation. Results from HeLa and HepG2 cell lines are in lanes 7 and 8. The products expected by EIII9 exon and skipping are indicated (EIII9+, EIII9-).

al., 1991) little was known about the linkage in the same molecule of the different exons (EIIIB, EIIIA, V). One of the original goals of this study was to investigate on the V-EIIIA-EIIIB molecular linkage. Unfortunately, it was not possible at the time to obtain reliable PCR products covering the whole region so it was necessary to split the linkage between the EIIIA-V and EIIIB-EIIIA associations.

Although the splicing regulation is tissue specific and is not co-ordinated between the EIIIA and the V regions neither in brain, nor in kidney, or in liver (result exposed in section 5.3) and it is independent from age. As discussed in chapter 3 splicing of EDA and EIIIB seems to be independent and are regulated by different cis-acting sequences. However the results show a clear correlation between the expression of the two exons in both 6 and 24 months old rat brain, whereas no linkage is observed in kidney. Figure 4.6 shows that in the rat's brain there is a clear decrease of expression of the EIIIA exon in the EIIIB+ messengers with respect to the EIIIB- ones. Nevertheless, the EIIIA ratio variation between the two classes of messengers (EIIIB+ and EIIIB-) in each single rat is independent from the EIIIA individual variability observed in the 24 months old rats (data not shown).

It was also shown that most of the splicing regulation is autonomous for each individual region but we have also observed a partially co-ordinated pattern in the case where the EIIIA+/EIIIB+ mRNA is less abundant than expected from the statistical combination of the ratios given by the single forms . The EIIIA and EIIIB cis-acting elements studied up to now have been found to be different (see chapter 3) and the differential response to the regenerative process of the two sites (see section 5.3) strongly support the hypothesis of a different regulative pathway for the two alternatively spliced exons. However only a complete understanding of the trans acting factors involved in the alternative splicing process may help to understand these complex regulations.

4.2.2 Pattern of expression of the single alternatively spliced regions

To have a clear picture of the variations that occur in the FN molecule during development and senescence an independent analysis of each splicing site was performed.

In this work was demonstrated that, there is an invariant expression ratio of the EIIIA, EIIIB and V mRNA forms in kidney in the 6 and 24 months old rats. On the other hand there are at least two significant variations in brain. FN provides stimuli for specific gene expression, cell spreading, differentiation and growth via integrins receptors (Schwartz et al., 1991; Ingber et al., 1990). The V120 form , that carries the CS1 cell binding site for the integrin $\alpha 4\beta 1$, shows a significant decrease in the 24 months old rats, suggesting a correlation between the loss of plasticity of the cerebral tissue during senescence and a decrease in V120 expression. This finding well correlates with previous observations showing active role of the IIICS region in promoting neurite outgrowth from peripheral nervous system neurones (Humpries et al., 1988), and embryonic neural crest cell migration (Dufour et al., 1988).

The high individual variability of the EIIIA forms in the brain of 24 months old rats with respect to the 6 months old animals has not an obvious explanation as, at least in the cerebral tissue, there is no clear function assigned to the EIIIA region (see section 1.5.6.2). Nevertheless is extremely interesting the relation emerging between the loss of tight control made on individual bases over the EIIIA levels of expression and the ageing process occurring in the brain. Further studies with larger number of individual and in other species are needed to establish the extent of this phenomena.

The search for a fourth alternative splicing site gave negative results, even using a very highly sensitive technique. No alternative splicing of the EIII9 exon has been detected in a various extent of tissues and cell lines (Fig. 4.7). This was surprisingly since the EIII9 exon bear an extremely weak polypyrimidine tract with only one pyrimidine residue, both in human and rat. At least at low level, the usage of an alternative 3′ splice site or exon skipping was expected. It can be hypothised for the EIII9 exon the presence of an enhancer mechanism that can overcame the presence of a weak 3′ splice site helping its recognition in a strong, constitutive, fashion. This mechanism will be probably different from the ones already pointed out for the human EDA and rat EIIIB exons (see chapter 3) since no evident sequence homologies have been reported between the discovered cis-acting sequence and the EIII9 region.

Chapter 5

Fibronectin Expression in the Rat Regenerating Liver

The hepatic tissue is the major source of circulating plasma FN. As described in 1.5.3.2 plasma FN is characterised by the absence of the EIIIA and EIIIB elements. Previous studies (Pagani et al., 1991; Magnuson et al., 1991) have shown that the adult liver presents almost undetectable levels of FN messengers containing one or both EIIIA and EIIIB exons. Nevertheless, in vitro studies have evidenced the ability of the hepatic cell types to produce the cellular FN form in which EIIIA and EIIIB are present (Odenthal et al., 1992; Ramadori et al., 1992).

Up to now several studies have investigated the changing in the FN expression in physiological processes such as wound healing (ffrench-Constant et al., 1989; Brown et al., 1993), epithelial fibrosis (Barnes et al., 1994), vascular intima proliferation (Glukhova et al., 1989), peripheral nerve injury (Mathews and ffrench-constant, 1995), liver fibrosis (Jarnagin et al., 1994) and myocardial infarction (Knowlton et al., 1992). In most of the model studied there is an increased deposition of matrix protein modulating tissue fibrosis in response to injury. In this study a different approach has been chosen, the FN expression has been studied in a physiological response involving active cellular proliferation and tissue regeneration.

Using partial hepatectomy (Waynfort and Flecknell, 1992) we have evaluated in an in vivo experimental model the ability of the liver to undergo reprogramming in case of massive cell proliferation and tissue reorganisation. The rat regenerating liver is an excellent model to study the regulatory mechanisms involved. Particularly the absence of the EIIIA and EIIIB alternatively spliced

regions in the liver tissue gives the possibility to identify clearly any change in the expression of these regions, giving new clues on their functions.

5.1 PARTIAL HEPATECTOMY SURGERY

The liver is composed of four lobes (Fig. 5.1). The median and left lateral lobes comprise about 70% of the liver and their removal is recognised classically as partial hepatectomy. A ventral abdominal incision is made on the anaesthetised rat (Fig. 5.2 A). The liver is exposed and the suspensory ligament attaching the left lateral lobe to the diaphragm is cut (Fig. 5.2 B). The left lateral and the median lobe are now moved out of the abdominal cavity and other two suspensory ligaments are cut (Fig. 5.2 C). A ligature is tied at the base of the two lobes (Fig. 5.2 D). The lobes can, thus, be excised and the incision is closed. No special therapy is required other than a high standard of postoperative care. The liver will regenerate in toto within 10-20 days.

Four months old rats were subject to partial hepatectomy as described above, the hepatic tissue taken was immediately frozen in liquid nitrogen. Two control rats were sham. All the rats were then sacrificed after 24 hours and the hepatic tissue recovered. All the regenerating liver tissues were carefully taken avoiding the scar tissue. Histological examination of the regenerating liver tissue showed several mitotic figures but no evidence of inflammatory cell infiltration.

5.2 FIBRONECTIN mRNA QUANTIFICATION

Previous results obtained by Northern Blot (Helen Mardon, unpublished) on hepatectomized rat sacrificed at 8, 16, 24, 48 and 72 hours after the operation evidenced a noticeable increase in the FN messenger in the regenerating liver after 24 hours. Although no proper quantification was carried out in these preliminary

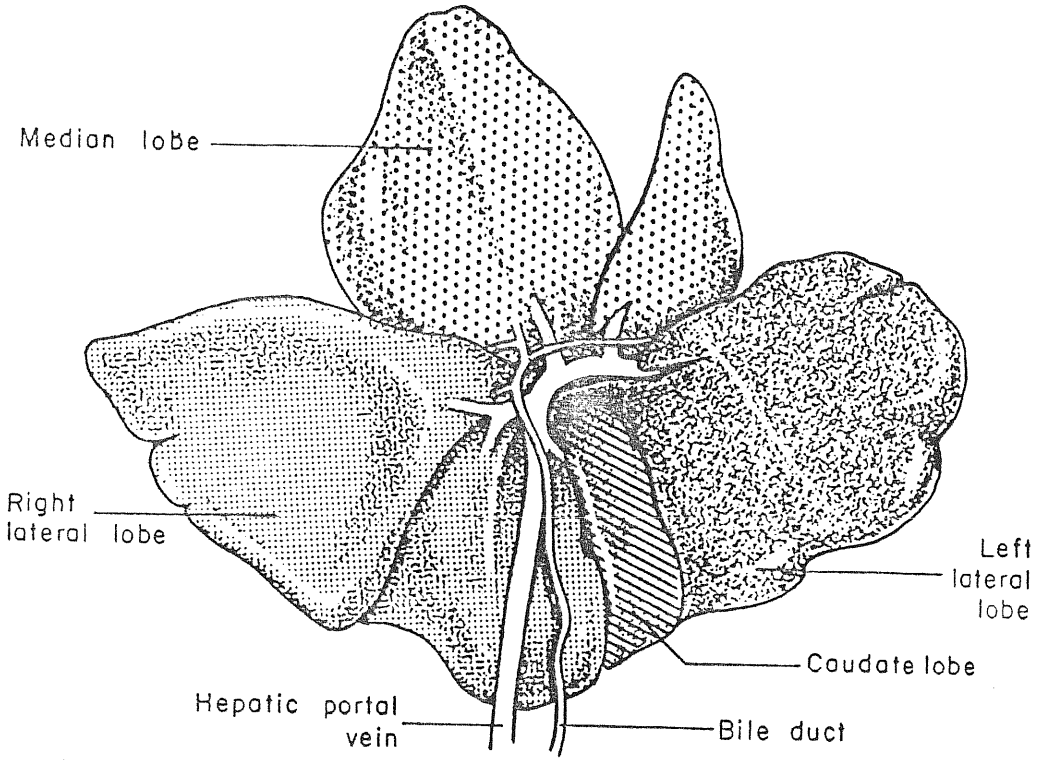


Figure 5.1 View of the dorsal surface of the liver.
(Waynfort and Flecknell, 1992)

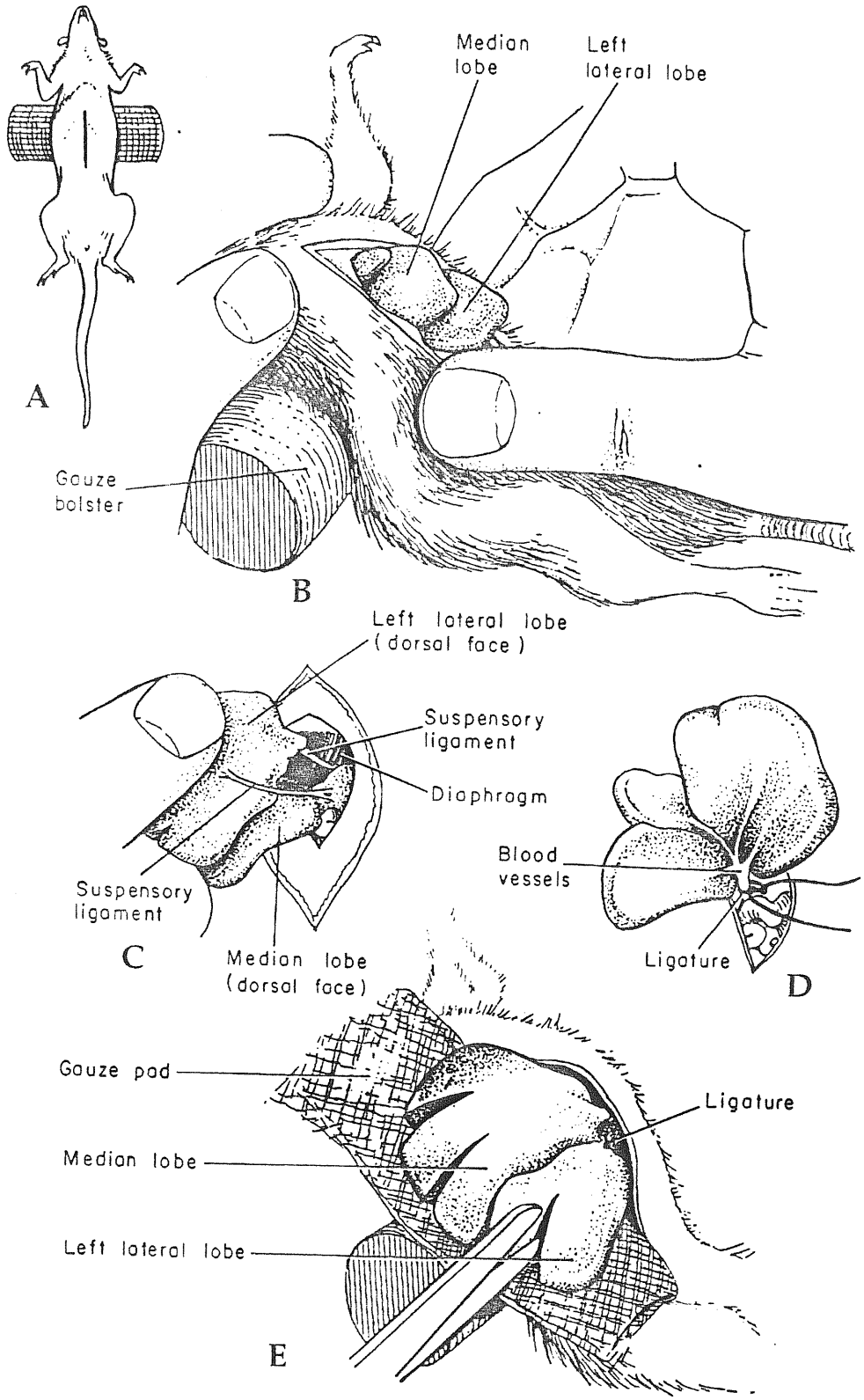


Figure 5.2 Partial hepatectomy.

A) Site of incision for partial hepatectomy. B) Median and left lateral lobes are exposed. C) Two of the three suspensory legaments which have to be cut to free the liver. D) Ligating the blood vessels to the median and left lateral lobes. E) Cutting the lobes.

(Waynfort and Flecknell, 1992)

experiments. This was further studied by quantitative PCR analysis (Gilliland et al., 1990)). The technique, which has been successfully applied to viral RNA quantification (Menzo et al., 1992), consists in coamplifying two templates that slightly differ in length but are recognised by the same set of primers. This ensure the same amplification efficiency for both template species. If the quantity of one template is known the amount of the other can be extrapolated from the ratio between the two products of the amplification reaction.

This method allows a reliable quantification of small amounts of DNA and RNA. We have therefore developed a competitive RT-PCR assay to detect variations in FN expression levels between the normal and the regenerating liver tissue. The efficiency of the RT step and the equal amount of the total mRNA used in the reactions were controlled by comparing the results with the competitive PCR analysis of the glyceraldehyde-phosphate-dehydrogenase (GAPDH) mRNA levels. GAPDH is a housekeeping enzyme whose mRNA was seen to remain constant regardless of the replication stimuli received by the cell.

5.2.1 Quantitative PCR assay

The quantification of the FN mRNA levels in the normal, control and hepatectomized liver was carried out by coamplification of the cDNA synthesised from the total RNA extracted from the tissues and the plasmid pRFout260 with the primers 5'OUT and 3'OUT. Quantification of the GAPDH mRNA levels in the same samples was achieved by coamplifying the cDNAs and the plasmid pGAPDH260 (see 2.2.12.5 and Fig. 5.3) with the primers 5'GAPDH and 3'GAPDH. Five independent RNA extractions and cDNA synthesis were carried out for each sample. A constant amount of reverse transcribed total mRNA was used in each reaction whereas a series of dilution, ranging from 1 ng to 0.1 pg, of the competitor plasmid was used. Figure 5.4 (rows B, D, F) shows that the equivalence point between the GAPDH mRNA template and its competitor (plasmid pGAPDH260) remains constant (2-4 pg) both in the liver taken during the hepatectomy and 24 hours after the partial hepatectomy. Moreover in the sham operated control there is no detectable variation in the overall transcriptional activity of GAPDH during the regenerating process. On the other hand, for the FN messenger (rows A,C,E

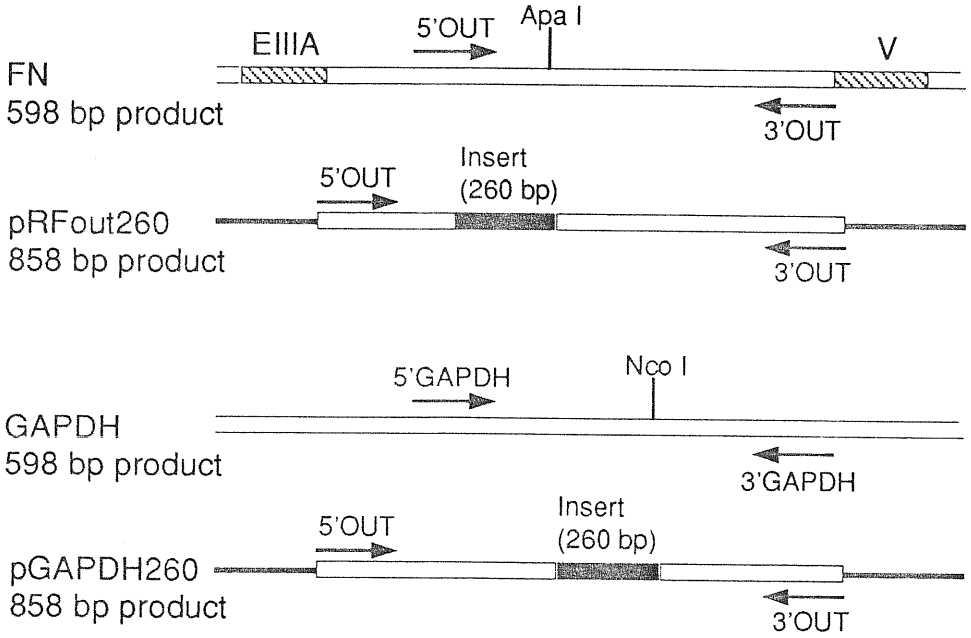


Figure 5.3. Constructs pRFout260 and pGAPDH260.
Locations of the primers 5'OUT - 3'OUT on the FN cDNA and pRFout 260 construct and 5'GAPDH - 3'GAPDH on the GAPDH cDNA and pGAPDH260 construct. The size of the amplified products are also indicated.

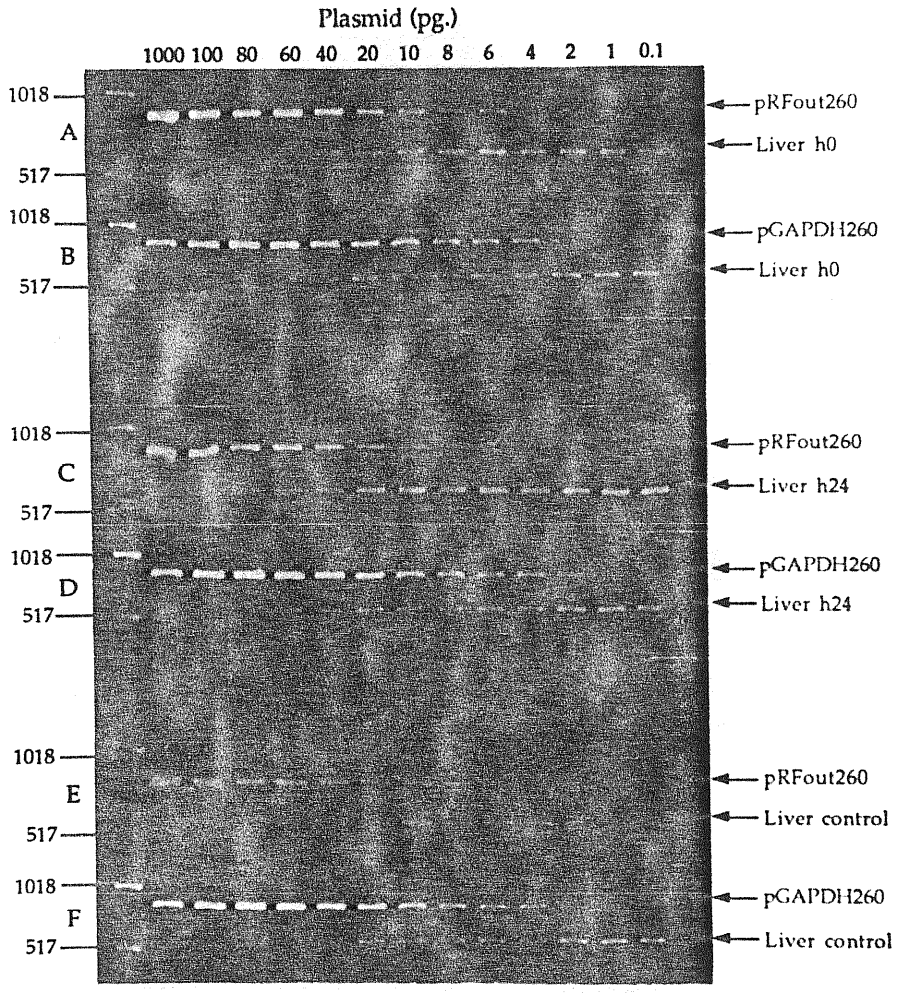


Figure 5.4. Quantitative PCR assay.

Row A shows the PCR set of reactions in which an equal amount of cDNA, synthesised from the mRNA of the liver taken at the time of the operation (liver h0), is titrated, in a series of reactions, with a decreasing amount of the competitor plasmid pRFout260. In row C the same set of reactions is represented for the regenerating liver taken 24 hours after the operation (Liver h24). Row E show the reactions carried out for the control liver analysed 24 hours after the sham operation. In row B, D and F there are the results of the competitive PCR carried out to quantify the GAPDH messenger in the liver taken during the operation, the regenerating, and the control one, respectively.

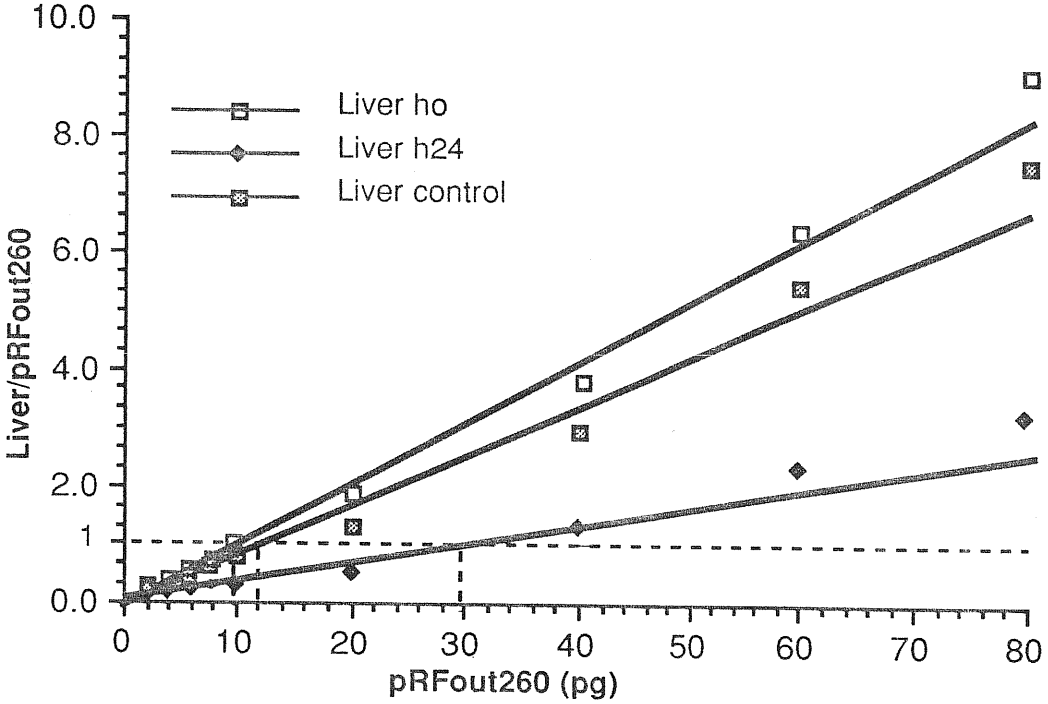


Figure 5.5. Quantification of the FN messenger.

The graph shows the ratio between the O.D. of the liver FN PCR product and the competitor (pRFout260) PCR product plotted against the competitor quantity in each reaction. From the intersection of each line with the value 1 of the axis (equivalence point between the two template species) a more accurate quantification is reached. In the sample of liver taken during the operation the equivalence point is at 10 pg the control sample shows an increase to 12 pg whereas the regenerating liver shows an increase of up to 30 pg.

and Fig. 5.5), its equivalence point with the competitor (plasmid pRFout260) is reached at 10 pg in the sample taken at the time of the hepatectomy (row A). There is an increase up to 30 pg in the FN transcript expression levels measured in the regenerating liver after 24 hours (row C). Conversely, the control sample taken 24 hours after the sham operation shows only a slight increase up to 12 pg (row E).

5.3 FIBRONECTIN PRE-mRNA PROCESSING

Since the highest FN levels were detected 24 hours after the operation, the regenerating hepatic tissue of those rats was analysed to characterise any change in the FN's splicing pattern when compared with the tissue taken at the time of the operation. The tissues were amplified by PCR to analyse the three different splice sites (EIIIB, EIIIA and V), their level of expression and the linkage of the different forms. The analysis of the alternative splicing pattern of the liver tissues has been carried out as for the brain and the kidney in chapter 4. The same PCR strategy has been used. Location of the primers on the FN molecule is shown in figure 4.1 while the products obtained by each primer pair and primer sequences are summarised in table 4.1 and 1.1 respectively.

5.3.1 EIIIA and EIIIB exons expression

The amplification of the fragment including the EIIIB site was performed with the primers EIIIB5' and EIIIB3'. This reaction gives two products, one of 543 bp which includes the EIIIB exon (EIIIB+), and the other of 270 bp lacking the EIIIB exon (EIIIB-). The tissues taken at time 0 and after 24 hours in both hepatectomized and control rats show only the shorter product. Therefore only the EIIIB- FN mRNA forms are present in the liver even during the regenerative process (Fig. 5.6 A).

The amplification of the region that includes the EIIIA site was performed with the primers EIIIA5' and EIIIA3'. This amplification gave two different products one of 200 bp in which the EIIIA region is excluded (EIIIA-) and one of

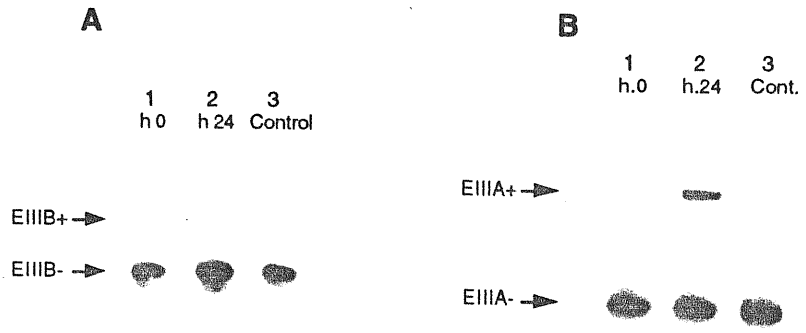


Figure 5.6 Analysis of the EIIIB and EIIIA regions.

A) Line 1 shows the result of the gel electrophoretic analysis of the amplification products on the EIIIB region in the liver taken during the partial hepatectomy. Line 2 shows the remaining hepatic tissue of the rat analysed 24 hours after the operation. In lane 3 is shown the analysis of the liver taken 24 hours after the sham operation.

B) Lanes 1 and 2 show the results of the gel electrophoretic analysis of the amplification products on the EIIIA region in a 6 months old rat on the hepatic tissue taken during the partial hepatectomy (lane 1) and 24 hours after the operation (lane 2). Lane 3 shows analysis of the control tissue 24 hours after the sham operation.

470 bp in which the region EIIIA is included (EIIIA+). In the hepatic tissue taken at the time of operation the EIIIA+ form is below 1% of the total (Fig. 5.6 B, lane 1), but the tissue taken at 24 hours shows a remarkable increase of that form, up to 17% of the total (lane 2). In contrast no changes were observed in the sham operated control tissue (lane 3). This result shows a strong correlation between the involvement of the EIIIA region, normally absent in the FN synthesised by the liver, and the regenerating event. This phenomena is not due to the surgical stress as proved by the barely detectable level of EIIIA+ in the control tissue.

5.3.2 V region expression

The changes in the synthesis of the V variants (V120, V95 and V0) and their linkage to the appearance of the EIIIA region was also studied by two different PCR reactions for each sample. The primers PREIIIA- and V3' were used to detect the V pattern linked to the EIIIA- form while the primers PREIIIA+ and V3' were used to detect the V pattern linked to the EIIIA+ form (Fig. 4.1). As expected for the tissue extracted at the time of operation and for the surgical stress control there were no mRNAs carrying the EIIIA region, the EIIIA- mRNA shows a relative abundance of the V0 form (57%) (Fig. 5.7, lane 1). In the analysis of the regenerating tissue there was a significant variation in the splicing pattern of the V region that was equal in both the EIIIA+ and the EIIIA- forms with a relative increase of the V120 form from 11% of the total in the tissue taken at the time of the operation to 32% of the total in the regenerating tissue (Fig. 5.7, lanes 1,3). It is clear therefore, that the new types of messengers carrying the EIIIA region arising with the regeneration process are not associated to a particular V form.

5.4 DISCUSSION

5.4.1 Fibrotic and regenerating liver

The rat regenerating liver provides an excellent model to study the changes that occur to FN transcription and pre mRNA processing during the in vivo

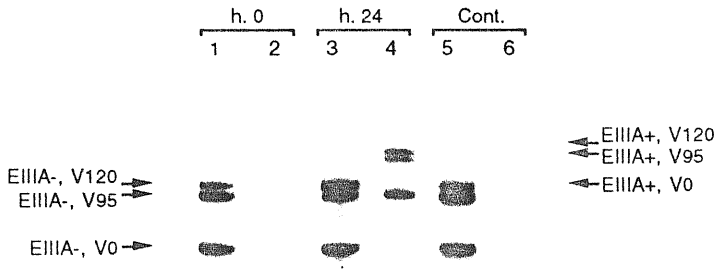


Figure 5.7. Analysis of the linkage between the EIIIA and the V region.
Lanes 1 and 2 present the result of the gel electrophoretic fractionation of the PCR amplification products of the hepatic tissue taken from a 6 months old rat during a partial hepatectomy. Lines 3 and 4 present the analysis of the hepatic tissue of the same rat 24 hours after the operation. Lines 5 and 6 present the result of the analysis of the hepatic tissue taken from a control rat 24 hours after a sham operation. In lines 1, 3 and 5 the samples were amplified using the primers PREIIIA- and V3', that detected the EIIIA- messengers linked to the three V forms. In lines 2, 4 and 6 the samples were amplified using PREIIIA+ and V3', that detected the EIIIA+ messengers linked to the three V forms.

physiological response to environmental changes. Several in vivo experimental systems showed an alteration of the FN expression pattern in pathological situations as discussed in section 1.5.6 (ffrench-Constant et al., 1989; Brown et al., 1993; Barnes et al., 1994; Glukhova et al., 1989; Mathews and ffrench-constant, 1995; Jarnagin et al., 1994; Knowlton et al., 1992). Most of these observations concern physiological responses involving massive tissue fibrosis in response to injury. As exposed in section 1.5.6.2, a brilliant experimental work performed by Jarnagin et al. (1994) on the rat fibrotic liver started to uncover EIIIA functions showing its role in lipocyte activation. In this study a different approach has been chosen, the FN expression has been studied in a physiological response involving active cellular proliferation and tissue regeneration.

Liver fibrosis and liver regeneration are two distinct well known pathophysiological processes. Liver fibrosis reflects quantitative and qualitative changes within the ECM and represents an important aspect of chronic liver disease. Chronic injury (cirrhosis), inflammation, toxic cell injury, tumour growth can perturb a delicate balance that occur among the ECM and cells of different types. In response liver cells produce and deposit an abnormal connective tissue that influence cellular functions stimulating conversion of lypocites to myofibroblasts. Unlike chronic injury, subsequent to an acute injury, the liver is capable of a complete regeneration, thereby restoring its structural and functional integrity. The result of regeneration following partial hepatectomy is not the growth of a new liver segment at the resection margin of an old, quiescent liver remnant; rather, the entire remaining liver undergoes proliferation until the hepatocyte cell mass is restablished and the original hepatic structure restored. During hepatic regeneration after partial hepatectomy no fibrosis develops and particularly there is no lipocyte activation.

5.4.2 Increased fibronectin transcription

It was of interest to analyse the absolute amount of FN mRNA during regeneration and to establish if during tissue regeneration there is an increase in the rate of FN transcription as already observed in other pathological conditions involving stress and acute phase reactions (Knowlton et al., 1992; Mamuya et al.,

1992; Murphy-Ullrich and Mosher, 1986). The quantitative PCR method used allowed, for the first time, this problem to be addressed in a precise and effective way. We demonstrate that there is a specific increase of FN transcripts in the regenerating liver up to 3 fold 24 hours after the operation while the sham operated control shows only a slight increase. This result well correlates with the observed increase of the TGF-β expression in the regenerating liver (Jakowlew et al., 1991). As discussed in section 1.5.4 the FN promoter contains sequence conferring response to TGF-β (Dean et al., 1990), thus, this may be one of the main stimulus to increase the FN production.

5.4.3 Variations in the FN alternative splicing pattern

The V120 FN form has been shown to be abundant in liver fetal tissue in human, chick and rat (Oyama et al., 1989; Pagani et al., 1991; ffrench-Constant and Hynes, 1988). We have now shown that V120 also increase in the regenerating liver. The functional significance of this variance probably involves the integrin α4β1, the cellular receptor for the CS1 cell binding site included in the V120 form. Indeed, FN can provide a stimulus for cellular spreading and growth via its integrin receptors (Schwartz et al., 1991; Ingber et al., 1990). Our results however do not establish definitively if the observed increase in expression is due to a reprogramming of the RNA processing in the hepatocyte, the endothelial cells or any other cell type that is actively contributing to liver regeneration. However, since the results described in chapter 4 clearly show that the V pattern is specific for different tissues, the invariance of the V0/V95 ratio, between normal and regenerating liver, suggests that the hepatocyte reprogramming hypothesis is the more likely. The same is valid for the appearance of the EIIIA+ exon, where the EIIIA+ forms are linked to the V0, V95 and V120 forms in the same proportion as to the EIIIA- forms. Thus, EIIIA+ and EIIIA- messengers are associated with the same V pattern, reinforcing the hypothesis of the same cellular origin for both messengers types. This is consistent with the observations of Odenthal et al. (1992) that, after 3 days in culture hepatocytes start producing EIIIA+ FN. Our results suggest a role in cell proliferation and tissue modelling for the EIIIA+ and V120. The two regions seem to be subjected to independent regulatory mechanism and

we have shown here that they are not exclusively linked in the same mRNA species. Similarly to the fibrotic liver there if not a full reprogramming to an embryo-like splicing pattern in the regenerating liver. Indeed, the EIIIB+ messenger, the fetal isoform by excellence, is absent. This confirms the results obtained in this work and by R. Hynes group on the independent regulation of the alternative splicing of the EIIIA and EIIIB regions (as discussed in chapter 3). This result makes unlikely the hypothesis for a common function, at least in this model, of the two sites and differentiates the liver regenerating process from others, such as wound healing in which there is an increase of both EIIIA and EIIIB (ffrench-Constant et al., 1989).

It is clear that the hepatocyte proliferative stimuli induced by the hepatectomy has resulted in a reprogramming of the FN transcription and RNA processing pathways. These variations are specifically due to the regenerative process following hepatectomy as the changes due to surgical stress are minimal and the tissues were carefully taken avoiding the scar tissue or any inflammation site. However the precise cellular type that undergoes to these changes is still unknown and should be uncovered performing in situ hybridisation studies of the regenerating tissue. Also unknown is the precise functional significance of the altered pattern of FN expression. Nevertheless these results give new clues on the active role of the EIIIA exon in the activation of cell proliferation and tissue modelling a role that should differ from the lipocyte activation a process not present in the regenerating liver.

References

Amendt, B.A., Zhi-Hai,S. and Stolzfus,C.M. (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. Cel. Biol. 15: 4606-4615.

Amrein, H., Hedley,M.L. and Maniatis,T. (1994). The role of specific protein-protein interactions in positive and negative control of pre-mRNA splicing by transformer 2. Cell 76: 735-746.

Aota, S., Nagai,T. and Yamada,K.M. (1991). Characterisation of regions of fibronectin besides the arginine-glycine-aspartic acid sequence required for adhesive function of the cell-binding domain using site-directed mutagenesis. J. Biol. Chem. 266: 15938-15943.

Bakhshi, A., Guglielmi,P., Siebenlist,U., Ravetch,J.V., Jensen,J.P. and Korsmeyer,P. (1986). A DNA intron/deletion necessitates an aberrant RNA splice accounting for a mu heavy chain disease protein. Proc. Natl. Acad. Sci. 83: 2689-2693.

Banroques, J. and Abelson, N.J. (1989). A protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. 9: 3710-3719.

Barnes, J.L., Hastings,R.R. and De La Garza,M.A. (1994). Sequential expression by cellular fibronectin by platelets, macrophages and mesanglial cells in proliferative glomerulonephritis. Am. J. Pathol. 145: 585-597.

Barone, M.V., Henchcliffe,C., Baralle,F.E. and Paolella,G. (1989). Cell type specific trans-acting factors are involved in alternative splicing of human fibronectin pre-mRNA. EMBO J. 8: 1079-1085.

Bennet, M., Pinol-Roma,S., Staknis,D. and Dreyfuss,G. (1992). Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. Mol. Cell. Biol. 12: 3165-3175.

Berget, S.M., Moore,C. and Sharp,P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. 74: 3171-3175.

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

Bernath, V.A., Muro, A.F., Vitullo, A.D., Bley, M.A., Barañao, J.L. and Kornblihtt, A.R. (1990). Cyclic AMP inhibits fibronectin gene expression in a newly developed

granulosa cell line by a mechanism that suppresses cAMP-responsive-element transcriptional activation. J.Biol.Chem. 265: 18219-18226.

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: 389-402.

Black, D.L. (1992). Activation of c-src neuron specific splicing by an unusual RNA element in vivo and in vitro. Cell 69: 795-807.

Blencowe, B.J., Sproat,B.S., Ryder,U., Barabino,S. and Lamond,A.I. (1989). Antisense probing of the human U4/U6 snRNP with biothinylated 2' -OME RNA oligonucleotides. Cell 59: 531-539.

Boggs, R.T., Gregor,P. Idriss,S., Belote,M.J. and McKeow,M. (1987). Regulation of sexual differentiation in Drosophila melanogaster via alternative splicing of RNA from the transformer gene. Cell 50: 739-747.

Bork, P. and Doolittle, R.F. (1992). Proposed acquisition of an animal protein domain by bacteria. Proc. Natl. Acad. Sci. 89: 8990-8994.

Borsi, L., Castellani,A., Risso,A.M., Leprini,A. and Zardi,L. (1990). Transforming growth factor-b regulates the splicing pattern of fibronectin messenger RNA precursor. FEBS Lett. 261: 175-178.

Bowditch, R.D., Hariharan,M., Tominna,E.F., Smith,J.F., Yamada,K.M., Getzoff,E.D. and Ginsberg,M.H. (1994). Identification of a novel integrin binding site in fibronectin. J. Biol. Chem. 269: 10856-10863.

Brandt, C.R., Morrison,S.L., Birshtein,B.K. and Milcarek,C. (1984). Loss of a consensus splice signal in a mutant immunoglobulin gene eliminates the the CH1 domain exon from the mRNA. Mol. Cell. Biol. 4: 1270-1277.

Breitbart, R.E. and Nadal Ginard, B. (1987). Developmentally induced, muscle-specific trans factor control the differential splicing of alternative and constitutive troponin T exons. cell 49: 793-803.

Brosi, R., Groning,K., Behrens,S.E., Luhrmannn,R. and Kramer,A. (1993). Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60 kD subunit to yeast PRP9. Science 262: 102-104.

Brown, L.F., Dubin,D., Lavigne,B., Logan,B., Dvorak,H.F. and Vandewater,L. (1993). Macrophages and fibroblast express embryonic fibronectins during cutaneous wound healing. Am. J. Pathol. 142: 585-597.

Brunak, S. and Engelbrecht, J. (1991). Prediction of human mRNA donor and acceptor sites from the DNA sequence. J. Mol. Biol. 220: 49-65.

Burke, E.M. and Danner, D.B. (1991). Changes in fibronectin mRNA splicing with in vitro passage. Biochem. Bioophys. Res. Commun. 178: 620-624.

Burtis, K.C. and Baker, B.S. (1989). Drosophila dublesex gene controls somatic sexual differentiation by producing alternative spliced mRNAs encoding related sex-specific polypeptides. Cell 56: 997-1010.

Burtis, K.C., Coschigano,K.T., Baker,B.S. and Wensink,P.C. (1991). The doublesex proteins of Drosophila melanogaster bind directly to a sex specific yolk protein enhancer. EMBO J. 10: 2577-2582.

Caceres, J.F. and Krainer, A.R. (1993). Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. EMBO J. 12: 4715-4726.

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

Carnemolla, B., Leprini,A., Allemanni,G., Saginati,M. and Zardi,L. (1992). The inclusion of the type III repeat ED-B in the fibronectin molecule generates conformational modifications that unmask a cryptic sequence. J. Biol. Chem. 267: 24689-24692.

Carothers, A.M., Urlaub,G., Grunberger,D. and Chasin,L. (1993). Splicing mutants and their second-site suppressors at the dihydrofolate reductase locus in chinese hamster ovary cells. Mol. Cell. Biol. 13: 5085-5098.

Cavaloc, Y., Popielarz,M., Fuchs,J.P., Gattoni,R. and Stevenin,J. (1994). Characterization and cloning of tthe human splicing factor 9G8: a novel 35 kDa factor of the serine/arginne protein family. EMBO J. 11: 2639-2649.

Chebli, K., Gattoni,R., Schmitt,P., Hildwein,G. and Stevenin,J. (1989). The 216-nucleotide intron of the E1A pre-mRNA contains a hairpin structure that permits utilization of unusually distal branch acceptor. Mol. Cell. Biol. 9: 4852-4861.

Chen, E.Y., Liao,Y.C., Smith,D.H., Barrera-Saldana,H.A., Gelinas,R.E. and Seeburg,P.H. (1989). The human growth hormone locus: nucleotide sequence, biology, and evolution. Genomics 4: 479-497.

Chen, I.T. and Chasin, L.A. (1994). Large exon does not limit splicing in vivo. Mol. Cell. Biol. 14: 2140-2146.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiociyanate-phenol-cloroform extraction. Anal. Biochem 162: 156-159.

Chow, L.T., Gelinas,R.E., Broker,T.R. and Roberts,R.J. (1977). An amazing sequence rearrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12: 1-8.

Company, M., Arenas,J. and Abelson,N.J. (1991). Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. Nature 349: 487-493.

Cunningam, S.A., Else,A.J., Potter,B.V.L. and Eperon,I.C. (1991). influences of separation and adjacent sequences on the use of alternative 5' splice sites. J. Mol. Biol. 217: 265-281.

Dean, D., Bowlus,C. and Bourgeois,S. (1987). Cloning and analysis of the promoter region of the human fibronectin gene. Proc. Natl. Acad. Sci. 84: 1876-1880.

Dean, D.C., Newby, R.F., and Bourgeois, S. (1988). Regulation of fibronectin biosynthesis by dexamethasone, transforming growth factor beta and cyclic AMP in human cell lines. J.Cell.Biol. 106: 2159-2170.

Dean, D.C., Blakely, M.S., Newby, R.F., Ghazal, P., Henninghauser, L., and Bourgeois, S. (1989). Forskolin inducibility and tissue-specific expression of the fibronectin promoter. Mol.Cell.Biol. 9: 1498-1506.

Dean, D.C., McQuillan, J.J. and Weintraub, S. (1990). Serum stimulation of fibronectin gene expression appears to result from rapid serum-induced binding of nuclear proteins to a cAMP response element. J.Biol.Chem 265: 3522-3527.

Del Gatto, F. and Breathnach, R. (1995). Exon and intron sequences, respectively, repress and activate splicing of a fibroblast growth factor receptor 2 alternative exon. Mol. Cell. Biol. 15: 4825-4834.

Del Gatto, F. and Breathnach, R. (1995). Exon and intron sequences, respectively repress and activate splicing of a fibroblast growth factor receptor 2 alternative exon. Mol. Cel. Biol. 15: 4825-4834.

Dirksen, W.P., Hampson,R.K., Qiang,S. and Rottman,F. (1994). A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. J. Biol. Chem. 269: 6431-6436.

Dominski, Z. and Kole, R. (1991). Selection of splice sites involved in pre-mRNAs with short internal exons. Mol. Cell. Biol. 11: 6075-6083.

Dominski, Z. and Kole, R. (1992). Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. 12: 2108-2114.

Dominski, Z. and Kole, R. (1994). Identification and characterization by antisense oligonucleotides of exon and intron sequences required for splicing. Mol. Cell. Biol. 14: 7445-7454.

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

Dufour, S., Duband,J.L., Humphries,M.J., Obara,M., Yamada,K.M. and Thiery,J.P. (1988). Attachment, spreading and locomotion of avian neural crest cells are mediated by multiple adhesion sites on fibronectin molecules. EMBO J. 7: 2661-2671.

Eperon, I.C., Ireland,D.C., Smith,R.A., Mayeda,A. and Krainer,A.R. (1993). Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. EMBO J. 12: 3607-3617.

ffrench-Constant, C. and Hynes, R.O. (1988). Patterns of fibronectin gene expressing and splicing during cell migrations in chicken embryos. Development 104: 369-382.

ffrench-Constant, C., Van De Water, L., Dvorak, H.F. and Hynes,R.O. (1989a). Reappearance of an embryonic pattern of fibronectin splicing during wound healing in the adult rat. J. Cell. Biol. 109: 909-913.

ffrench-Constant, C. and Hynes, R.O. (1989b). Alternative splicing of fibronectin is temporally and spatially regulated in the chicken embryo. Development 106: 375-388.

Flickinger, T.W. and Salz, H.K. (1994). The drosophila sex determination gene snf encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. Genes Dev. 8: 914-925.

Fu, X.D., Katz,R.A., Skalka, A.M. and Maniatis,T. (1991). The role of branch point and 3'-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. 5: 211-220.

Fu, X.D. and Maniatis, T. (1992). Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. Science 256: 535-538.

Fu, X.D. (1993). Specific comittment of different pre-mRNAs to splicing by single SR proteins. Nature 365: 82-85.

Garcia-Blanco, M.A., Anderson,G.J. Beggs,J. and Sharp,P.A. (1989). Identification and purification of a 62,000 dalton protein which binds specifically to the polypyrimidine tract of introns. Genes Dev. 3: 1874-1886.

Gattoni, R., Schmitt,P. and Stevenin,J. (1988). In vitro splicing of E1A transcripts: characterization of novel reactions and of multiple branch points abnormally far from the 3' splice sites. Nucleic Acid. Res. 16: 2389-2409.

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

Gibbs, R.A., Nguyen,P., Edwards,A., Civitello,A.B. and Caskey,C.T. (1990). Multiplex DNA deletion detection and exon sequencing of the hypoxantine phosphoribosyltransferase gene in Lesh-Nyhan families. Genomics 7: 235-244.

Gilliland, G., Perrin,S. and Bunn,H.F. PCR protocols a Guide to Methods and Applications. San Diego: Academic Press Inc, 1990.

Glukhova, M.A., Frid,M.G., Shekhonin,B.V., Vasilevskaya,T.D., Grunwald,J., Saginati,M. and Kotelianski,V.E. (1989). Expression of extra domain A fibronectin sequence in vascular smooth muscle cells is phenotype dependent. J. Cell. Biol. 109: 357-366.

Goguel, V. and Rosbash, M. (1993). Splice site choiche and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. Cell 72: 893-902.

Gontarek, R.R., McNally,M.T. and Beemon,K. (1993). Mutation of an RSV intronic element abolishes both u11/U12 snRNP binding and negative regulation of splicing. Genes Dev. 7: 1926-1936.

Goux-Pelletan, M., Libri,D., Carafa,Y.D., Fiszman,M., Brody,E. and Marie,J. (1990). In vitro splicing of mutually exclusive exons from the chicken b-tropomyosine gene: role of the branch point location and very long pyrymidine strectch. EMBO J. 9: 241-249.

Gozani, O., Patton,J.G. and Reed,R. (1994). A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. EMBO J. 13: 3356-3367.

Green, M.R. (1991). Biochemical mechanism of constitutive and regulated mRNA splicing. Annu. Rev. Cell Biol. 7: 559-599.

Grosveld, G.C., de Boer,E., Shewmaker,C.K. and Flavel,R.A. (1982). DNA sequences necessary for trancription of of the rabbit b-globin gene in vivo. Nature 295: 120-126.

Guan, J.L. and Hynes, R.O. (1990). Lymphoid cells recognize an alternatively spliced segment of fibronectin via integrin receptor 4-b. Cell 60: 53-61.

Gui, J.F., Tronchere,H., Chandler,S.D. and Fu,X.D. (1994). Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. Proc. Natl. Acad. Sci. 91: 10824-10828.

Gunderson, S.I., Beyer,K., Martin,G., Keller,W., Boelens,W.C. and Mattaj,I.W. (1994). The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. Cell 76: 531-541.

Guo, W., Mulligan,G.J., Wormsley,S. and Helfman,D.M. (1991). Alternative splicing of b-tropomyosin RNA. Genes Dev. 5: 2096-2107.

Gutman, A. and Kornblihtt, A. (1987). Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA. Proc. Natl. Acad. Sci. USA 84: 7179-7182.

Hartmuth, K. and Barta, A. (1988). Unusual branch point selection in processing of human growth hormone pre-mRNA. Mol. Cell. Biol. 8: 2011-2020.

Hawkins, J.D. (1989). A survey on intron and exon lenghts. Nucleic Acid Res. 16: 9893-9908.

Hedley, M.L. and Maniatis, T. (1991). Sex-specific splicing and polyadenilation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. Cell 65: 579-586.

Helfman, D.M., Ricci,W.M. and Finn,L.A. (1988). Alternative splicing of tropomyosin pre-mRNA in vitro and in vivo. Genes and Dev. 2: 1627-1638.

Henchcliffe, C. (1988). Alternative splicing of human fibronectin transcripts. D. Phil. Thesis

Hidada, Y., Palella,T.D., O'Toole,T.E., Tarle,S.A. and Kelly W.N. (1987). Human Adenosine phosphorylribosyl transferase- identification of allelic mutations at the nucleotide level as a cause of a complete deficiency of the enzyme. J. Clin. Inv. 80: 1409-1415.

Higgs, D.R., Goodburn,S.E.Y., Lamb,J., Clegg,J.B., Weatherall,D.J. and Proudfoot,N.J. (1983). a-thalassemia caused by a polyadenilation signal mutant. Nature 306: 398-400.

Hodges, P.E. and Rosemberg, L.E. (1989). The spf mouse: a missense mutation in the ornithine transcarbamylase gene also causes aberrant mRNA splicing. Proc. Natl. Acad. Sci. 86: 4142-4146.

Hoffman, B.E. and Grabowski, P.J. (1992). U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. Genes Dev. 6: 2554-2568.

Hook, M and Svitalski, L.M. In Fibronectin: Biology of the Extracellular Matrix, ed. D.F. Mosher. 295-308. San Diego: Academic Press, 1989.

Horowitz, D.S. and Krainer, A.R. (1994). mechanism for selecting 5' splice sites in mammalian pre-mRNA splicing. TIG 10: 106.

Huh, G. and Hynes, R.O. (1993). Elements regulating an alternatively spliced exon of the rat fibronectin gene. Mol. Cell. Biol. 13: 5301-5314.

Humphrey, M.B., Bryan,J., Cooper,T.A. and Berget,S.M. (1995). A 32-nucleotide exon splicing sites in a differential internal exon. Mol. Cell. Biol. 15: 3979-3988.

Humphries, M.J., Akiyama,S.K., Komoriya,A., Olden,K. and Yamada,K.M. (1988). Neurite extension of chicken peripheral nervous system neurons on fibronectin: relative importance os specific adhesion sites in the central cell-binding domain and the alternatively spliced type III connecting segment. J. Cell. Biol. 106: 1289-1297.

Hynes, R.O. (1985). Molecular biology of fibronectin. Annu. Rev. Cell Biol. 1: 67-90.

Hynes, R.O. Fibronectins. New York: Springer-Verlag, 1990.

Hynes, R.O. (1992). Integrins: versatility, modulation and and signalling in cell adhesion. Cell 69: 11-25.

Iazaurraide, E., Lewis,J., McGuigan,J., Jankowska,M., Darzynkiewicz,E. and Mattaj,I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell 78: 657-668.

Ignotz, R., Endo, T., and Massague, J. (1987). Regulation of fibronectin and type I collagen mRNA leveles by transforming growth factor ß. J. Biol. Chem. 262: 6443-6446.

Ingber, D.E., Prusty,D., Frangioni,J.V., Cragoe,E.J., Lechene,C. and Schwartz,M.A. (1990). Control of intracellular pH and growth by fibronectin in capillary endothelial cells. J. Cell Biol. 110: 1803-18011.

Jakowlew, S.W., Mead,J.E., Danielpour,JJ.W., Roberts,A.B. and Fausto,N. (1991). Tgansforming growth factor-beta (TGF-b) isoforms in rat liver regeneration: messenger RNA expression and activation of latent TGF-beta. Cell Regul 2: 535-548.

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

Jarnagin, W.R., Rockey,D.C., Koteliansky,V.E., Wang,S.S. and Bissel,D.M. (1994). Expression of variant fibronectins in wound healing: cellular source and biological activity of the EIIIA segment in rat hepatic fibrogenesis. J. Cell. Biol. 127: 2037-2048.

Jeffreys, A.J. and Flavell, R.A. (1977). The rabbit beta globin gene contains a large insert in the coding sequence. Cell 12: 1097-1108.

Katz, R.A. and Skalka, A.M. (1990). Control of retroviral RNA splicing through maintenance of suboptimal processing signals. Mol. Cell. Biol. 10: 696-704.

Kenan, J. Query, C.C., and Keene J.D. (1991). RNA recognition: towards identifying determinants of specificity. TIBS 16: 214-217.

Keyes, L.N., Cline,T.W. and Schedl,P. (1992). The primary sex determination signal of drosophila acts at the level of transcription. Cell 68: 933-943.

Knowlton, A.A., Connelly,C.M., Romo,G.M., Mamuya,W., Apstein,C.S. and Brecher,P. (1992). Rapid expression of fibronectin in the rabbit heart after myocardial infarction with and without reperfusion. J. Clin. Inv. 89: 1060-1068.

Kohtz, J.D., Jamison,S.F., Will,C.L., Zuo,P., Luhrmann,R., Garcia-Blanco,M.A., and Manley,J.L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature 368: 119-124.

Komoriya, A., Green,L.J., Mervic,M., Yamada,S.,Yamada,K. and Humphries,M.J. (1991). The minimal essential sequence for a major cell type-specific adhesion site (CS1) within the alternatively spliced type III connecting segment domain of fibronectin is leucine-aspartic acid-valine. J. Biol. Chem. 266: 15075-15079.

Konarska, M.M., Padgett,R.A. and Sharp,P.A. (1983). Trans splicing of mRNA precursors in vitro. Cell 42: 165-171.

Konarska, M.M. and Sharp, P.A. (1986). Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Celll 46: 845-855.

Konarska, M.M. and Sharp, P.A. (1987). Interactions between snRNP particles in formation of spliceosomes. Cell 49: 763-774.

Kornblihtt, A.R., Vibe-Pedersen,K. and Baralle,F.E. (1983). Isolation and characterisation of cDNA clones for human and bovine fibronectins. Proc. Natl. Acad. Sci. USA 80: 3218-3222.

Kornblihtt, A.R., Umezawa,K., Vibe-Pedersen,K. and Baralle,F.E. (1985). Primary structure of human fibronectin; differential splicing can generate at least 10 polypeptides from a single gene. EMBO J. 4: 1755-1759.

Kornblihtt, A.R. and Gutman, A. (1988). Molecular biology of the extracellular matrix proteins. Biol. Rev., 63: Biol. Rev.,.

Krainer, A.R., Maniatis,T., Ruskin,B. and Green,M.R. (1984). Normal and mutant human b-globin pre-mRNA are faithfully and efficiently spliced in vitro. Cell 36: 993-1005.

Krainer, A.R., Conway,G.C. and Kozak,D. (1990). Purification and characterization pre-mRNA splicing factors SF2 from HeLa cells. Genes Dev. 4: 1158-1171.

Krawczak, M. Reiss, J. and Cooper,D.N. (1992). Human Gen. 92: 41-54.

Kudo, S. and Fukudo, M. (1989). Structural organization of glycophorin A and B genes: glycophorin B gene evolved by homologous recombination at Alu repeat sequeces. Proc. Natl. Acad. Sci. 86: 4619-4623.

Kuhne, T. Wieringa, B., Reiser,J. and Weissmann,C. (1983). Evidence against a scanning model of RNA splcing. EMBO J. 2: 727-733.

Kuo, H.C., Nasim,F.H. and Grabowski,P.J. (1991). Control of alternative splicing by differential binding of U1 small nuclear ribonucleoprotein particle. Science 251: 1045-1050.

Lamm, G.M., Blencowe,B.J., Sproat,B.S., Iribarrem,I.M., Ryder,U. and Lamond,A.J. (1992). Antisense probes containing 2-aminoadenosine allow efficent depletion of U5 snRNP from HeLa splicing extracts. Nucleic Acids Res. 19: 3193-3198.

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

Lang, K.M. and Spritz, R.A. (1983). RNA splice selection: evidence for a 5' to 3' scanning mechanism. Science 220: 1351-1355.

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

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

Lefcort, F., Venstrom,K., McDonald,J.A., and reichardt,L.F. (1992). Regulation of expression of fibronectin and its receptor a5b1, during development and regeneration of periphreral nerve. Development 116: 767-782.

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

Liao, X.C., Tang,J. and Rosbash,M. (1993). An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. Genes Dev. 7: 419-428.

Libri, D., Piseri,A. and Fiszman,M.J. (1991). Tissue specific splicing in vivo of the b-tropomyosin gene: dependence on an RNA secondary structure. Science 252: 1842-1845.

Ligtenberg, M.J.L., Gennisen,A.M.C., Vos,H.L. and Hilkens,J. (1990). A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. Nucleic Acid Res. 19: 297-301.

Linch, K.W. and Maniatis, T. (1995). Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes Dev. 9: 284-293.

Luhrmann, R., Kastner,B. and Bach,M. (1990). Structure of spliceosomal snRNPs and their role in pre mRNA splicing. Biochim. Biophys. Acta 1087: 265-292.

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

Magnuson, V.L., Young,M., Schattenbery,D.G., Mancini, M.A., Chen,D., Steffensen,B. and Klebe,R.J. (1991). The alternative splicing of fibronectin pre-mRNA is altered during ageing and in response to growth factors. J. Biol. Chem. 266: 14654-14662.

Mamuya, W.S. and Brecher, P. (1992). Fibronectin expression in the normal and hypertrophic rat heart. J. Clin. Inv. 89: 392-401.

Maniatis, T. (1991). Mechanism of alternative pre-mRNA splicing. Science 251: 33-34.

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

Mardon, H.J. and Sebastio, G. (1992). Regulation of alternative splicing in the IIICS region of human fibronectin pre-mRNA encoding cell binding sites CS1 and CS5. J. Cell. Sci. 103: 423-433.

Marvitt, J., DiLella,A.G., Brayton,K., Ledley,F.D., Robson,K.J.H. and Woo,S.L.C. (1989). GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylketonuria. Nucleic. Acid. Res. 15: 5613-5628.

Mathews, G.A. and ffrench-Constant, C. (1995). Embryonic fibronectins are up-regulated following peripheral nerve injury in rats. J. Neurobiol. 26: 171-188.

Mayeda, A. and Krainer, A.R. (1992a). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell 69: 365-375.

Mayeda, A., Zahler,A.M., Krainer,A.R. and Roth,M.B. (1992b). Two members of a conserved family of nuclear phosphoprotein are involved in pre-mRNA splicing. Proc. Natl. Acad. Sci. 89: 1301-1304.

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

Mayeda, A., Helfman,D.M. and Krainer,A.R. (1993). Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol. Cell. Biol. 13: 2993-3001.

McKeow. (1992). Alternative mRNA splicing. Annu. Rev. Cell Biol. 8: 133-55.

McNally, L.M. and McNally,M.T. (1995). SR proteins interact with a splicing inhibitor sequence from RSV. RNA processing meeting, CSH.

McPheeters, D.S. and Abelson, J. (1992). Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. Cell 71: 819-831.

Menzo, S., Bagnarelli,P., Giacca,M., Manzin,A., Varaldo,P.E., and Clementi,M. (1992). Absolute quantitation of viremia in HIV-infected asymptomatic subjects by competitive reverse-transcription and polymerase chain reaction. J. Clin. Microbiol. 30: 1752-1757.

Michaud, S. and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes Dev. 5: 2534-2546.

Michaud, S. and Reed, R. (1993). A functional association between the 5' and 3' splicing sites is estabilished in the earliest prespliceosome complex (E) in mammals. Genes Dev. 7: 1008-1020.

Mitchell, P.A., Urlaub,G. and Chasin,L. (1986). Spontaneous splicing mutations at the dihydrofolate reductase locus in Chinese hamster ovary cells. Mol. Cel. Biol 6: 1926-1935.

Moore, M.J. and Sharp, P.A. (1992). Site specific modification of pre-mRNA: the 2' hydroxyl groups at the splice site. Science 256: 992-997.

Moore, M.J., Query,C.C. and Sharp,P.A. Splicing of precursors to mRNA by spliceosome. In The RNA World, ed. R.F. Gesteland and J.F. Atkins. 303- 357. New YorkNew York: CSH Press, 1993.

Mosher, D. Fibronectin. San Diego: Academic Press Inc., 1989.

Mould, A.P., Wheldon,L.A., Komoriya,A., Wayner,E.A., Yamada,K.M. and Humphries,M.J. (1990). Affinity chromatographic isolation of the melanoma adesion receptor for the IIICS region of FN and its identification as the integrin a4-b1. J. Biol. Chem. 265: 4020-4024.

Murata, Y., Seo,H., Sekiguchi,K., Imai,T., Lee,J. and Matsui,N. (1990). Specific induction of fibronectin gene in rat liver by thyroid hormone. Mol. Endocrinol. 4: 693-699.

Murphy-Ullrich, j.E. and Mosher, D.F. Fibronectin and disease processes. In Extracellular Matrix Biochemistry. New York: Elsevier, 1986.

Nakai, K. and Sakamoto, H. (1994). Construction of a novel database containing aberrant splicing mutations of mammalian genes. Gene 141: 171-177.

Nakamura, T., Nakaijma,T.,Tsunoda,S., Nakada,S.,Tsurui,H. and Wada,A. (1992). Induction of E1A-responsive negative factors for transcription of the fibronectin gene in adenovirus E1-transformed rat level. J. Virol 66: 6436-6450.

Nandabalan, K., Price,L. and Roeder,S.R. (1993). Mutations in U1 snRNA bypass the requirement for a cell type-specific RNA splicing factor. Cell 73: 407-415.

Nelson, K.K. and Green, M.R. (1988). Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. Genes Dev. 2: 319-329.

Nelson, K.K. and Green, M.R. (1989). Mammalian U2snRNP has a sequence specific RNA-binding activity. Genes Dev. 3: 1562-1571.

Newman, A.J. and Norman, C. (1992). Mutations in yeast U5 snRNA alter the specificity of 5' splice site cleavage. Cell 65: 115-123.

Niwa, M. and Berget, S.M. (1991). Mutation of the AAUAAA polyadenilation signal depresses in vitro splicing of proximal but not distal introns. Genes and Dev. 5: 2086-2095.

Noble, J.C.S., Prives,C. and Manley,J.L. (1988). Alternative splicing of SV40 early pre-mRNA is determined by branch site selection. Genes Dev 2: 1460-1475.

Norton, P.A. and Hynes, R.O. (1987). Alternative splicing of chicken fibronectin in embryos and in normal and in transfected cells. Mol. Cell. Biol. 7: 4297-4307.

Odenthal, M., Neubauer,K, Baralle,F.E., Peters,H., Meyer,K.H. and Ramadori,G. (1992). Rat hepatocytes in primary culture synthesize and secrete cellular fibronectin. Exp. Cell Res. 203: 289-296.

Ohno, K. and Suzuki, K. (1988). Multiple abnormal beta-hexaminidase alpha-chain mRNAs in a compound heterozygous Ashkenazi Jewish patient wit Tay Sachs disease. J. Biol. Chem. 263: 18563-18567.

Owens, R., Kornblihtt,A.R. and Baralle,F.E. Fibronectin, the generation of multiple pypeptides from a single gene. In Oxford surveys on eukariotic genes, ed. N. McLean. 141-160. Oxford: Oxford University Press, 1986.

Oyama, F., Murata,Y., Suganuma,N., Kimura,T., Titani,K. and Sekiguchi,K. (1989). Patterns of alternative splicing of fibronectin pre-mRNA in human and fetal tissues. Biochemistry 28: 1428-1434.

Oyama, F., Hirohashi,S., Shimosato,Y. Titani,K. and Sechiguchi,K. (1989). Deregulation of alternative splicing of fibronectin pre-mRNA in malignant human liver tumors. J. Biol. Chem. 264: 10331-10334.

Padget, R.A., Hardy,S.F. Aand Sharp,P.A. (1983). Splicing of adenovirus RNA in a cell-free transcription system. Proc. Natl. Acad. Sci. 80: 5230-5235.

Pagani, F, Zagato,L., Vergani,C., Casari,G., Sidoli,A. and Baralle,F.E. (1991). Tissue-specific splicing pattern of fibronectin messenger RNA precursor during development and ageing in rat. J. Cell. Biol. 113: 1223-1229.

Pagani, F., Zagato,L., Majer,J.A., Ragnotti,G., Coviello,D.A. and Vergani,C. (1993). Expression and alternative splicing of fibronectin mRNA in human diploid endothelial cells during aging in vitro. Biochim. Biophys. Acta 1173: 172-178.

Paolella, G., Barone,M.V. and Baralle,F.E. Fibronectin. In Extracellular Matrix, ed. M. and Reid Zern L.M. 3-24. New YorK: Marcel Dekker, 1993.

Patel, R.S., Odermatt,E., Schwarzbauer,J.E. and Hynes,R.O. (1986). Organization of the rat fibronectin gene provides evidence for 'exon shuffling' during evolution. EMBO J. 6: 2565-2572.

Patton, J.G., Porro,E.B., Galceran,J., Tempst,P. and Nadal-Ginard. (1993). Cloning and characterisation of PSF, a novel pre-mRNA splicing factor. Genes Dev. 7: 393-406.

Paul, J.I., Schwarzbauer,J.E., Tankum,J.W. and Hynes,R.O. (1986). Cell-type-specific fibronectin subunits generated by alternative splicing. J. Bil. Chem. 261: 12258-12265.

Pershbacher, M.D. and Rouslahti, E. (1984). Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature 309: 30-33.

Peterson, M.L., Bryman,M.B., Peiter,M. and Cowan,C. (1994). Exon size affects competition between splicing and cleavage-polyadenylation in the immunoglobulin m gene. Mol. Cell. Biol. 14: 3434-3445.

Polly, P. and Nicholson, R.C. (1993). Sequence of the mouse fibronectin-encoding gene promoter region. Gene 137: 353-354.

Ramadori, G., Knittel,T., Odenthall,H., Schwogler,S., Neubauer,K. and Meyer zum Buschenfelde,K.H. (1992). Gastroenterology 103: 11313-1321.

Ramchatesingh, J., Zahler,A.M., Neugebauer,K.M., Roth,M.B. and Cooper,T. (1995). A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. Mol. Cell. Biol. 15:

Reed, R. and Maniatis, T. (1986). A role for exon sequences and splice site proximity in splicie site selection. Cell 46: 681-690.

Reich, C.I., VanHoy,R.W., Porter,G.I. and Wise,J.A. (1992). Mutations at the 3' splice site can be supressed by compensatory base pair changes in U1 snRNA in fission yeast. Cell 69: 1159-1169.

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

Roscigno, R.F., Weiner,M. and Garcia-Blanco,M.A. (1993). A mutational analysis of the polypyrimidine tract of introns. J. Biol. Chem 268: 11222-11229.

Roth, M.B., Zahler,A.M. and Stolk,J.A. (1991). A conserved phamily of nuclear phosphoproteins localized to sites of polymerase II transcription. J. Cell. Biol. 115: 587-596.

Rouslahti, E. (1988). Fibronectin and its receptors. Annu. Rev. Biochem. 57: 375-413.

Ruby, S.W. and Abelson, J. (1991). Pre-mRNA splicing in yeast. Trends Genet. 7: 79-85.

Ruskin, B., Zamore,P.D. and Green,M.R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. cell 52: 207-219.

Sakamaki, K., Tomonaga,M., Tsukui,K. and Nagata,S. (1989). Molecular cloning and characterisation of a chromosomal gene for human eosinophil peroxidase. J. Biol. Chem. 264: 16828-16836.

Scherly, D., Boelens,W., Dathan,N.A., van Venrooij,W.J., Mattaj,I.W. (1990). Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B" and their cognate RNAs. Nature 345: 502-506.

Schmid, S.R. and Linder, P. (1992). DEAD protein family of putative RNA helicases. Mol. Microbiol. 6: 283-291.

Schwarzbauer, J.E., Tamkun,J.W., Lemischka,I.R. and Hynes,R.O. (1983). Three different fibronectin mRNAs arise by alternative splicing within the coing region. Cell 35: 421-431.

Schwarzbauer, J.E., Patel,R.S., Fonda,D. and Hynes,R.O. (1987). Multiple sites of alternative splicing of the rat fibronectin gene. EMBO J. 6: 2573-2580.

Schwarzbauer, J.E., Spencer,C.S. and Wilson,C.L. (1989). Selective secretion of alternatively spliced fibronectin variants. J. Cell. Biol. 109: 3445-3453.

Seidman, J.G. and Leder, P. (1980). A mutant immunoglobulin light chain is formed by aberrant DNA and RNA splicing events. Nature 286: 779-783.

Seraphin, B. and Rosbash, M. (1989). Identification of functional U1 snRNP-pre-mRNA complexes committed to spliceosome assembly and splicing. Cell 59: 349-358.

Seraphin, B., Kretzner,l. and Rosbash,M. (1991). Genetic depletion indicates a late role for U5 snRNP during in vitro spliceosome assembly. Nucleic Acid Res. 19: 3857-3860.

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

Sharp, P.A. (1994). Split genes and RNA splicing. Cell 77: 805-815.

Shelley, C.S., Colin,R.S., Baralle,F.E. and Shoulders,C.C. (1985). Comparison of the human apolipoprotein genes. J. Mol. Biol. 186: 43-51.

Siebel, C.W., Fresco,L. and Rio,D.C. (1992). The mechanism of somatic inhibition of Drosophila P-element pre-mRNA splicing: multiprotein complex at an exon pseudo-5' splice site control U1 snRNP binding. Genes Dev. 6: 1386-1401.

Siebel, C.W., Kanaar,R. and Rio,D.C. (1994). Regulation of tissue specific P-element pre-mRNA splicing requires the RNA-binding protein PSI. Genes Dev. 8: 1713-1725.

Sikder, S.K., Kabat,E.A. and Morrison,S.L. (1985). Alternative splicing patterns in an aberrantly rearranged immunoglobulin Kappa-light chain gene. Proc. Natl. Acad. Sci. 82: 4045-4049.

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

smith, C.W.J., Porro,E.B., Patton,J.G. and Nadal-Ginard,B. (1989b). Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. Nature 342: 243-247.

Sontheimer, E.J. and Steitz, J.A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science 262: 1989-1996.

Southern, E. (1975). Detection of specific equences among DNA fragments separated by gel electrophoresys. J. Mol. Biol. 98: 503.

Staffa, A and Cochrane, A. (1995). Identification of positive and negative splicing regulatory alaments within the terminal tat-rev exon of human immunodeficiency virus type 1. Mol. Cel. Biol. 15: 4597-4605.

Staknis, D and Reed, R. (1994). SR proteins promote the first specific recognition of the pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol 14: 7670-7682.

Steingrimsdottir, H., Rowley,G., Dorado,G., Cole,J., and Lehmann,A.R. (1992). Mutations which alter splicing in the human hypoxanthine-guanine phosphorybosyltransferase gene. Nucleic Acid Res. 20: 1201-1208.

Sun, Q., Mayeda,A., Hampson,R.K., Krainer,A.R. and Rottman,F.R. (1993). General splicing factor ASF/SF2 promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. 7: 2598-2608.

Sun, Q., Dirksen,W.P. and Rottman,F.M. (1995). An exhonic splicing repressor in bovine growth hormone pre-mRNA. RNA Processing Meeting, CSH.

Talerico, M. and Berget, S.M. (1990). Effect of 5' splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. 10: 6299-6305.

Talerico, M and Berget, S.M. (1994). Intron definition in splicing of small drosophila introns. Mol. Cell. Biol. 14: 3434-3445.

Tanaka, K., Watakabe,A., and Shimura,Y. (1994). Polypurine sequence within a downstream exon function as a splicing enhancer. Mol. Cell. Biol. 14: 1347-1354.

Tian, M. and Maniatis, T. (1992). Positive controls of pre-mRNA splicing in vitro. Science 256: 237-240.

Tian, M. and Maniatis, T. (1993). A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. Cell 74: 105-114.

Treisman, R., Proudfoot,N.J., Shander,M. and Maniatis,T. (1982). A single-base change at a splice in a b-thalassemic gene causes abnormal RNA splicing. cell 29: 903-911.

Tsukahara, T.C., Casciato,C. and Helfman,D.M. (1994). Alternative splicing of b-tropomyosin pre-mRNA: multiple cis-elements can contribute to the use of the 3'- and 5'- splice sites of the nonmuscle/smooth muscle exon 6. Nuclei Acid Res. 22: 2318-2325.

Utans, U., Behrens,S.E., Luhrmann,R. and Kramer,A. (1992). A splicing factor that is inactivated during in vivo heat shock is functionally equivalent to the [U4/6.U5] three snRNP -specific protein. Genes Dev 6: 631-641.

Valcarcel, J., Singh,R., Zamore,P.D. and Green,M.R. (1993). The protein sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature 362: 171-175.

van Oers, C.C.M., Adema,G.J., Zandberg,H., Moen,T.C. and Baas,P.D. (1994). Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. Mol. Cell. Biol. 14: 951-960.

Vartio, T., Laitinen,L., Narvavev,O., Cutollo,M., thornell,L.E., Zardi,L. and Virtanen,I. (1987). Differential expression of the ED sequence containing form of cellular fibronectin in embryonic and adult tissue. J. Cell. Sci. 88: 419-430.

Vellard, M., Sureau,A., Soret,J.,Martinerie,C. and Perbal,B. (1992). A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myb exon. Proc. Natl. Acad. Sci. 89: 2511-2515.

Vibe-Pedersen, K., Kornblihtt,A.R. and Baralle,F.E. (1984). Expression of a human a-globin/fibronectin gene hybrid generates two mRNAs by alternative splicing. EMBO J. 3: 2511-2516.

Vilardell, J. and Warner, J.R. (1994). Regulation of dplicing at an intermediate step in the formation of the spliceosome. Genes Dev. 8: 211-220.

Wang, A.M., Doyle,M.V. and Mark,D.F. (1989). Quantitation of mRNA by the polymerase chain reaction. Proc. Natl. Acad. Scie. 86: 9717-9721.

Wang, Z., Hoffmann,H.M. and Grabowski,P.J. (1995). Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. RNA 1: 21-35.

Watakabe, A., Tanaka,K. and Shimura,Y. (1993). The role of exon sequences in splice site selection. Genes Dev. 7: 407-418.

Weil, D., Bernard,M., Combates,N., Wirtz,M.K., Hollister,D.W., Steinmann,W. and Ramirez,F. (1988). Identification of a mutation that causes exon skipping during collagen pre-mRNA splicing in an Ehlors-Danlos syndrome variant. J. Biol. Chem. 263: 8561-8564.

Wieringa, B., Meyer,F., Reiser,J. and Weissmwnn,C. (1983). Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit b-globin gene. Nature 301: 38-43.

Wilson, C.L. and Schwarzbauer, J.E. (1992). The alternatively spliced V region contributes to the differential incorporation of plasma and cellular fibronectin into fibrin clots. J. Cell. Biol. 119: 923-933.

Wolff, T. and Bindereif, A. (1993). Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first step of splicing. Genes Dev. 7: 1377-1389.

Wu, J. and Manley, J.L. (1989). Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev. 3: 1553-1561.

Wu, J. and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell 75: 1061-1070.

Xu, R., Teng,. and Cooper,T.A. (1993). The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. Mol. Cell. Biol. 13: 3660-3674.

Zahler, A.M., Lane,W.S., Stolh,J.A. and Roth,M.B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev. 6: 837-847.

Zahler, A.M., Neugebauer,K.M., Stolk,J.A. and Roth,M.B. (1993a). Human SR proteins and isolation of a cDNA encoding SRp75. Mol. Cell. Biol. 13: 4023-4028.

Zahler, A.M. Neugebauer, K.M., Lane,W.S. and Roth,M.B. (1993b). Distinct functions of SR proteins in alternative pre-mRNA splicing. Science 260: 219-222.

Zahler, A.M. and Roth, M.B. (1995). Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5′ splice sites. Proc. Natl. Acad. Sci., 92: 2642-2646.

Zamore, P.D. and Green, M.R. (1991). Biochemical characterisation of U2 sn RNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intranuclear distribution. EMBO J. 10: 207-214.

Zardi, L., Carnemolla,B., Siri,A., Petersen,T.E., Paolella,G., Sebastio,G. and Baralle,F.E. (1987). Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon. EMBO J. 6: 2337-2342.

Zheng, Z.M. and Baker,C.C. (1995). Identification of exonic splicing enhancers in the bovine papilloma virus type 1 early region. RNA Processing Meeting, CSH.

Zhuang, Y., Goldstein,A.M. and Weiner,A.M. (1989a). The natural 5' splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. Mol. Cell. Biol. 7: 3018-3020.

Zhuang, Y., Goldstein,A.M. and Weiner,A.M. (1989b). UACUAAC is the preferred branch site for mammalian mRNA splicing. Proc. Natl. Acad. Sci. 86: 2752-2756.

Zuo, P. and Manley, J.L. (1993). Functional domains of the human splicing factor ASF/SF2. EMBO J. 12: 4727-4737.

Zuo, P. and Manley, J.L. (1994). The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. Proc. Natl. Acad. Sci. 91: 3363-3367.

Most of the work exposed in this thesis has been published in the following articles:

Caputi, M., Casari,G., Guenzi,S., Tagliabue,R., Sidoli,A., Melo,C.A. and Baralle,F.E. (1994). A novel bipartite splicing enhancer modulates the differential processing of the human FN EDA exon. Nucleic Acid Res. 22: 1018-1022.

Caputi, M., Melo,C.A., and Baralle,F.E. (1995a). Regulation of Fibronectin expression in rat regenerating liver. Nucleic Acid. Res. 23: 238-243.

Caputi, M., Baralle,F.E. and Melo,C.A. (1995b). Analysis of the linkage between FN alternative spliced. sites during ageing in rat tissues. Biochim. Biophis. Act. 1263: 53-59.