ALTERNATIVE SPLICING OF THE *GRIN1* CI CASSETTE EXON: SILENCING MECHANISM

by

Kyoungha Han

BA, Seoul National University, 1993

MS, Seoul National University, 1995

Submitted to the Graduate Faculty of

Faculty of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Kyoungha Han

It was defended on

August 16, 2004

and approved by

Karen M. Arndt

Jeffrey L. Brodsky

Craig L. Peebles

John L. Woolford

Paula J. Grabowski
Dissertation Director

ALTERNATIVE SPLICING OF THE *GRIN1* CI CASSETTE EXON: SILENCING MECHANISM

Kyoungha Han, PhD

University of Pittsburgh, 2004

In higher eukaryotes, alternative splicing is used extensively to regulate gene expression in a cell- and developmental stage-specific manner. Regulation affects the ratio of protein isoforms and the functional diversity of the corresponding proteins. Although its importance has been underscored, control mechanisms of alternative splicing are poorly understood. The N-methyl-D-aspartate receptor NR1 subunit (NMDAR1, *GRIN1*) contains three cassette exons (NI, CI and CII), and which are specifically expressed in mammalian brain in various combinations. In this study, a minigene splicing reporter system containing the CI cassette exon and flanking exon – intron region was utilized to study the control mechanism of alternative splicing in mammalian cell lines. This work focuses primarily on the identification of exonic UAGG motifs and a 5' splice site region G cluster. Individual motifs play a silencing role for splicing but the combination of all three is required for strong silencing. UV-crosslinking shows that the G cluster interacts with hnRNP F or H/H' proteins and the UAGG motif interacts with hnRNP A1 protein. Coexpression of hnRNP F or H/H' switches the major splicing pattern from skipping to inclusion of the wild type splicing reporter and this activity is dependent in larger part on the G cluster at the 5' splice site. Surprisingly, the strong silencing requires the G cluster although no direct interaction with hnRNP A1 is observed. Insertion of an extra UAGG motif in the exon shows almost complete

silencing, suggesting that splicing control is dependent on the number of silencers. Removal of all three silencers results in complete loss of splicing repression. Therefore, the UAGG and G cluster motifs confer flexibility of control of CI cassette exon splicing and the number of the silencer motifs is likely to determine the silencing strength in various tissues. The splicing silencing mechanism defined here is an important aspect of the tissue-specificity of the CI cassette exon, which has previously been shown to involve the enhancing function of the mouse neuroblastoma apoptosis-related RNA-binding protein (NAPOR). Using RT-PCR analysis from candidates found by a splice site database search of the human and mouse genome, new groups of alternative exons containing the UAGG and GGGG (as a G cluster) motifs (Gene ID: *HNRPH1*, *HNRPH3*, *NCOA2*, *MEN1* and *Hp1bp3*) were identified. Autoregulation of hnRNP H (*HNRPH1*) mediated by the G cluster motif, which is a positive regulator of the CI cassette exon, an implication of these results. Overall, this work provides evidence for a splicing silencing mechanism that controls brain-region specific CI cassette exon by exonic UAGG and 5' splice site G cluster motifs. This work also shows that the motif pattern can be used computationally to identify additional skipped exons that contain combinations of UAGG and GGGG motifs, but are otherwise unrelated to the CI cassette exon.

ACKNOWLEDGEMENTS

I am immeasurably grateful to Dr. Paula Grabowski for giving me the opportunity to work in her lab. She has been a wonderful research advisor. I'll always remember the white roses of truth, virtue and knowledge in my mind.

I would also like to thank my committee members; Dr. Craig Peebles, Dr. Jeffrey Brodsky, Dr. Karen Arndt and Dr. John Woolford for their time and help along the way.

I'm grateful to my collaborators Gene Yeo and Dr. Chris Burge for their efforts and support. I have enjoyed working together with them. I thank Serkan Alkan, Dr. Christine Milcarek, Dr. Gideon Dreyfuss and Dr. Jungskik Yong for their materials and discussion.

My thanks to all of the members of the Grabowski lab, past and present, who have helped me.

I thank Jennifer Bayracktar for her friendship as my favorite classmate. My thanks also to Brandi Baros, Shruthi Vembar, Kathryn Kumer, Stephanie Winter, Paul Sapienza and Bob Youker for helping me and having a good time. I particularly thank Paul Cantalupo and Dr. Nick Quintyne for helping me with my scientific writing.

A special thank you to my family and friends for their continued love and support. I would particularly like to acknowledge my parents, Guirea and Moonhee Han, and my brothers, Jaehyung and Kyongkeun, since they have always been with me in everything. Also, I am thankful to Myungkeun Hong for her eternal friendship, Chulho Choi for his advice and am especially thankful to Yuyong Choi for his encouragement. I could not have achieved any of this without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Alternative pre-mRNA splicing: biochemistry and tissue-specific regulation

The central dogma is an essential pathway of the flow of genetic information, as Francis Crick proclaimed in 1958. The discovery of interruptions in eukaryotic genes by non-coding intervening sequences (introns) in 1977 launched a new understanding for the process of gene expression (Berget et al., 1977). The removal of introns from a precursor mRNA (pre-mRNA) to produce a mature mRNA is required for the expression of eukaryotic genes through the process called RNA splicing. The process of pre-mRNA splicing is a chemical reaction catalyzed by a protein and RNA complex, termed the spliceosome. Significant changes in processed mRNA could occur in a variety of patterns by inserting or deleting exons, generating multiple isoforms from a single gene in various tissues, thereby increasing the diversity of function from the encoding gene (Figure 1A). Therefore, pre-mRNA splicing is not only an essential process to produce functional gene products, but it is also a critical step of post-transcriptional regulation. RNA splicing is thought to take place co-transcriptionally, and recent research shows that splicing factors remain on the mRNA after splicing and couple pre-mRNA splicing with mRNA export and localization. Interestingly, findings of the Human Genome Project suggest that a lower number of genes exist in humans than previously proposed. The public database represents over 1.5 million human expressed sequence tags (ESTs) from various tissues, while the number of proteins are predicted to be 0.2-2 million (Brett et al., 2000; Service, 2001; Roberts and Smith, 2002). Therefore, the proteomic complexity is 10-100 fold greater

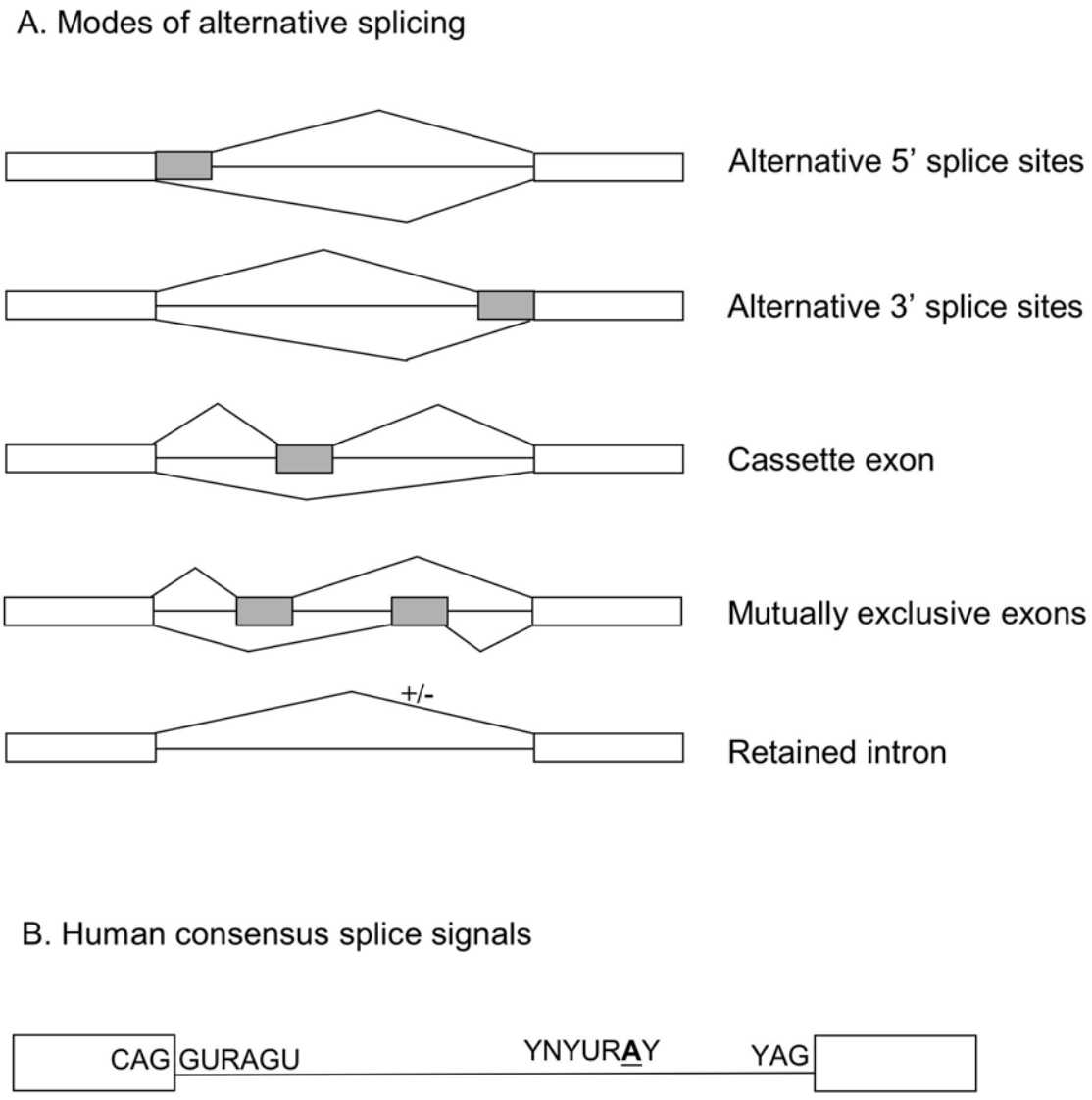


Figure 1. Modes of alternative splicing and splice signals.

(A) Schematic of the patterns of alternative splicing (adapted from Gravely, 2001). Boxes indicate exons: white boxes, constitutive exons; gray, alternative exons. Lines represent introns. (B) Consensus sequences of the splicing signals (Burge et al., 1999). Exon sequences are in boxes. Intron sequences are above the line. A branch site is underlined in bold. A, adenosine; C, cytidine; G, guanosine; U, uridine; R, purine; Y, pyrimidine.

than the estimated 32,000 genes (Lander et al., 2001). Bioinformatics analysis estimates that alternative splicing occurs in 35% to 75% of human genes (Brett et al., 2000; Lander et al., 2001; Johnson et al., 2003). Alternative splicing is therefore the principal mechanism to generate mRNA diversity from the limited number of genes. For example, the calcium-activated potassium channel gene, *slo,* is thought to mediate hair cell tuning by its alternative splicing, which generates over 500 different mRNAs of splice variants (Black, 1998). Thus, a current challenge is to understand how tissue-specific alternative splicing events occur in various tissues and how the flexibility of splicing is related to human diseases. An additional issue is to identify alternative exons on a genome-wide scale.

1.1. Specificity of RNA splicing signals

A major question about splicing reactions is how the splicing machinery recognizes splice sites in pre-mRNAs, because only the GU at the 5' splice site and the AG at the 3' splice site are rigorously conserved in the consensus sequences. Splice signals required for splicing include the 5' splice site, the 3' splice site and the branch point, and their consensus sequences are mainly in the intron region (Figure 1B). A major class of introns, starting with GU at the 5' splice site and ending with AG at the 3' splice site, is processed by the U2-type spliceosome, thus named U2-type introns (AG-GT rule). Although splice signals are highly conserved from yeast to vertebrates, the consensus sequences are short and relatively degenerate (Burge et al., 1999). In addition, since human genes have multiple introns of normally larger sizes (3,365 nt on average, Lander et al., 2001), potential splice signals frequently occur. In humans, the

5’ splice site consensus sequence is CAG:GURAGU (where ‘:’ denotes the exon:intron boundary); the branch point, YNYURAY (branch site is underlined); the 3’ splice site, YAG: (Figure 1B, Burge et al., 1999). For example, the consensus sequence of the 5’ splice site is a maximum of 9 nucleotides and the sequence is degenerate except for the GU sequence at the junction. Thus, the amount of information needed to specify 5’ splice junctions is not sufficient in most human pre-mRNA (Burge et al., 1999). The 3’ splice signal is specified by the polypyrimidine rich intron region upstream from the 3’ splice site as well as the consensus sequence at the 3’ splice site. However, the polypyrimidine rich region is variable, and the consensus sequence of the 3’ splice site is shorter than the 5’ splice site. These signals are unlikely to contribute to the specification of the 3’ splice site. Therefore, the splicing machinery in vertebrates requires auxiliary sequence elements other than the splice sites and the branch point in order to specify the splice junction (Burge et al., 1999). In contrast to human introns, yeast introns are short, and the consensus sequence of splicing signals is highly conserved. The splice signals contain all or most of the information for recognition of the consensus sequence at splice junctions, thus might be sufficient for spliceosome assembly and the splicing reaction. However, in humans, not only the consensus sequences at the splice sites, but also base compositional differences between exons and introns, open reading frame information, and interactions of *trans*-acting factors involved in transcription and pre-mRNA processing should be critical (McCracken et al., 1997; Bentley, 1999; Hirose et al., 1999, reviewed in Reed, 2003; Hirose and Manley, 2000).

Distinct from the U2 type introns, another group of metazoan introns (AT-AC

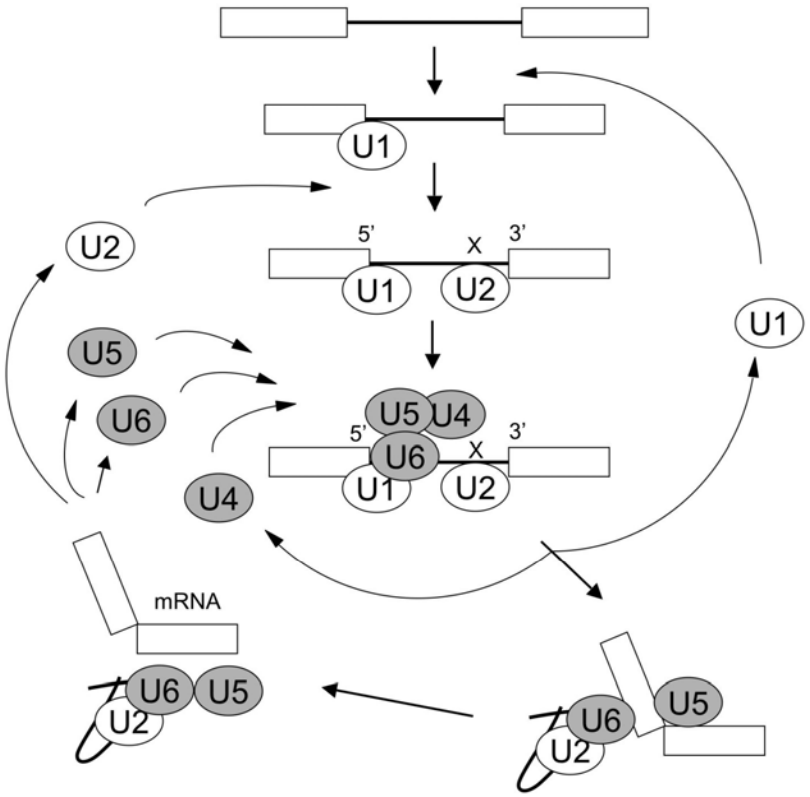
introns) has been identified, where the consensus sequences of splice sites are strongly conserved (Jackson, 1991; Hall and Padgett, 1994, reviewed in Sharp and Burge, 1997). The consensus sequence of the 5' splice site of this class is :AUAUCCUUU, the branch point UUCCUURACYCY and the 3' splice site YCCAC but these introns lack a polypyrimidine tract (Hall and Padgett, 1996; Tarn and Steitz, 1996b; Sharp and Burge, 1997). These introns are recognized by low abundant U11 and U12 snRNA groups and are spliced by the U12-type spliceosome including U5, U4atac and U6atac (Hall and Padgett, 1996; Tarn and Steitz, 1996a; Tarn and Steitz, 1996b). Although these two classes of introns are recognized by distinct sequences of snRNAs, the splicing mechanisms are conserved (Wu and Krainer, 1996; Dietrich et al., 1997). U5 snRNA is common to both spliceosomes and the interactions with other snRNPs are similar. Both sets of snRNAs bind the same core Sm proteins. The significance of U12-type splicing could be to confer a flexibility of splicing (or splicing regulation) in diverse conditions since cells could take advantage of two distinct splicing systems in order to regulate gene expression in response to varied environmental signals.

1.2. The spliceosome assembly and the snRNP recycle

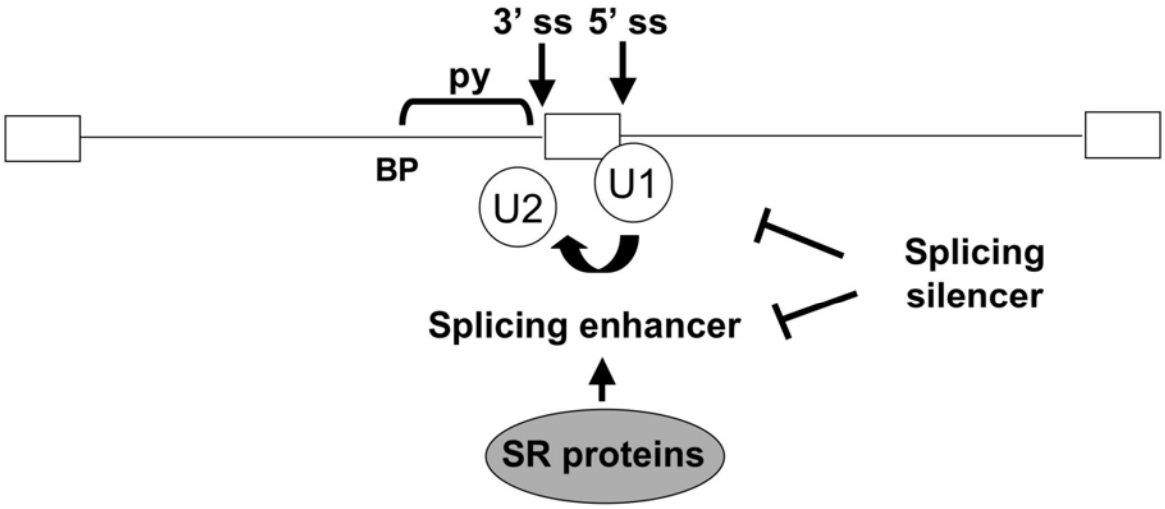
Pre-mRNA splicing is a biochemical reaction catalyzed by the spliceosome which a highly complex macromolecule containing core Sm proteins, small nuclear ribonucleoprotein particles (snRNPs U1, U2, U4, U5 and U6) and numerous non-snRNP proteins (Hartmuth et al., 2002; Zhou et al., 2002; Jurica and Moore, 2003, reviewed in Moore et al., 1993; Burge et al., 1999). In humans, approximately 145 spliceosomal proteins are identified in the spliceosome (Zhou et al., 2002). Assembly of the

spliceosome requires the recognition of splice signals: the 5’ splice site, the 3’ splice site and the branch point (Reed, 1996). Pre-mRNA interacts with heterogeneous nuclear ribonucleoprotein (hnRNP) proteins forming the H complex before the commitment step, which is not associated with snRNPs and other splicing factors. Spliceosome assembly initially occurs by recognition of the 5’ splice site (Figure 2A) and the complexes at different stages of assembly can be discriminated *in vitro*. For a major class of introns, E complex is formed by association of U1 snRNP with the 5’ splice site, and is converted to A complex when U2 is stably associated with the 3’ splice site. Recruiting a U4/U5/U6 tri-snRNP particle forms B complex, and a final catalytic complex is C compex containing U2, U5 and U6 snRNPs. Spliceosome assembly is initiated by recognition of the 5’ splice site by U1 snRNP through base pairing of U1 snRNA (Zhuang and Weiner, 1986; Michaud and Reed, 1991). Although U1 snRNP association with the 5’ splice site is required, it is not sufficient to specify the splice site (Seraphin et al., 1988; Seraphin and Rosbash, 1990). The branch site is initially recognized by U2AF65, the non-snRNP protein, not by U2 snRNP (Ruby and Abelson, 1988; Rosbash and Seraphin, 1991). This protein interacts with U2AF35 and U1 snRNP 70K proteins, thereby promoting U2AF and U1 snRNP binding to splice sites (reviewed in Chabot, 1996; Lopez, 1998; Tacke and Manley, 1999; Graveley, 2000). U1 snRNP binds to the 5' splice site, splicing factor 1 (SF1/BBP) binds to the branch point, and the 65 and 35 kDa subunits of U2 auxiliary factor (U2AF) bind to the polypyrimidine tract and 3’ splice site to form the early (E) complex (Bennett et al., 1992; Abovich and Rosbash, 1997) in vertebrates. This assembly is ATP-independent (Michaud and Reed, 1991). The formation of the E

A. Splicesome assembly



B. Exon definition model



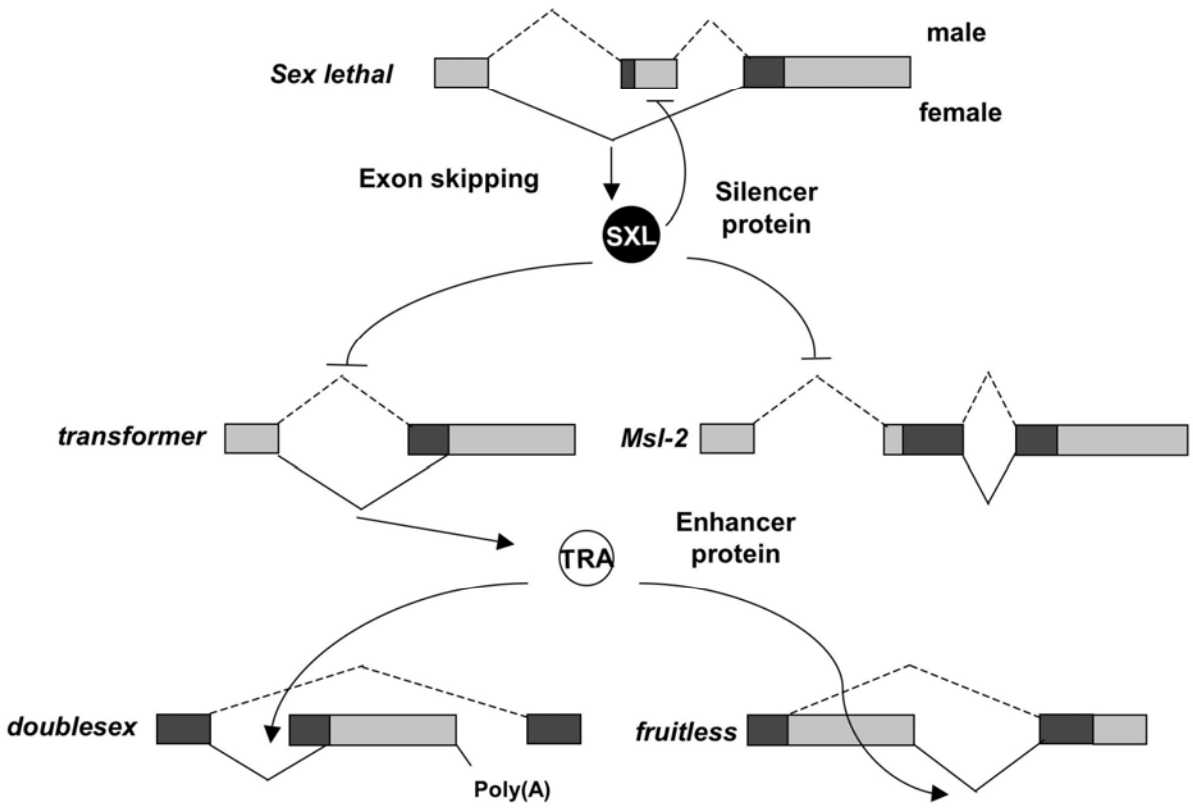


Figure 2. Spliceosome assembly and cascades of alternative splicing.

(A) Exons and introns are indicated by boxes and lines. Circles represent snRNPs. A major class of introns are shown as substrates of the U2-type spliceosome (Burge et al., 1999). (B) Exon definition model. Splice signals are shown; 5'ss, 5' splice site; 3'ss, 3' splice site; BP, branch point; py, polypyrimidine rich region. Circles represents proteins; U1 snRNP, U2 snRNP, SR proteins. (C) Sex determination in *Drosohphila*. Pathway for deferentiation of Female is shown. Dotted lines indicate splicing for male sex determination.

complex is a commitment step of splicing reactions, thus it is an important step of splicing regulation. A U2 snRNP complex recognizes the polypyrimidine tract in the upstream intron across the exon, which is promoted by U1 snRNP (Barabino et al., 1990; Seraphin and Rosbash, 1990). Although U2 snRNP is able to recognize the branch point in the absence of U1 snRNP, it is not efficient for most pre-mRNAs (Query et al., 1997). E complex is converted to A complex by stable association of U2 snRNP with the branch point. This transition is rapid and ATP-dependent (Liao et al., 1992). The putative ATP-dependent RNA helicase Prp5 (human homologue, hPrp5) is thought to mediate ATP hydrolysis (O'Day et al., 1996). Stable association of U2 snRNP with branch point requires human ternary splicing factor complex SF3a (Bennett et al., 1992; Kramer et al., 1995, reviewed in Burge et al., 1999). Then the B complex is formed by association of a U4/U5/U6 tri-snRNP particle with the spliceosome in an ATP-dependent manner (Cheng and Abelson, 1987; Konarska and Sharp, 1987; Reed and Palandjian, 1997). Finally, the spliceosome is rearranged to form the catalytically active C complex for the first step of splicing, where the phosphodiester bond is broken at the junction between the exon and the 5' splice site and the lariat is formed (Teigelkamp et al., 1994; Roy et al., 1995). The catalytic spliceosome is associated with U6, U2 and U5 snRNPs whereas U1 and U4 become destabilized. U4 plays a role in bringing U6 to the catalytic site of the spliceosome and blocking undesired catalysis by base pairing with U6 (Guthrie and Patterson, 1988). The catalytic site is created by base pairing of U6 with the 5' splice site by replacing U1, interaction of U6 with U2, and bridging of the splice sites by U5. The second step of splicing joins the exons by forming the phosphodiester bond and detaching the lariat intron (Horowitz and Krainer, 1997; Zhou

and Reed, 1998). Both steps of catalysis require ATP-dependent conformational rearrangements. By release of the ligated exons and the lariat intron, the splicing reaction is completed. The spliceosome disassembly ensues snRNP recycle for the next round and the lariat intron is subsequently degraded.

1.3. The exon definition model

In vertebrates, initial recognition of splice sites is promoted by interactions between the 3' and 5' splice sites across the exons that are short and separated by long introns, which is termed exon definition (Figure 2B, reviewed in Rio, 1993; Berget, 1995; Black, 1995; Adams et al., 1996; Reed, 1996). For human pre-mRNAs, on average, the exons are 145 nt and the introns 3,365 nt in their length (Lander et al., 2001). In this model, the splicing machinery defines an exon as a unit due to interactions of splicing factors spanning the short exon at the early step of spliceosome assembly (Robberson et al., 1990). The efficiency depends upon the strengths of the adjacent splice sites as well as auxiliary RNA control elements: this step is a major target for regulating alternative splicing. Based on the exon definition model, the strength of a 5’ splice site is predicted to influence the efficiency of 3’ splice site recognition. Enhanced binding of U1 snRNP to the 5’ splice site increases exon inclusion of rat preprotachykinin exon 4, which is mediated by activation of the upstream intron splicing (Kuo et al., 1991; Grabowski et al., 1991). Mutations at the ends of an internal exon show more exon skipping than mutation at the 5’ splice site in the three exon construct, containing identical intron and exon sequences (Talerico and Berget, 1990). This suggests that exons, rather than splice sites, are the recognition units for assembly of the

spliceosome. Consistent with these results, binding of U2AF65 to the 3’ splice site is prompted by association of U1 snRNP with the 5’ splice site in UV crosslinking analysis (Hoffman and Grabowski, 1992). These interactions bridging an exon are thought to activate pre-mRNA splicing mediated by the RS domains of SR proteins functioning (Wu and Maniatis, 1993; Graveley and Maniatis, 1998). Metazoan 5' and 3' splice sites are efficiently trans-spliced when the 3' trans-splicing substrate contains a downstream exonic enhancer or a 5’ splice site (Chiara and Reed, 1995), This suggests that the 3’ splice site are specified by exon definition through 5' and 3' splice-site pairing in metazoans.

One prediction of this model is that exon length should be restricted. Shortening the exon leads to a decrease in exon inclusion, probably due to steric hindrance and loss of exonic regulatory sequences (Dominski and Kole, 1991; Dominski and Kole, 1992). In addition, increasing the exon length inhibits spliceosome assembly (Robberson et al., 1990). Although most vertebrate exons belong to the exon definition model, many small exons, called microexons, are constitutively spliced. For example, the chicken cardiac troponin T exon 17 contains only 6 nucleotides and is constitutively spliced (Carlo et al., 2000). This constitutive exon inclusion depends on six copies of an intronic splicing enhancer downstream of the target exon, which is a binding site for the bridging mammalian splicing factor SF1. Therefore, spliceosome assembly can occur on the extended exon by SF1, thus enhancing recognition of the upstream exon. Splicing of a minor class of introns flanking classical splice signals is another issue for exon definition because these introns occur in between a major class of introns. Human voltage-gated skeletal muscle sodium channel α subunit (SCN4A) intron 21 is an AT-AC

type intron and its splicing is stimulated by the 5’ splice site of the downstream conventional intron (Wu and Krainer, 1996). This effect is dependent on U1 snRNA, indicating that splicing factors can mediate interactions across the exon between major and minor classes of introns. In this model, a challenging question is how exons, which are separated by a long intron, are brought together by the spliceosome.

Although this model explains internal exon splicing well, exon definition requires a special mechanism for defining terminal 5’ and 3’ exons. Pre-mRNAs transcribed by RNA polymerase II are modified by 7-methylguanosine triphosphate (m7Gppp) cap structure (Shatkin, 1976). Addition of cap analogs or capped RNAs competes for splicing of pre-mRNA and the cap structure affects splicing of the 5’ proximal intron (Konarska et al., 1984; Edery and Sonenberg, 1985; Patzelt et al., 1987; Ohno et al., 1987; Inoue et al., 1990), which suggests that recognition of the first 5' splice site is promoted by the cap structure. In the same context as the 5’ end of pre-mRNA, it is proposed that interaction of the spliceosome with the polyadenylation signal defines the 3’ terminal intron (Robberson et al., 1990). The 3’ end formation is promoted by polyadenylation signals when placed within an exon but not within an intron (Adami and Nevins, 1988; Brady and Wold, 1988; Levitt et al., 1989). In addition, efficient polyadenylation occurs by the presence of an upstream 3’ splice site and the polyadenylation signal affects 3’ terminal intron splicing. These effects can be seen *in vivo* (Niwa et al., 1990; Niwa and Berget, 1991; Niwa et al., 1992; Nesic and Maquat, 1994). Interaction of U1 snRNA with pre-mRNA requires both an upstream 3’ splice site and a downstream polyadenylation site, which is shown by UV crosslinking (Wassarman and Steitz, 1993).

The alternative intron definition model is originated from the abundance of short introns (typically 100-200 bp) in invertebrates, such as *Drosophila* and *C. elegans*. Introns in some pre-mRNA are relatively short compared to exons. Therefore, the splicing machinery may initiate splicing reactions by pairing splice sites across the introns rather than exons (Talerico and Berget, 1994). As expected in this model, a mutation on the 5' splice site leads to intron retention, and expansion of short introns represses splicing (Talerico and Berget, 1994). Short introns may contain all the information required for splicing since inserted short introns into intronless transcripts are capable of splicing in yeast (Gatermann et al., 1989). From the results of a computational approach (Lim and Burge, 2001), splice signals of every intron in yeast accurately identify the introns although they do not determine the location of 3' cleavage. In *Drosophila* and *C. elegans,* splicing signals specify exact junctions of most short introns but in humans and plants, additional enhancer motifs are required for the specification of short introns.

1.4. RNA Splicing Regulation: Alternative mRNA Splicing

The most interesting outcome of the human genome project is that the number of genes is much smaller than the huge set of diverse proteins. The answer to this discrepancy emphasizes the importance of generating multiple mRNAs from a single gene. More than half of the genes (59%) contain alternatively spliced exons based on the alignment of expressed sequence tags (ESTs) and cDNAs to genomic sequences (Lander et al., 2001). More recently, at least 74% of human multi-exon genes have been detected to contain alternative exons using exon-junction microarrays (Johnson et al.,

2003). Before the recognition of the prevalence of alternative RNA splicing, it had been thought to increase proteomic diversity primarily in the nervous system (reviewed in Grabowski, 1998; Black, 2000; Grabowski and Black, 2001; Graveley, 2001; Stamm, 2002). The consequence of alternative RNA splicing has been noted by two landmark discoveries (Berget et al., 1977; Lander et al., 2001), interruptions of genes by introns and the lower than expected number of human genes. This signifies not only the prevalence of alternative RNA splicing in all tissues rather than just the nervous system as previously thought, but also the crucial contribution to the regulation of gene expression in various circumstances, thereby further determining tissue-specificity. Therefore, experimental tools are necessary to investigate alternative splicing patterns on a global scale in different tissues, at different developmental stages, in response to cellular signals, and in the progression of human diseases. Affymetrix microarray technologies have been used to analyze alternative splicing on a genome scale (Johnson et al., 2003). Using 20 probe pairs (25 nt) representing different exons of a gene, this method can indicate potential alternative exons by detecting reduced signals compared to other probes. However, there is a limitation to detect regulated exons using the consensus sequence of the splice junctions because the consensus sequences at the junction are short, degenerate and mainly positioned in the intron as discussed above, and because regulatory sequence elements are in the upstream and/or downstream introns as well as in the exon adjacent to the junctions (Lim and Burge, 2001; Fairbrother et al., 2002, reviewed in Blencowe, 2000; Cartegni et al., 2002; Ladd and Cooper, 2002).

Alternative splicing is not only involved in regulation of gene expression but also its malfunction is related to human diseases. A large-scale analysis proves that protein isoforms arising from alternative splicing tend to alter domain architecture and the functional sites of the protein by inserting, deleting or substituting functional residues (Kriventseva et al., 2003). This suggests that alternative splicing modulates the function of protein domains, and therefore possibly serves as a dominant-negative regulation mechanism, such as apoptotic cell death (Boise et al., 1993; Kriventseva et al., 2003). Alternative splicing can also introduce premature stop codons, regulating the level of gene expression (Hillman et al., 2004; Lamba et al., 2003; Lewis et al., 2003; Green et al., 2003; Wollerton et al., 2004). It is regulated in a cell type- or development-specific manner, which mediates fine control of gene expression and quantitative regulation of protein functions in the nervous system (Wang and Grabowski, 1996; Grabowski, 1998; Schmucker et al., 2000; Zhang et al., 2002). Splicing errors induced by genetic mutations are responsible for protein defects and are found in human diseases. Such mutations, which account for up to 15% of inherited point mutations, result in mRNA splicing defects (Krawczak et al., 1992; Cooper and Mattox, 1997). Therefore, understanding regulation of alternative splicing would be required for gene therapy to treat human diseases.

1.4.1. RNA regulatory elements

Most alternatively spliced exons contain weak splicing signals and are regulated by additional regulatory *cis*- and *trans*-acting elements. RNA control elements, which are distinct from splice sites, include the positive-acting exonic and intronic splicing

enhancers (ESEs or ISEs), and the negative-acting splicing silencers (ESSs and ISSs) (Table 1, reviewed in Grabowski, 1998; Blencowe, 2000; Cartegni et al., 2002; Ladd and Cooper, 2002; Smith and Valcarcel, 2000). In addition to splicing signals, regulatory sequences or the structure in the vicinity of the splice sites has been identified to play a pivotal role in splicing (reviewed in Black, 1995).

ESS is important for splicing control. Frontotemporal dementia with Parkinsonism linked to chromosome 17 (FTDP-17) is caused by mutations in the *tau* gene, which encodes a microtubule-associated protein Tau. FTDP-17 is caused by intronic mutations immediately adjacent to the 5' splice site of Tau exon 10 that increase exon inclusion (D'Souza et al., 1999). It is also caused from exonic missense mutations that affect microtubule (MT) assembly or a silent mutation. (Hutton et al., 1998; Hong et al., 1998; D'Souza et al., 1999). The regions that are involved in enhancing and silencing of immunoglobulin M (IgM) exons M1 and M2 splicing in mouse, but the exact sequences are not defined although the inhibitor is reported to be evolutionarily conserved (Kan and Green, 1999). The UUAG sequence of Tau exon 10, which is changed to U<u>G</u>AG (mutated sequence is underlined) by a silent mutation, is identical to a silencer motif found in HIV *tat* exon 3 (ESS3). However, Tau exon 10 is not repressed by hnRNP A1 in COS cells by co-transfection (D'Souza et al., 1999; Si et al., 1998; Gao et al., 2000). One good example for a defined ESS is a UAGG sequence in the K-SAM exon of fibroblast growth factor receptor-2 gene (FGFR-2) (Del Gatto et al., 1996). The FGFR-2 gene contains two alternative exons, K-SAM and BEK, which are spliced in a cell type-specific manner (Miki et al., 1992). UV crosslinking identifies hnRNP A1 as a binding protein of the UAGG sequence.

Many identified ESSs have a bipartite architecture, possibly representing a general feature in the regulation of alternative splicing (Amendt et al., 1995; Staffa and Cochrane, 1995; Zheng et al., 1996; Konig et al., 1998). It has been thought that most ESSs are located downstream of ESEs, which are different from those identified in the rat beta-tropomyosin (β-TM) exon 7 and in the human cell surface molecule CD44 exon 5 (Konig et al., 1998; Chen et al., 1999). However, the importance of the order of these ESEs and ESSs has not been proven. In the rat β-TM gene, exons 6 and 7 are cassette exons in a mutually exclusive manner. Exon 7 is silenced by an exonic sequence at the 5' end of the exon and an upstream intron sequence (Guo and Helfman, 1993). The sequence UGUGGG at the 5' end of exon 7 is part of an ESS, and the binding of a trans-acting factor is responsible for the activity of the ESS (Chen et al., 1999). Adjacent to the ESS, a purine-rich sequence GAGGAGGAG lies immediately downstream, that is identical to previously characterized ESEs (Watakabe et al., 1993; Xu et al., 1993).

Most naturally occurring ESEs are purine-rich sequences, such as (GAR)n, and play positive roles in both constitutive and regulated splicing associated with specific SR-family proteins, thus promoting the utilization of adjacent splice sites. Since ESEs are known to interact with SR proteins, consensus sequences of the motifs are well defined by the SELEX method (Liu et al., 1998; Schaal and Maniatis, 1999; Liu et al., 2000). ESE motifs are short (6-8 nt), degenerate and partially overlap each other (Blencowe, 2000; Cartegni and Krainer, 2003). Consensus winner sequences of human ASF/SF2, SC35, SRp40 and SRp55 have been obtained and web-based resources are

Table 1. Splicing regulatory RNA cis-acting elements and trans-acting factors.

Type	Regulated exon	Sequence motif	Splicing factor	References
ESS	Tau exon 10	UUAG	hnRNP A1	1
	FGFR-2 K-SAM exon	UAGG	hnRNP A1	3
	β-TM exon 7	UGUGGG	hnRNP H	4
ESE[a]	β-TM exon 7	GAGAGAG	Tra	4
	Doublesex *fruitless*	(GAR)n	Tra	15
ISS	HIV *tat* exon 3	UAG	hnRNP A1	2
	c-src N1 exon *GRIN1* NI exon γ2 24-nt exon	UCUU CUCU	PTB	5 6 6
	Adenovirus L1 unit IIIa	upstream intron	ASF/SF2	8
	GRIN1 NI exon	upstream intron	NAPOR	6
	α-TM exon 3	upstream intron		7
ISE	GH exon 3 α-globin β-TM	G(A/U)GGG		9 14 16
	c-src N1 exon	DCS C-rich	KSRP	10
	c-src N1 exon	DCS	hnRNP F	11
	c-src N1 exon	DCS	hnRNP H	12
	GRIN1 CI exon	downstream intron	NAPOR	17
	cTNT exon 5	UGUG, UUGUU	ETR-3	13

a. *ESEfinder* database for ESE motifs http://exon.cshl.edu/ESE/, Cartegni et al., 2003; 1. D'Souza and Schellenberg, 2002; 2. Tange et al., 2001; 3. Del Gatto et al., 1996; 4. Chen et al., 1999; 5. Chan and Black, 1997; 6. Ashiya and Grabowski, 1997; 7. Guo and Helfman, 1993; 8. Kanopka et al., 1996; 9. Cogan et al., 1997; 10. Markovtsov et al., 2000; 11. Min et al., 1995; 12. Chou et al., 1999; 13. Philips et al., 1998; 14. McCullough and Berget, 1997; 15. Lynch and Maniatis, 1996; 16. Sirand-Pugnet et al., 1995; 17. Zhang et al., 2002.

available (http://exon.cshl.edu/ESE/, Cartegni et al., 2003). Separately, using a computational method, RESCUE-ESE, 10 representative hexamer motifs which are categorized as enhancing 5' and 3' splice site recognition, respectively, are found in human (Fairbrother et al., 2002). This study hypothesizes that weak splice sites require ESEs to strengthen the splice signal, resulting in a higher frequency of ESEs in weak exons (Fairbrother et al., 2002). Although the precise mechanism of ESEs involved in spliceosome assembly is controversial, the motifs occur frequently in the exon region. In addition, SR proteins that bind to the ESE motifs are ubiquitously expressed, increasing the possibility of important roles in splicing (Liu et al., 1998; Liu et al., 2000; Schaal and Maniatis, 1999; Cartegni et al., 2003). A well-characterized ESE function is described in studies on the cascade of alternative splicing that is involved in determining the sex of *Drosophila*, which regulates inclusion of exon 4 in the *doublesex* (*dsx*) gene (Figure 2C, reviewed in Graveley, 2002; Lopez, 1998). The *sex lethal* (Sxl) exon 3 is skipped in females, producing SXL protein, which maintains protein expression by autoregulation of *Sxl* alternative splicing pattern. SXL protein induces TRA protein expression by repressing the male-specific splice site of the *transformer* (*tra*) and *male-specific lethal 2* (*msl-2*) genes. Female-specific exon inclusion of *doublesex* and *fruitless* requires six 13-nucleotide repeats as well as a purine-rich element (PRE) located over 300 nucleotides downstream of the 3' splice site (Lynch and Maniatis, 1996). Binding of this ESE by TRA protein positively regulates splicing of *doublesex* and *fruitless* pre-mRNAs.

Intronic RNA regulatory elements, intronic splicing enhancers (ISE) and intronic splicing silencers (ISS), are also involved in the control of alternative splicing. A representative ISS is a polypyrimidine tract at the 3' splice site which interacts with

polypyrimidine tract binding protein (PTB), thereby repressing splicing. The consensus sequences are UCUU or CUCU repeats in either the upstream or downstream intron (Mullen et al., 1991; Patton et al., 1991; Mulligan et al., 1992; Chan and Black, 1997; Ashiya and Grabowski, 1997). These motifs are found in various systems, for example, intron 6 of rat β-tropomyosin (Mulligan et al., 1992), intron 2 of rat α-tropomyosin (Mullen et al., 1991), the introns upstream and downstream N1 exon of c-src (Chan and Black, 1995; Chan and Black, 1997) and the upstream intron of the GABA (A) receptor γ2 subunit 24 nucleotide (24-nt) exon (Ashiya and Grabowski, 1997).

CUG repeats occur in intronic muscle-specific splicing enhancers (MSEs) of cardiac troponin T (cTNT) exon 5 (Ryan and Cooper, 1996; Philips et al., 1998). Enhancement of proximal upstream exon splicing by this ISE is mediated by CUG binding protein (CUGBP). CUG expansion in the 3' untranslated region (UTR) causes myotonic dystrophy (DM), dominant for neuromuscular disorder, by inappropriate binding of the proteins to the CUG repeats. In the small α-globin second intron, a G triplet is a minimal intronic enhancer that additively improves splicing (McCullough and Berget, 1997). G triplets are suggested as a common component of the intron downstream from human 5' splice sites and assist in the intron definition model. In intervening sequence 3 (IVS3) of the human growth hormone (GH) gene, (A/U)GGG repeats located 28 bp downstream from the 5' splice site repress mRNA splicing, which resembles A1 winner sequences (Cogan et al., 1997). Another report shows that an intronic (A/U)GGG repeat is important for mRNA splicing of intron 7 in chicken β-TM pre-mRNA (Sirand-Pugnet et al., 1995). This effect is additive and this repeat is necessary for spliceosome assembly. Analysis of a 1000 nucleotide window of around

664 primate 5' exon/intron junctions revealed high frequency of G-rich runs (G-rich quartets) downstream of the 5' splice sites and some C-rich quartets upstream of the 5' splice site (Nussinov, 1988; Nussinov, 1989). Although the data set is relatively small, it proposes that the pre-mRNAs contain unique hairpin loop structures in which conserved sequences occupy specific sites, thereby exposing either of the splice sites. Thus, the G-rich quartet plays an enhancing role in pre-mRNA splicing in this model and in some introns as mentioned above. However, since mutations destroying the secondary structure do not affect splicing efficiency, this model requires further testing.

In addition to sequence elements specific for RNA binding proteins, the secondary structure formed by the sequences in the vicinity of a 5' splice site has been identified to repress alternative splicing. Mutations in the 5' splice site of Tau exon 10 increase splicing of this exon by acting through disruption of the stem-loop structure, providing evidence that the stability of the stem-loop structure underlies the alternative splicing of exon 10 (Grover et al., 1999). The alternative 3' splice site A3, required for *tat* mRNA, is in the terminal loop of a stem-loop structure (SLS2), which is highly conserved in HIV-1 (Jacquenet et al., 2001). A mutation of the A3 downstream sequence, which reinforces SLS2, decreases A3 utilization, suggesting that the secondary structure as well as the sequence may cooperate in limiting utilization of the site A3.

1.4.2. RNA-binding *trans*-acting factors

The *cis*-acting splicing enhancers and silencers of pre-mRNAs interact with *trans*-acting factors controlling alternative splicing, which are spliceosomal components involved in specifying splice signals (Table 1). Most ESEs are identified as binding

motifs for SR family proteins, which are non-snRNP splicing factors and contain N-terminal RNA recognition motifs (RRM) and C-terminal arginine/serine-rich (RS) domains (reviewed in Graveley, 2000; Black, 2003; Manley and Tacke, 1996; Fu, 1995). SR-related proteins that also contain RS domains are functionally and structurally distinct from the SR family. These proteins are mostly positive splicing regulators, general splicing factors (U2AF, U1 70K and SRm160/300), or rarely specific inhibitors (Reviewed in Graveley, 2000; Black, 2003). The RS domains of these proteins interact with the RS domains in U1 snRNP when U1 snRNP associates with the 5' splice site (reviewed in Chabot, 1996; Lopez, 1998; Tacke and Manley, 1999; Graveley, 2000). SR and SR-related proteins function in the recruitment of the tri-snRNP particle and after the first catalytic step (Chew et al., 1999; Roscigno and Garcia-Blanco, 1995). ASF/SF2-specific ESEs activate the second catalytic step of a trans-splicing reaction in a bimolecular exon ligation assay (Chew et al., 1999). As an exception, ASF/SF2 inhibits adenovirus L1 unit IIIa pre-mRNA splicing by binding to an intronic repressor element and by preventing recruitment of U2 snRNP to the spliceosome (Kanopka et al., 1996).

The polypyrimidine tract binding protein (PTB, or hnRNP I) is a negative splicing regulator of multiple targets in the nervous system and muscles (reviewed in Valcarcel and Gebauer, 1997), and represses neuron-specific splicing in nonneural cells. PTB has been proposed to be an important splicing factor for recognition of the 3' splice site, which is shown by UV crosslinking (Garcia-Blanco et al., 1989). The PTB functions as a repressor by binding ISSs in either the upstream or downstream intron (Mullen et al., 1991; Patton et al., 1991; Mulligan et al., 1992; Chan and Black, 1997). PTB contains

four non-canonical RRMs that lack conserved aromatic residues, and RRM4 is required for coupling its splicing repression and RNA-binding activities (Dreyfuss et al., 1993; Liu et al., 2002). The $GABA_A$ receptor γ2 subunit 24-nt exon, the clathrin light chain B exon EN, and the N-methyl-D-aspartate receptor NR1 (NMDA R1) subunit NI cassette exon (describes as exon 5) pre-mRNAs are regulated by PTB during rat cerebellum development (Zhang et al., 1999). In nonneural cells, the c-*src* N1 exon, which is neuron-specific, is repressed by binding of PTB to multiple elements (Chan and Black, 1997; Chou et al., 2000).

KH-type splicing regulatory protein (KSRP), which is similar to hnRNP K, is also a neural-enriched RNA-binding protein identified from a neuronal splicing extract. It is an intronic enhancer of the c-*src* N1 exon contrary to PTB, thereby promoting the assembly of the intronic enhancer complex (Min et al., 1995). Another splicing activator with three hnRNP K homology (KH) domains and specifically expressed in brain is Nova-1 (Buckanovich and Darnell, 1997). In Nova-1 knockout mice, the ratio of the exon-included to -skipped mRNA products, glycine receptor α2 exon 3A and $GABA_A$ receptor exon γ2L, is reduced in the corresponding midbrain and hindbrain regions where Nova-1 protein has been genetically depleted (Jensen et al., 2000). In *rous sarcoma* virus, the 55-kDa protein (p55) is crosslinked to specifically bind NRS5', a negative regulator of splicing (NRS), and is identified as hnRNP H (Caputi and Zahler, 2001; Fogel and McNally, 2000). In the c-src N1 exon, hnRNP H is a component of a neuron-specific enhancer complex containing hnRNP F, KSRP, and an unidentified 58 kD protein (p58) downstream of the 5' splice site. ESS activity of the β-TM exon 7 is correlated with specific binding by hnRNP H. When its binding is inhibited through antibodies against

hnRNP H or partial depletion of hnRNP H from nuclear extract, exon 7 splicing *in vitro* is activated. This effect can be reversed by addition of purified recombinant hnRNP H. These results indicate that hnRNP H participates in exclusion of exon 7 in nonmuscle cells (Chen et al., 1999).

Mouse neuroblastoma apoptosis-related RNA-binding protein (NAPOR; mNAPOR-1) has been reported to have dual functions, a silencing function in the regulation of the NI cassette and an enhancing function in CI cassette splicing events of the NMDAR1 pre-mRNA (Zhang et al., 2002). NAPOR (also called CELF (CUGBP and ETR3-like factors) or BRUNOL (Bruno-like proteins), Good, 1997; Ladd et al., 2001) is identified by its increased expression in a neuroblastoma cell line during apoptosis and is closely related to human CUG-binding protein 2 (CUG-BP2) (Choi et al., 1999). NAPOR contains three RRMs and a divergent domain and was initially identified as an embryonic lethal abnormal vision (ELAV) type RNA binding protein 3 (ETR-3) in the fetal heart (Hwang et al., 1994). CUG-BP1 promotes splicing of exon 5 of the cardiac troponin T (cTNT) pre-mRNA by interacting directly with an intronic enhancer element rich in CUG repeats (Philips et al., 1998). In Drosophila, the ELAV has a neuron-specific expression pattern and is believed to play a crucial role in the regulation of alternative splicing of neuroglian (Koushika et al., 1996). In mammals, the ELAV related Hu-type RNA-binding proteins, which are expressed in brain and muscle, exhibit brain region- and cell-specific expression patterns (Keene, 2001; Okano and Darnell, 1997). RRM2 with an additional 66 amino acids of ETR-3 can function as a full length protein in activating cTNT exon 5 by binding the enhancer sequences (Singh et al., 2004). RRM 3

contains a nuclear localization signal and is important for splicing activity with the divergent domain (Ladd and Cooper, 2004).

Tissue-specific splicing decisions are thought to occur by synergistic or antagonistic activities of complexes that assemble on multiple RNA sequence elements, which is called combinatorial control (reviewed in Lopez, 1998; Smith and Valcarcel, 2000). A neuron-specific N1 exon of *c-src* contains six different RNA regulatory elements that negatively or positively affect exon inclusion (Chan and Black, 1997; Modafferi and Black, 1999). Intronic PTB, hnRNP F and hnRNP H binding sites, and exonic enhancers contain redundant signals and overlap with RNA elements of opposite effects in different cell lines, which increases exon inclusion in neuronal cells by the balance of their mutual activities. Two models were proposed to illustrate neuron-specific splicing of a cassette exon, a positive regulation model and a negative regulation model (Grabowski, 1998). Enhancing activity promotes the positive regulation of neuron-specific exons in neuronal tissues whereas this activity is minimal in non-neuronal tissues. Although the expression is not exclusively neuron-specific, KSRP is thought to shift the equilibrium of alternative splicing from exon skipping to exon inclusion by binding to splicing enhancer elements, which is a good example for the positive control model. On the other hand, the negative regulation model proposes that silencing activity of neuron-specific exons occurs in non-neurons while this activity is absent in neurons. The negative control model is attractive because no specific factors are required for neuron-specific splicing and also because a difference in the silencing activity could account for the tissue specific splicing. PTB is a representative factor for this negative regulation model.

1.5. Background for the NMDA R1 receptor: biology and splicing pattern

In the nervous system, many alternatively spliced mRNAs are translated into proteins playing roles in learning and memory, neurotransmission, ion channel function, and synaptic specificity (Grabowski and Black, 2001). N-methyl-D-aspartate receptors (NMDARs) are excitatory neurotransmitter receptors, belonging to the subclass of ionotropic glutamate receptors. NMDARs are present at many excitatory synapses and are widely distributed in the mammalian brain (Laurie et al., 1995; Winkler et al., 1999, reviewed in Cull-Candy et al., 2001). Molecular cloning has revealed that NMDARs consist of NR1, NR2, and NR3 subunits (O'Brien et al., 1998; Moriyoshi et al., 1991; Sugihara et al., 1992), and their surface expression plays a crucial role in neuronal development of the brain, synaptogenesis, synaptic plasticity, and excitotoxicity (Zukin and Bennett, 1995; Ehlers, 1998 #330]). Therefore, inappropriate surface expression, signaling or synaptic targeting of NMDARs could trigger neuronal disorders such as stroke, addiction and neurodegenerative diseases (Cull-Candy et al., 2001).

Alternative splicing of a NR1 gene generates eight isoforms via the insertion or deletion of two exon cassettes in the N-terminal (NI) and C-terminal (CI), as well as alternate 3' splice site selection (CII) domains (Hollmann et al., 1993; Zukin and Bennett, 1995). Splice variants show differences in their spatial and temporal expression patterns, pharmacological properties, ability to be phosphorylated, subcellular localization in transfected fibroblasts and responsiveness to calmodulin-mediated inactivation (Hollmann et al., 1993; Ehlers et al., 1995; Ehlers et al., 1998; Zukin and Bennett, 1995). Splicing of the NI (described as exon 5) and CI (described as

exon21) cassette exons are highly tissue-specific. Each shows a substantial reversal of the splicing pattern in rat forebrain and hindbrain tissue (Zhang et al., 2002). PTB is involved in the repression of r2, CLCB, and NMDAR1 NI cassette exon during neuronal development (Wang and Grabowski, 1996). In contrast, NMDAR1 CI cassette exon exhibits the different pattern of regulation where sequences of the upstream intron lack PTB binding motifs but are G-rich (Ashiya and Grabowski, 1997). NAPOR, which is highly expressed in rat forebrain, shows a positive correlation with CI cassette exon inclusion (Zhang et al., 2002). In transient coexpression *in vivo*, NAPOR exhibits a positive effect on CI cassette exon splicing but a negative effect on the NI cassette exon, suggesting dual function as a splicing regulator. Therefore, understanding the regulatory mechanism involved in CI cassette exon splicing would reveal a novel developmental control of gene expression through alternative splicing in the nervous system.

Given the central dogma of gene expression, the functionality of CI domain originates from the DNA primary sequence, thus control of splice variants should be related to protein functions. NMDARs are selectively localized at the postsynaptic membrane in the brain but are also found at extrasynaptic sites. Previous studies have suggested that the splice variant form may determine the localization of the NMDA receptor NR1 subunit (Al-Hallaq et al., 2001). The CI cassette has been identified to contribute to trafficking of NR1 subunit in neurons, which contains an ER retention signal (RRR) (Okabe et al., 1999; Standley et al., 2000; Scott et al., 2001). The CI cassette associates with a neurofilament (NF) subunit NF-L, suggesting a possible mechanism for anchoring to the cytoskeleton or localizing NMDA receptors in the

neuronal plasma membrane (Ehlers et al., 1996; Ehlers et al., 1998). The interaction of NF-L with NMDA receptors could provide such a linkage by anchoring NMDA receptors in the plasma membrane to cytoplasmic neurofilaments. The CI domain encoded by the CI cassette exon serves as a substrate for phosphorylation by protein kinase C (PKC) at serine 890 (S890) and 896 (S896), and protein kinase A (PKA) at S897 (Figure 6A, Ehlers et al., 1995; Ehlers et al., 1996; Tingley et al., 1997), and contains a high affinity calmodulin (CaM) binding site (Ehlers et al., 1995; Ehlers et al., 1996). PKC phosphorylation in this exon disrupts receptor clustering which can modulate CaM-dependent inactivation of NMDA receptors (Ehlers et al., 1995). Coordinated phosphorylation by PKA and PKC promotes surface delivery of NMDA receptors by masking the ER retention signal (Scott et al., 2003). The NMDA NR1 gene, which is indispensable for the formation of functional NMDA receptors, is widely expressed in the central nervous system and the expression is varied during development. Altogether, the absence of the CI domain of NMDAR1 in the cytoplasmic face of postsynaptic neurons would change receptor functions. Therefore, functional analysis of the CI cassette exon indicates the importance to study its splicing regulation mechanism in order to better understand how the function of NMDA receptors is controlled in neurons. It has been shown that the CI cassette exon is both positively and negatively regulated in different brain regions. NAPOR also exemplifies the positive regulation model of the CI cassette exon. Because splicing control regulated by PTB has been a major mechanism to explain neuron-specific splicing, studying the non-canonical splicing mechanism for the CI cassette exon regulated by NAPOR promises to reveal a novel mechanism for the positive regulation model.

1.6. Project Goals

The major goal of this study is to understand a splicing silencing mechanism controlling the NMDAR1 CI cassette exon. Specifically, this work investigates RNA sequence elements involved in blocking recognition of splice sites and their related factors, and further expands the silencing mechanism to a genome wide scale. This research is designed to answer the following questions: What is the role of the 5' splice site and the proximal downstream intron in alternative splicing of the CI cassette exon? What RNA sequence elements within the CI cassette exon are responsible for the splicing silencing mechanism? What are the roles of RNA secondary structure and RNA binding proteins? How prevalent is this splicing silencing mechanism in the human and mouse genomes? Is this mechanism commonly associated with neuron-specific exons? To answer these questions, an *in vivo* splicing assay system established, which utilizes a splicing reporter constructed from the *GRIN1* gene. Mutational analysis is employed to identify *cis*-acting elements and to test a hypothesized secondary structure model that might be involved in the silencing mechanism. Interactions of *trans*-acting factors with the identified silencers are examined by UV crosslinking and immunoprecipitation with site-specific labeled RNA substrates. Using bioinformatics approaches, cassette exons that could be regulated by the same silencing mechanism are searched and tested by biochemical analysis.

Chapter 2. Silencing mechanism of the CI cassette exon alternative splicing

2.1. Introduction

Alternative splicing is a mechanism of generating multiple mRNAs, and consequently, proteins, from a single gene. Despite its importance, little is understood about the molecular mechanisms controlling alternative splicing. Some research groups have focused on neuronal genes involved in information processing in mammals (reviewed in Black, 2000; Black, 2003; Grabowski and Black, 2001). These studies have been focused on tissue-specific splicing regulators that are differentially expressed in various tissues, such as proteins involved in the silencing or enhancement of spliing. Studies of alternative splicing in the nervous system have identified negative regulators that require specific binding sequences. In particular, PTB has been involved in mechanisms of negative control of brain and muscle specific splicing events (Ashiya and Grabowski, 1997; Chan and Black, 1997; Charlet et al., 2002; Grabowski, 1998; Polydorides et al., 2000). PTB is involved in the regulation of NMDAR1 NI cassette exon (described as exon 5) but not the regulation of NMDAR1 CI cassette exon (described as exon 21) during development of the rat cerebellum. Consistently, PTB is not associated with the upstream intron region containing the 3' splice site of the CI cassette exon. The NI and CI cassette exons are reciprocally regulated by NAPOR in rat forebrain, which acts as a negative regulator for the NI cassette but a positive regulator for the CI cassette in an *in vivo* transient splicing assay. Mouse NAPOR has been found to interact with PTB in a yeast two hybrid screen (Zhang et al., 2002).

This study seeks to investigate the molecular and biochemical basis for the mechanism of splicing silencing of the CI cassette exon. The silencing mechanism is an important aspect of regulation of the CI cassette exon by NAPOR. Moreover, the complexity of this mechanism could be explained in part by the differential expression of NAPOR in the brain regions. The CI cassette exon and its flanking intronic sequences are investigated to search for *cis*-acting splicing regulatory elements that weaken splice signals. From the exon definition model, which suggests an exon as a bridge for spliceosome assembly, the splice site and exonic region of the CI cassette would be considered likely targets for splicing silencing. Furthermore, the functional importance of the CI domain of NMDA receptor protein encoded from the CI cassette exon prompts us to examine the exonic sequences involved in splicing silencing.

The novel findings of this work include the identification of a pattern of guanosine cluster (G cluster) and UAGG motifs required for silencing of the CI cassette exon. This study also demonstrates that the G cluster is required for the silencing function of hnRNP A1. In addition, potential new candidate exons are found in human and mouse using a genome wide search for a pattern of the G cluster and UAGG motifs, and alternative exons are validated by biochemical approaches.

2.2. Materials and methods

2.2.1. Mutagenesis of splicing reporter plasmids and construction of protein expression plasmids

All the 5' splice site mutants generated for this study were derivatives of the wild-type exon 21 splicing reporter E21wt (Zhang et al., 2002) in the pBPSVPA$^+$ vector (Nasim et al., 1990). Site-directed mutations were done using QuikChange® Site-Directed Mutagenesis Kit (Stratagene). The mutated sequences were designed in the primer sets used for polymerase chain reaction (PCR) reactions, which are listed in Appendix A. The region of mutated sites on plasmids was sequenced to confirm the expected base changes. The clone pA$^+$E21wt was converted to pA$^+$E21wto by point mutation on an engineered XhoI site in the exon region that was used for the convenience of cloning. The plasmid pA$^+$E21wto then was used as a template for site-directed mutations generating ESE, double and/or triple mutants. Chimeric construct rGγE21 was created by inserting the NotI and BamHI fragment of PCR product containing C1 cassette and the flanking introns into the NotI and BamHI sites of rGγ25 (Zhang et al., 1996). Primers used for amplifying the insert are listed in Appendix A (E21-U164 and E21-D103). The construct contains 165 nucleotides of upstream and 103 nucleotides of downstream intron regions of C1 cassette. The EcoRI fragment of rGγE21 was transferred into the EcoRI site of the pBPSVPA$^+$ vector for transient expression assays. The EcoRI site of rGγE21 was eliminated by filling the ends with Klenow and was ligated into the blunt ends of the vector pBS-. The blunt ends of vector pBS- were generated by end filling with Klenow followed by BamHI digestion. Plasmid pBS-rGγE21 was used as a parent plasmid for intron substitution constructs pBS-rGγE21-up and pBS-rGγE21-dn. The construct pBS-rGγE21-up was made by replacing the XbaI and NotI fragment of pBS-rGγE21 with the XbaI and NotI fragment of the PCR product from E21wt with the primers, 3021-XbaI and E21-U164A-NotI. The γ2 exon 1

and its downstream intron region was replaced with the upstream exon of E21wt and its downstream intron. Substituting the BamHI and EcoRI fragment of γ2 exon 3 and the upstream intron with the BamHI and EcoRI digested PCR product of the E21wt template with primers (E21-Dn-103-BamHI and 3255A-EcoRI) generated pBS-rGγE21-dn. The blunt ends of rGγE21-up or rGγE21-dn were generated by filling the EcoRI digested ends and were ligated at EcoRV ends of pBPSVPA^{+} vector. The pA^{+}rGγE21-dn-5m2 and pA^{+}rGγE21-dn-5m4 mutants were produced by site-directed mutagenesis. All constructs were verified by restriction mapping and DNA sequencing. Plasmids pFlag-F containing hnRNP F (Chou et al., 1999) and pFlag-DSEF-1 containing hnRNP H/H' (DSEF-1 in Bagga et al., 1998; Arhin et al., 2002) were obtained from the Milcarek lab and the hnRNP A1 clone, Myc-A1, from the Dreyfuss lab. The hnRNP F and hnRNP H/H' coding regions were individually subcloned into the BamHI site of pcDNA4/HisMax vectors (Invitrogen). The hnRNP A1 expression plasmid was also constructed between the EcoRI and XhoI sites of the pcDNA4/HisMax vector using the entire coding region.

2.2.2. Transient expression and coexpression assays

C2C12 cells were grown in Dulbecco's modified Eagle medium (DMEM, Gibco BRL) with 10% (v/v) fetal bovine serum. For each transfection, 1x10^{6} cells were seeded on 60-mm plates to achieve 60–80% confluency. For transient expression assays, 3 μg of wild type or mutant reporter plasmids were used for transfection. For transient coexpression assays, the pcDNA4-hnRNP F, pcDNA-hnRNP H/H' or pcDNA4-hnRNP A1 expression plasmid was cotransfected with the reporter in a ratio of 2/1, 4/1 or 6/1. Each cotransfection contained a total of 3.5 μg of plasmid DNA. Plasmids were mixed

with 300 μl of Opti-Mem followed by the addition of an equal volume of Opti-Mem mixed with 20 μl of Lipofectamine (Gibco BRL). The methods for transfection, total RNA purification and RT-PCR analysis were as described in a previous study (Zhang et al., 2002). PCR was carried out for 20 cycles in the transient expression assays and 24 cycles in the transient coexpression assays.

2.2.3. Transcription and site-specific labeling

Single stranded DNA oligonucleotides containing T7 promoter sequence were synthesized as templates for transcription. Sequence of complementary DNA oligonucleotides is indicated in Appendix A. Complementary DNA oligonucleotides were annealed by slowly cooling down to room temperature after heating. Labeled substrate RNAs, cs1, cs2, 3h1, 3h3 and A1 winner, were synthesized from 100 ng of the annealed DNA template in a 25 μl reaction of 0.4 mM each of ATP, UTP and CTP, 0.3 mM GTP, 2.5 μl of [α-^{32}P]GTP (3000 Ci/mmol, 10 μCi/μl), 0.5 mM diguanosine and T7 RNA polymerase (Stratagene). Following 2 hours incubation at 37ºC, the reactions were digested with Dnase I (Promega) for 10 minutes at room temperature. Labeled RNAs were collected by Sephadex G25 chromatography. Site-specific labeled substrate RNAs were prepared by transcribing and ligating two transcripts (Moore and Sharp, 1992; Moore and Query, 2000). Sequences of complementary DNA oligonucleotides used as templates for transcription are indicated in the order of sense followed by antisense strand (Appendix A). 5' half RNAs of wt3 and mt3 and 3' half RNA were transcribed respectively in a 125 μl reaction under the conditions of 500 ng annealed DNA template, 0.4 mM each of GTP, ATP, UTP and CTP, 2 mM guanosine and T7 RNA polymerase

(Stratagene). Reactions were incubated at 37ºC for 2 hours and digested with Dnase I (Promega) for 10 minutes at room temperature. Transcripts were resolved on 7M urea/10% polyacrylamide gel and purified by incubating the gel slice at 42ºC for at least 6 hours in an elution buffer containing 1 mM EDTA, 2% SDS and 300mN NaOAc. An incorporated guanosine at the 5' end of 3' half RNA was phosphorylated using 1:1 molar ratio of [γ-^{32}P] ATP (6000 Ci/mmol, 150 μCi/ul) to RNA. Unincorporated nucleotides were removed by Sephadex G25 chromatography. Labeled 3' half RNA (25 pmol) and 5' half RNA were hybridized to an antisense splint DNA (Appendix A) by heating a mixture of the three and then slowly cooling down to ambient temperature in T4 DNA ligase reactions. The ligation reactions (10 μl) were incubated with 1.5 weiss unit/μl of T4 DNA ligase (NEB) at 16 ºC for 4 hours. The ligation reactions were digested with DNase I and gel-purified. The concentrations and the integrity of RNAs were confirmed by electrophoresis of 7M urea/10% polyacrylamide gels.

2.2.4. Biotinylation of RNA substrates and biotin-streptavidin affinity selection

RNA substrates were biotinylated at random positions using biotin-16-UTP and T3 RNA polymerase under the *in vitro* transcription reaction conditions as above. After pre-incubation of HeLa nuclear extract under splicing reaction conditions for 5 min at 4°C, biotinylated RNA substrates were added and incubated for 5 min at 30°C to form splicing complexes in a 25 µl volume. Reactions that were then adjusted to 0.4 mg/ml of heparin were incubated for 3 minutes at 30°C and stopped by addition of 30 µl of 1X NET (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5mM EDTA) containing 0.01% NP-40 to

the reaction mixture. Streptavidin-agarose beads (50 µl) were equilibrated in a 1X NET buffer containing 0.01% NP-40 and preblocked by incubating twice in a blocking buffer (0.5mg/ml BSA, 0.05% glycogen, 0.05% tRNA and 0.01% NP-40) for 15 min. After washing out with 1X NET, the reaction solutions above were incubated with 25 µl pre-blocked beads for 2 hours at 4°C and washed five times with 1X NET. Biotin-affinity selected RNAs were eluted from the beads by digestion of biotin-streptavidin complexes with proteinase K for 20 minutes. The 3' end of RNAs was labeled using 5'-^{32}P-pCp (cytidine-3',5'-bis-phosphate) and RNA ligase. The RNA samples were resolved in a 10% polyacrylamide, 7M urea gel and visualized by autoradiography.

2.2.5. UV crosslinking and Immunoprecipitation analysis

UV crosslinking was done in 12.5 µl splicing reactions containing HeLa nuclear extract. In vitro splicing reactions contained 36 % (v/v) HeLa nuclear extract (4 mg/ml), 1 mM $MgCl_2$, 1.5 mM ATP, 5 mM creatine phosphate, 60 mM KCl, 20 mM Hepes, pH 7.4, and 150 kcpm or 100 kcpm of substrate RNA. After incubation at 30 ºC for 5 minutes, heparin was added to the reactions to a final concentration of 0.25 mg/ml, and incubated on ice for 5 minutes. Samples were then irradiated on ice with 1.2 J of UV for 10 minutes and digested with 1 mg/ml ribonuclease A for 20 minutes at 30 ºC (Ashiya and Grabowski, 1997). The samples in SDS sample buffer were boiled for 5 minutes and loaded onto a 12.5 % SDS-polyacrylamide gel. Gels were run at 30 mA for approximately 4 hours, fixed in 45 % methanol, 9 % acetic acid for 1-12 hours and then dried. Crosslinked proteins were detected by autoradiography. For immunoprecipitation analysis, protein A beads (Sigma) were equilibrated in 1 X Buffer A containing 10 mM

Tris-HCl, pH7.5, 100 mM NaCl, 2.5 mM $MgCl_2$ and 1 % Triton X-100. The 25 μl of equilibrated beads were incubated with antibodies, 5 μL each of R7263 (Veraldi et al., 2001), R7264 (Veraldi et al., 2001) and preimmune serum respectively, or 1 μl each of 9H10 (Burd and Dreyfuss, 1994b) and nonimmune Ig G in the same way. Antibodies used in this study, R7263 (anti-hnRNP F antibody) and R7264 (anti-hnRNP H/H' antibody), were gifts from the Milcarek lab and 9H10 (anti-hnRNP A1 antibody) from the Dreyfuss lab. After incubation on ice for 1 hour, the beads were washed three times with 1X Buffer A. Equal volumes of UV crosslinking reactions were incubated with the beads for 20 minutes on ice. The reactions were precipitated, washed with 1X Buffer A and resolved on a 12.5 % SDS-polyacrylamide gel.

2.2.6. Immunoblotting

After SDS-PAGE, gels were soaked in Transfer buffer (25 mM Tris, 192 mM glycine and 20% methanol) and the resolved protein samples were transferred to Immobilin P membrane at 100 mA for at least 4 hours at 4ºC. The blot was incubated with 5% blocking milk in PBST (1X PBS and 0.05% TweenX20) at room temperature. After washing the blot, the primary antibody, anti-Xpress (1:20,000 dilution) in PBST buffer, were incubated for 1hour with shaking and washed with PBST buffer. The secondary antibody, anti-mouse anti IgG (1:25,000) conjugated to alkaline phosphatase in PBST buffer, were incubated for 1 hour and washed. By adding Western Lighting reagents, chemiluminescence signal was developed and visualized by autoradiography.

2.2.7. RT-PCR analysis for the validation of cassette exons

Poly A^+ RNAs of mouse heart and brain were purchased from Clontech and total RNA of C2C12 cells was isolated with RNAzol (Biotecx). A series of 0.1 μg, 0.2 μg and 0.4 μg of $poly^+$ RNA or 1 μg, 2 μg and 4 μg of total RNA were hybridized with 400 ng, 800 ng and 1,600 ng of hexanucleotide random primers. Primed RNA samples were then used for a 20 μl reverse transcription (RT) reaction containing 0.5 mM dNTPs, 0.01MDTT, and 80U M-MLV reverse transcriptase (Stratagene), respectively. The reactions were incubated at 37 ºC for 1 hour followed by 72 ºC for 10 minutes to inactivate the enzyme. PCR reactions contain 1/20th of the reverse transcription reaction, 0.2 μM primers, 0.2 mM dNTPs, 2 mM MgCl2, 2U Taq polymerase and 1 μCi [α-^{32}P]dCTP in 10 μl total volume. Cycling parameters included: denaturation at 94 ºC, 30 seconds, annealing at 50 or 55 ºC, 30 seconds, and elongation at 72 ºC, 30 seconds for 22 ~ 24 cycles, followed by a final step at 72 ºC, 10 min. A total of 1/4th of each reaction was resolved on 5 M urea/6 % polyacrylamide gels. The bands were quantified directly from the gel using a Fuji phosphorimager system. PCR primers were designed using the PRIMER3 program (*http://frodo.wi.mit.edu/*) and are listed in Appendix A.

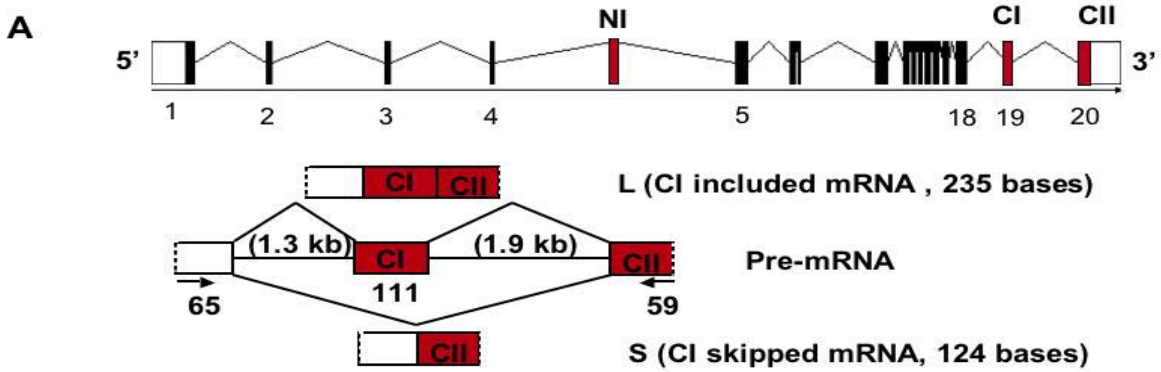
2.3. Results

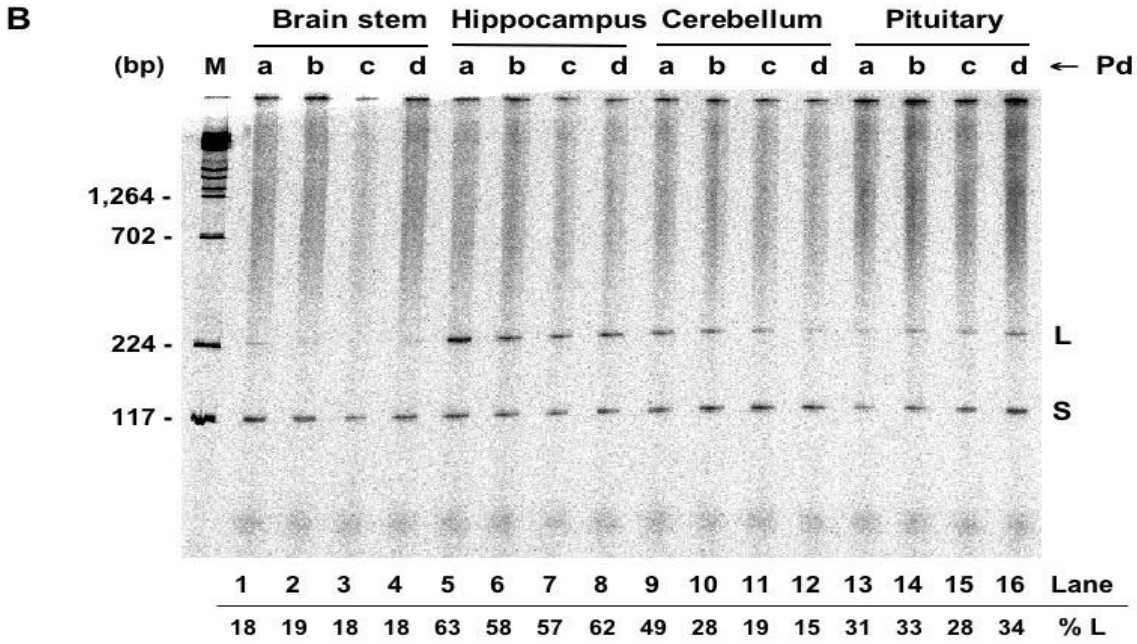
2.3.1. Splicing of the CI cassette exon of NMDARI (*GRIN1*) pre-mRNA is tissue- and developmental stage-specific in the rat brain.

Alternative splicing of the NMDA R1 subunit NR1 gene (*GRIN1*) is regionally regulated in the rodent brain, which may contribute to functional diversity of NMDARs in distinct brain areas (Laurie and Seeburg, 1994; Standaert et al., 1994; Hollmann et al., 1993). Given the limitations of *in situ* hybridization techniques, accurate quantification of distinct *GRIN1* splice variants was determined by a RT-PCR assay (Winkler et al., 1999; Zhang et al., 2002). Results from this method agree with previous *in situ* hybridization studies reporting ubiquitous distribution of *GRIN1* mRNA in the rat brain. However, the ratio of CI cassette exon-included to -skipped mRNA of the *GRIN1* splice variants could not be precisely quantified, because the PCR reaction also contains an additional standard in a single tube. In another study, developmental splicing regulation of the CI cassette exon was shown in a single Purkinje neuron cell by single-cell RT-PCR (Wang and Grabowski, 1996). The splicing regulation in the Purkinje neuron cell represented the splicing patterns of the CI cassette exon as well as the GABA receptor γ2 subunit neuron-specific exon in the cerebellum.

This study shows quantitative analysis of alternative splice variants for the CI cassette exon in rat brain tissues. The quantitative RT-PCR assay was optimized to yield PCR products representative of mRNA splice variants using conditions where only the amount of the RT reaction is a limiting factor in the PCR reaction. For this purpose, PCR reactions were carried out for 20 cycles, where the rate of amplification is in the

linear range. The PCR products were labeled with [α-^{32}P] dCTP because cytidine is distributed evenly in the amplified CI cassette exon region and in both the upstream and downstream flanking exons. The percent exon inclusion was calculated by the percentage of radioactivity in the exon-included band divided by the total radioactivity in both the included and skipped bands. To investigate the spatial and temporal expression patterns of the *GRIN1* mRNA CI cassette exon in the rat brain, RT-PCR was performed with total RNA isolated from various rat brain tissues (Figure 3). The predicted sizes of the PCR products are 235 base pairs (bp) and 124 bp which correspond to the CI included- (L) and skipped-isoforms (S), respectively, when using primers specific for the adjacent exons (Figure 3A). Two predicted PCR products migrate as expected when compared to the DNA molecular marker (M) lane (Figure 3B). The ratios of the two products, which show variations in tested tissues, are presented as the percentage of exon inclusion (%L) in the graph corresponding to the gel panel (Figure 3C). C1 cassette exon inclusion is above 50% only in the hippocampus through all the stages from postnatal days (Pd) 14 to 45 (lanes 5-8), which indicates the major site where the alternative exon is positively regulated among the tested rat brain tissues. By contrast, exon skipping is the predominant form in the brain stem (lanes 1-4). The CI cassette exon shows predominantly skipping patterns and the percentages are different in the cerebellum and pituitary, revealing tissue-specific alternative splicing. CI cassette exon splicing decreases in the cerebellum during postnatal development from 14 to 45 days, indicating developmental stage-specific splicing (Lanes 9-12). This feature of the CI cassette splicing is different from


A
5'
NI
CI
CII
3'
1
2
3
4
5
18
19
20
CI
CII
L (CI included mRNA , 235 bases)
(1.3 kb)
CI
(1.9 kb)
CII
Pre-mRNA
65
111
59
CII
S (CI skipped mRNA, 124 bases)


B
Brain stem
Hippocampus
Cerebellum
Pituitary
(bp)
M
a b c d a b c d a b c d a b c d
← Pd
1,264 -
702 -
224 -
117 -
L
S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 Lane
18 19 18 18 63 58 57 62 49 28 19 15 31 33 28 34 % L

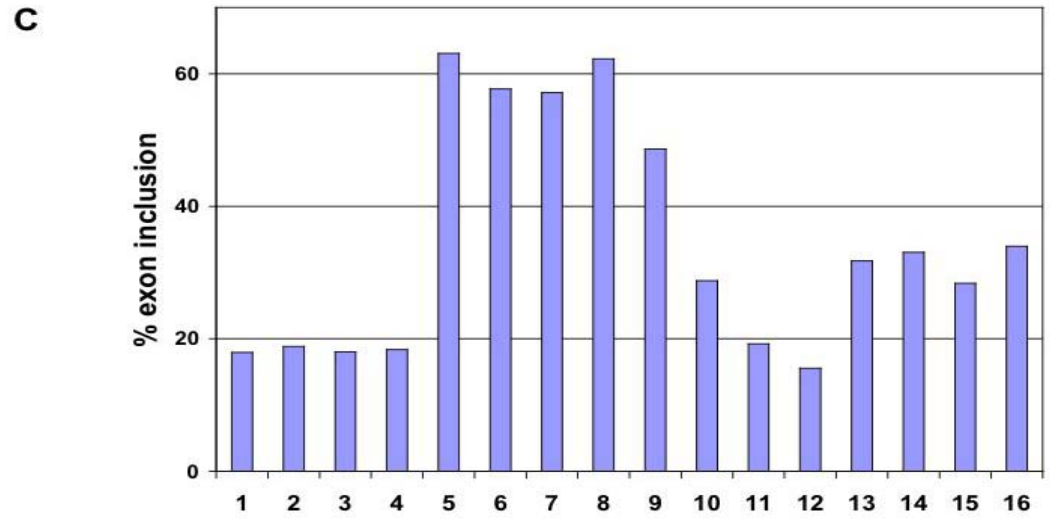

C
% exon inclusion
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 3. Analysis of CI cassette exon splicing pattern in various rat brain tissues during postnatal development.

(A) The genomic structure of rat *GRIN1* (Glutamate NMDA receptor subunit zeta 1 precursor, NR1) is shown in schematic at top. Exon numbering is according to the Ensemble database (*www.ensembl.org*); 5' untranlsated region (UTR) and 3' UTR, white boxes; constitutive exons, black boxes; alternative exons (red boxes) include NI, CI and CII cassettes. Alternative splicing of the CI cassette exon generates exon included (L) or skipped (S) products from one pre-mRNA. The CI cassette exon length is 111 nucleotides. Arrows indicate primers used for PCR amplification (NMDA3021; NMDA 3255A, sequences are shown in Appendix B). Numbers below pre-mRNA indicate the regions amplified by RT-PCR (20 cycles). Sizes of amplified DNA products are indicated in base pairs (bp). (B) The PCR products are resolved on a 6% polyacrylamide, 5M urea gel (lanes 1-16). Total RNA purified from brain stem, hippocampus, cerebellum and pituitary, and postnatal day 14 (a), 21 (b), 28(c), and 45 (d) are indicated above the gel panel. The exon included product is 235 base pairs (bp) in length and the skipped product is 124 bp in length. DNA molecular weight markers, Phage Lambda DNA digested by BstEII; lane M. (C) Graph shows, for each sample, the percentage of exon inclusion (%L) as calculated by, %L = 100 [L / (L + S)]. Radioactivity of the band representing either the exon included or skipped product was quantified by phosphorimager analysis.

several other neural pre-mRNAs, such as the GABA receptor γ2 subunit, of which splicing increases in the cerebellum during postnatal development and is regulated by PTB (Wang and Grabowski, 1996). The brain stem, hippocampus, and pituitary shows overall slight variations in exon inclusion during postnatal development. These results demonstrate that CI cassette splicing is subject to both tissue- and developmental-specific regulation in the rat brain, which is different from pre-mRNA targets controlled by PTB.

2.3.2. An *in vivo* splicing assay for the study of splicing regulatory elements of the CI cassette exon.

An *in vivo* splicing assay system was established by initially screening for alterations in the splicing pattern of the CI cassette exon. The splicing reporter of CI cassette exon was constructed as a three-exon substrate. The flanking intron lengths are long (1,257 bp, upstream intron; 1,913 bp, downstream intron) as shown in Figure 4A. Two pieces of PCR fragments were amplified, ligated at the XhoI-end and subcloned into the mammalian expression vector pBSVA$^+$ to generate the splicing reporter. Regions of the CI cassette exon, the upstream flanking exon and its intron were amplified by PCR from the rat genomic DNA (primers listed in Appendix A). The downstream intron and its flanking exon were also obtained. The insert including entire intron regions in the vicinity of the regulated exon and its flanking exons was sequenced (Genebank accession number AY090615, published in Zhang et al., 2002).

Cell lines that originated from neuronal tissue ST15A (CNS) and RC-4B/C (pituitary) and non-neuronal tissues HeLa (epithelial) and C2C12 (muscle) were

transfected with the CI cassette exon splicing reporter E21wt. The splicing reporter was subcloned into the mammalian expression vector pBSVPA^{+} for transfection of C2C12 cells (Figure 4A, E21wt). RT-PCR analysis was performed using total RNA purified from the transfected cells. The splicing patterns are summarized in Table 2, which illustrates that exon skipping is predominant regardless of cell line origin. For the reason that RC4-B/C is derived from non-neuronal origin, only ST15A exhibits neuronal properties in this study (Zhang et al., 1996). The partial neuronal phenotype of ST15A may explain why CI cassette exon splicing is not significantly different in between ST15A (33.4%) and the non-neuronal cell lines (20~21%). Since a previous study showed a positive role for NAPOR in CI cassette exon splicing (Zhang et al., 2002), the non-neuronal property of the tested cell lines is useful to establish an *in vivo* system to scrutinize the silencing mechanism of CI

Table 2. Percentage of CI cassette exon inclusion in various mammalian cell lines.

Cell Type	Cell Line	% Exon Inclusion [a] ± STDEV (sample #)	Neuronal Phenotype
human epithelium	HeLa [b]	20.3 ± 4.53 (n=2)	No
mouse myoblast	C2C12 [c]	20.8 ± 0.10 (n=2)	No
rat pituitary	RC4B/C [d]	21.7 (n=1)	No
mouse cerebellum	ST15A [d]	33.4 ± 1.44 (n=2)	Yes (partial)

[a] The percentage of exon inclusion was determined by transient expression of the splicing reporter plasmid, E21-wt. Refer to figure 4 (A); [b] Salzman, 1961; [c] Zhang et al., 2002 ; [d] Zhang et al., 1996.

cassette exon splicing. In this study, C2C12 cells were chosen for further analysis because they exhibit high transfection efficiency as well as a non-neuronal phenotype. The CI cassette exon splicing reporter is a useful minigene system to study its silencing mechanism by mutational analysis in C2C12 cells.

2.3.2.1. The CI cassette exon is moderately sensitive to engineered restriction enzyme sites.

For the convenience of cloning splicing reporters, an XhoI endonuclease restriction enzyme site was generated in the CI cassette exon region, and a NotI and a BamHI site were engineered into the intron. To rule out the possibility of the effect of the introduced restriction sites on CI cassette exon splicing, several derivatives of the wild type plasmid were tested in the transfection assay (Figure 4A). CI cassette exon inclusion is moderately sensitive to engineered restriction enzyme sites. A generated XhoI, which disrupts a SC35 motif at position 73 in the exon (Figure 11A), induces to slightly decrease exon inclusion (Figure 4, lanes 1 and 2). Generation of NotI and BamHI sites in the introns increases exon inclusion (Lanes 3 and 4). Because of the predominant exon skipping pathway and the small variation of splicing patterns, these restriction enzyme sites were utilized to generate deletion, shuffled mutants and chimeric constructs.

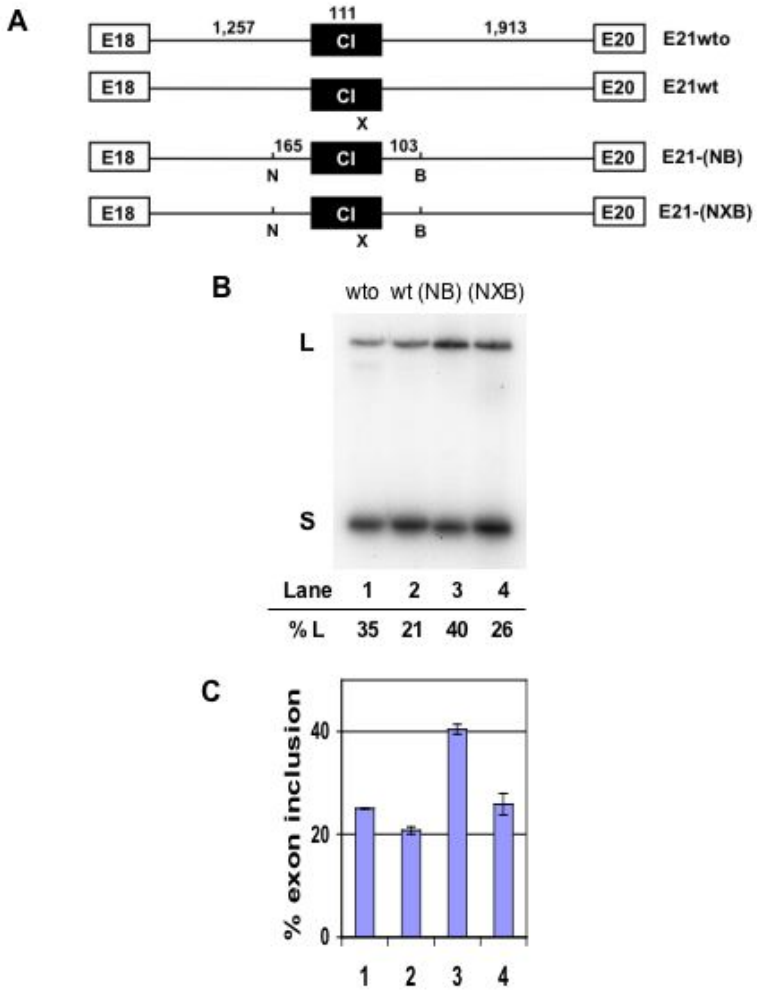


Figure 4. Sensitivity of CI cassette exon inclusion to engineered restriction enzyme sites.

(A) The splicing reporter constructs containing restriction sites in the CI exon and the introns are shown. An XhoI site (X) is at position 78 in the CI cassette exon (E21-wt and E21-NXB), which is generated by a point mutation (GCCTCCAGCT → GCCTCgAGCT). NotI (N) and BamHI (B) sites are engineered at 164 nucleotides upstream from the 3' splice site and at 103 nucleotides downstream from the 5' splice site (E21-NB and E21-NXB), respectively. The exon and the intron lengths are indicated above the construct. (B) Splicing reporter plasmids were cotransfected into mouse myoblast C2C12 cells and the splicing patterns analyzed by RT-PCR. The PCR products are shown in the gel panel (6% polyacrylamide, 5M urea gel). (C) Graph shows the percent of exon inclusion (%L) as described in Figure 3.

2.3.2.2. The CI cassette exon is insensitive to PTB.

PTB, a well-known splicing repressor, is involved in regulation of the $GABA_A$ receptor $\gamma 2$ subunit neuron-specific exon. The CI cassette exon does not contain functional PTB binding motifs in the upstream and downstream intron and the regulation profile during development of rat shows no similarity to that of the exons regulated by PTB (Wang and Grabowski, 1996). Based on the lack of significant PTB binding motifs we would expect that CI cassette exon splicing should not be regulated by PTB. To test this idea, C2C12 cells were co-transfected with a splicing reporter E21wt and a PTB expression plasmid pcDNA/PTB(A1) in a 1:2 or 1:6 ratio of splicing reporter to expression plasmid (Figure 5A). Whereas coexpression of PTB at a two or six-fold excess shows the complete repression of $\gamma 2$ exon inclusion, the same ratio of PTB does not seem to affect the exon inclusion of E21wt (Figure 5B, compare lanes 1-3 and 4-6). Because we cannot rule out that the effect on E21wt could not be detected, a chimeric construct rGγE21, where the splicing pattern of is prominently exon inclusion, was employed to more rigorously test the effect of PTB.

The construct, rGγE21, contains the exon and adjacent intron region, excluding the PTB binding motif of rGγ25. As a comparison, rGγ4-1(-)up containing the PTB binding motif of $\gamma 2$ was used, which is a derivative of rGγ25. PTB has been shown to repress exon inclusion of this construct significantly in a previous study (Zhang et al., 1996). PTB does not repress the splicing of rGγE21 whuch has no PTB binding motif in contrast to rGγ4-1(-)up (compare lanes 7-9 and 10-12). Therefore, these results indicate

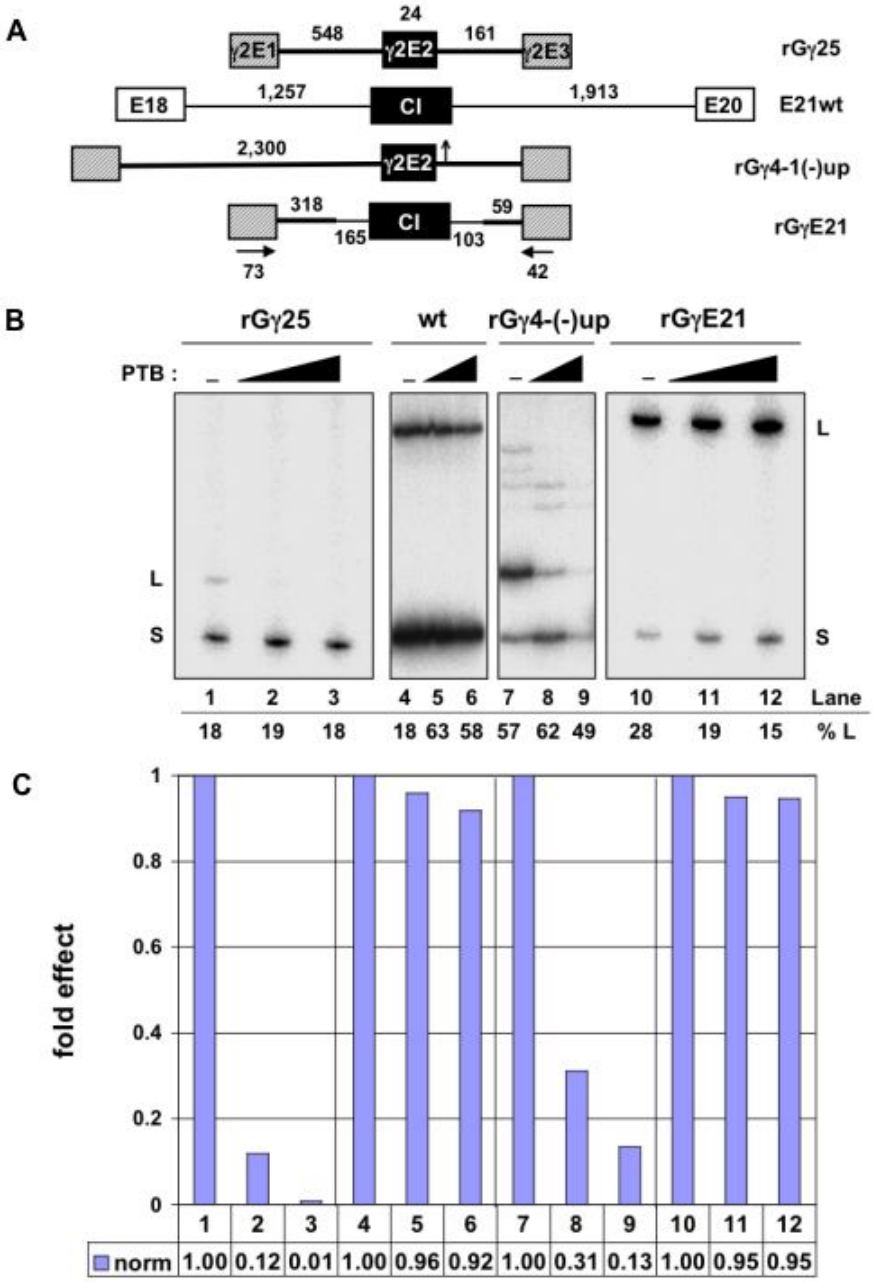


Figure 5. CI cassette exon splicing is relatively insensitive to regulation by PTB.

(A) Schematic of the splicing reporter plasmids used in the transfection assays is shown. An arrow (↑) indicates a 5' splice site mutation that increases base pairing with U1 snRNA in rGγ4-1(-)up (Zhang et al., 1996). The splicing reporter plasmid of the CI cassette exon, E21wt, contains the full upstream and downstream introns and partial flanking exons. The chimeric contstruct, rGγE21, consists of the CI cassette exon and the adjacent introns (thin lines) inserted between constitutive exons of rGγ25 (thick lines indicate γ2 introns). Exon lengths amplified by RT-PCR are indicated immediately below the arrows. The length of γ2 neuron-specific exon is marked above. Intron lengths are marked above or below the introns. (B) Splicing reporter plasmids were cotransfected into mouse myoblast C2C12 cells with a pcDNA vector backbone (lanes -) or a pcDNA/PTB expression plasmid (wedge, 1/2 or 1/6 ratio of splicing reporter plasmid to pcDNA/PTB). The PCR products are resolved on a 6% polyacrylamide, 5M urea gel. The percent exon inclusion values are shown (bottom). (C) Graph shows the fold effect of PTB expression on the pattern of splicing. Each value in the graph was calculated by dividing the percent exon inclusion (%L) of the splicing reporter co-transfected with PTB by the percent exon inclusion (%L) with vector backbone.

that CI cassette exon is not regulated by PTB although the fold effect in the graph (Figure 5C) shows a slight decrease in exon inclusion. This implies that the regulation mechanism of CI cassette exon splicing could be different from a repression mechanism by PTB.

2.3.2.3. The CI cassette exon sequence is not required for positive regulation by NAPOR.

Positive regulation of the CI cassette exon is correlated with the regions of expression of ETR-3 family proteins and transient overexpression of NAPOR increases exon inclusion (Zhang et al., 2002). ETR-3 is known to activate exon inclusion by interacting with UG-rich motifs in muscle-specific enhancer elements (MSEs) in cardiac troponin T (cTNT) pre-mRNA (Ladd et al., 2001; Charlet et al., 2002). The CI cassette exon contains potential binding motifs for ETR-3 GUGGU, CUG and GUAGGU. To test whether or not the CI cassette exon region contains sequence elements involved in an enhancing activity of NAPOR, the E21-ΔABC mutant was designed where most exonic region (from position 15 to position 95) containing the potential NAPOR binding sequences was deleted (Figure 6A). E21-ΔABC was coexpressed in C2C12 cells with a NAPOR expression plasmid. The E21-ΔABC mutant is positively regulated to the same extent as that of the wild type (figure 6B and 6C), suggesting that the CI cassette exonic region is not required for an enhancing role of NAPOR. This result, combined with the result of PTB effect, demonstrates that the CI cassette exon does not contain NAPOR binding motifs in the exon region and PTB binding motifs in the intron regions.

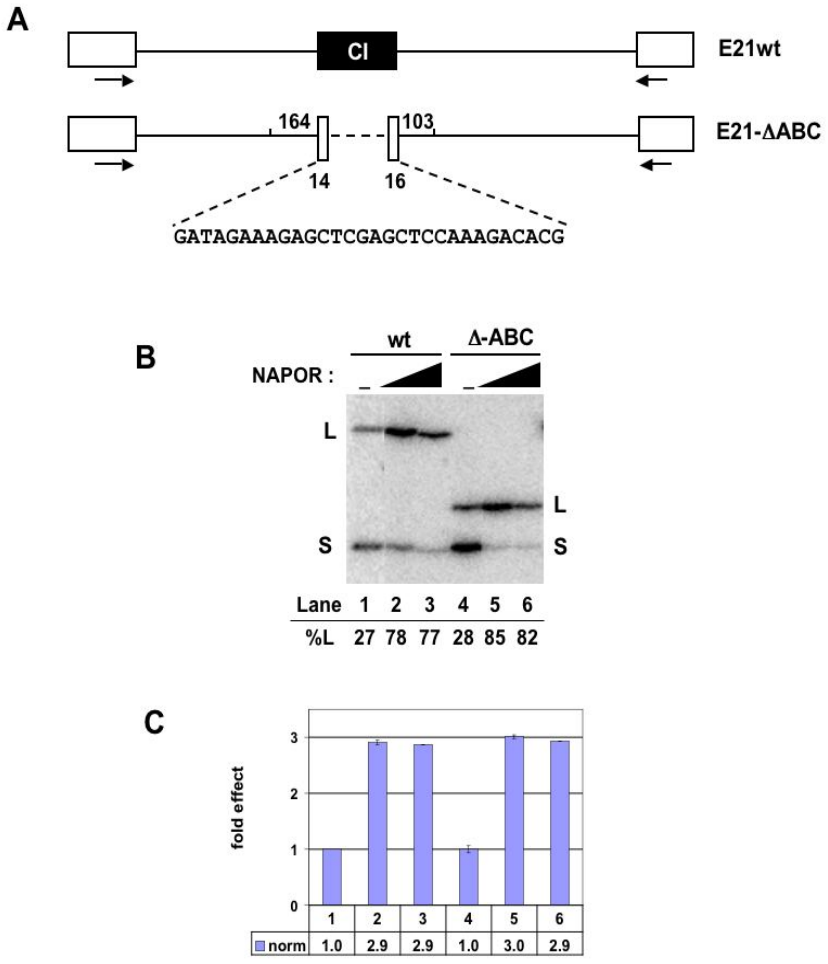


Figure 6. Lack of requirement for internal CI cassette exon sequences for enhancing effect of NAPOR.

(A) Schematic of wild type and deletion mutant splicing reporter constructs of the CI cassette exon; dashed line indicates deleted region. Most of the exon region is deleted in E21-ΔABC mutant except 11 nucleotides at the 5' end and 13 nucleotides at the 3' end of the CI exon plus 3 nucleotides of a XhoI restriction enzyme site for each end. (B) Splicing reporter plasmids were transiently expressed by co-transfection of C2C12 cells with a pcDNA/NAPOR expression plasmid. Control lanes (-) stand for mock transfection with the pcDNA backbone plasmid. Wedges indicate 2/1 or 6/1 mass ratio of protein expression plasmid to splicing reporter. The percent of exon inclusion (%L) is shown below the gel. (C) Graph shows the fold effect of NAPOR on the splicing reporters. Each value was normalized as described in Figure 5.

2.3.3. The CI cassette exon 5' splice site sequence is unique and conserved in mammals.

As the CI cassette exon and the flanking introns are sensitive to restriction enzyme sites, there is a possibility that splicing regulatory sequence elements exist in these regions. To identify the RNA *cis*-acting elements required for regulation of CI cassette exon splicing, two criteria were considered for searching the regions. Based on the idea that alternative exons have weak splice signals compared to constitutive exons, splicing regulatory elements are predicted in the vicinity of the regulated exon. Computational analysis of 25 brain-specific alternative exons confirmed this idea by the finding that key features of control sequences within 100 nucleotides of the adjacent introns were as follows: divergent 5' splice sites; highly pyrimidine-rich upstream introns; lack of GGG motifs in the first 100 bases of downstream intron; over-representation of an enhancer motif UGCAUG in the downstream intron (Brudno et al., 2001). From their analysis, alternative 5' splice sites are classified as consensus (:GURAGU, 28%), rare consensus sites (56%) occurring in less than 1% of the control exon and non-consensus sites (16%). Rare consensus sites are found in approximately 50% of brain-specific exons but the CI cassette 5' splice site (:GUAAGG) is not a rare consensus sequence. Unlike other brain-specific exons, the intron upstream from the CI cassette exon is less pyrimidine rich within 100 nucleotides (48% vs 70%) and the intron downstream from the CI cassette exon is missing the enhancer motif UGCAUG. However, one GGG motif at position +6 in the downstream intron of the CI cassette exon is observed, which is a feature of brain-specific exons. Also, the intron further downstream contains C-rich sequences at position +17, +23 and +41. Therefore, compared to brain-specific exons

that have been studied, sequence features of the CI cassette exon are unique. This is consistent with a feature of the developmental regulation profile that is different from a regulation by PTB. Given the observation that conservation of important functions from lower to higher eukaryotes is derived from the conserved DNA sequence of genes, important splicing regulatory elements are also expected to be conserved in alternative exons. The sequences of the CI cassette exon, the splice sites and 180 nucleotides of the upstream and downstream intron were aligned and compared among vertebrates including human, mouse and rat (Figure 7). The comparison shows a high degree of conservation of the CI cassette exon. The rat exon sequence is identical to the orthologous mouse exon sequence and is 97.3% identical to the human ortholog. The 20 intronic nucleotides upstream from the 3' splice site and the 8 nucleotides downstream from the 5' splice site are identical among the three species and are similar to the splice site consensus sequence (Figure 7B). The sequence conservation is still seen farther into the downstream intron but the extent of homology is not as good as in the region surrounding the 5' splice site. The U1 snRNA complementary region of the 5' splice site is normally bases +1 to +6 of the intron in mammals, which is weakly conserved when compared to yeast (Burge et al., 1999). The CI cassette exon contains 6 nucleotides complementary to U1 snRNA,

A

```
    -180                -160                -140
Human  caggagaagaggccaccctcgaacgtccg-ctgtcggcccgtctgtccagcacagggagg
mouse  cag-a-a-gaggccactcttgaaagtctggctgttggtccgtcttttcagcacagtgagg
rat    cag-a-a-gaggccactcttgaaagtctggctgttggtccgtctgttgagcacagtgagg
       *** * * ******** ** *** *** * **** ** ****** *  ******* ****

    -120                -100                -80
Human  caggcaggagcgagggcccaagtggccggccaggctgggcagcggcccatgcaggagcag
mouse  caggcaggagtcagggcctgagtgacctgccaggcctggcagcggccc-ca-aggagcag
rat    caggcaggagtcagggcctgagtgaccggccaggcctggcagcggccc-ca-aggagcag
       **********  ******  **** ** *******  ***********    ********

     -60                -40                -20
Human  gcgagggcaggtgtggccaccaccct-agccatctaatcacttatacatattcattttag
mouse  gcaagggcaggtgtggccgccacccttagttgtctaatcgcttatacatattcattttag
Rat    gcaagggcaggtgtggccgccacgcttagctgtctaatcgcttatacatattcattttag
       ** ***************** **** ** **   ******* ********************

                          20                  40                  60
Human  GATAGAAAGAGTGGTAGAGCAGAGCCTGACCCTAAAAAGAAAGCCACATTTAGGGCTATC
mouse  GATAGAAAGAGTGGTAGAGCAGAGCCCGACCCTAAAAAGAAAGCCACATTTAGGGCTATC
rat    GATAGAAAGAGTGGTAGAGCAGAGCCCGACCCTAAAAAGAAAGCCACATTTAGGGCTATC
       ************************** ********************************
        D  R  K  S  G  R  A  E  P  D  P  K  K  K  A  T  F  R  A  I
                       780                 880

                          80                  100          +1
Human  ACCTCCACCCTGGCTTCCAGCTTCAAGAGGCGTAGGTCCTCCAAAGACACGgtaaggggg
mouse  ACCTCCACCCTGGCCTCCAGCTTCAAGAGACGTAGGTCCTCCAAAGACACGgtaagggga
rat    ACCTCCACCCTGGCCTCCAGCTTCAAGAGACGTAGGTCCTCCAAAGACACGgtaagggga
       ********************************************************
         T  S  T  L  A  S  S  F  K  R  R  R  S  S  K  D  T
                       890                 900

              +20                 +40                 +60
Human  agagcaccccagtcccgcgtc-c-gactccacctgccctgccctgcgtgtgtctcccgc-
mouse  ag--caccccagtcccgcgtctctgactcc-cc--ccttgccccatgtctgtctccctct
rat    agagcaccccagtcccgcgtctctgactcc-tc--ccttgccccgtgtctgtctccctct
       **  ****************** * ******  *  ********   ** ******** *

              +80                 +100                +120
Human  ccc--atcaccccgccccggaccctgggctc--ctgtggcccactctgc-ccctgtct-c
mouse  cccttaccatcccacccctgcccccatcttcaaccctggcc-a-tctgtgccaagtcagt
rat    cccttaccatcccacctctgcccc-atcttcaaccctggcc-a-tctgtgccaagt----
       ***  * ** *** ** * *****     **  *  ***** * **** **  **

              +140                +160                +180
Human  cctgtggcggccgctctgcccagcccgcccatgctgctctctctcactctctggacctttt
mouse  ccttcacccacc-ctccaccca-ttaaagggtgctgtcctttagaacgctccagacctct
rat    ccttcactcaccctccacccattaaagg--gtgctgtcctttagaacgctccagacctct
       ***       **   *  **           *****  ** *   ** ***  ***** *
```

B

C

Figure 7. Sequence alignment of the CI cassette exons and the flanking introns from human, mouse and rat *GRIN1*.

(A) Nucleotide sequences of human (ENSG00000176884), mouse (ENSMUSG00000026959) and rat (ENSRNOG00000011726) *GRIN1* were obtained from the Ensembl database. Alignment was performed using the MacVector 7.0 program (Gap opening penalty, 10.0; Gap extension penalty, 5.0Oxford Molecular Ltd.). Uppercase and lowercase letters represent the exon and intron sequence, respectively. Deduced amino acid sequence from the DNA sequence of the CI cassette exon is shown as a one-letter abbreviation code. An asterisk (*) is used to represent identical nucleotides in human, mouse and rat. A dash (-) is used to indicate the positions of the gaps between nucleotides. Numbers indicate the positions of the nucleotides. A plus (+) sign indicates the positions of the nucleotides in the downstream intron from the 5' splice site, and a minus (-) sign indicates the positions of the nucleotides in the upstream intron from the 3' splice site. Arrows indicate the splice junctions. The complementary sequence of the 5' splice site to U1 snRNA is shown in red. Note that nucleotide sequences of the exon and the surrounding intron regions are highly conserved. (B) The consensus sequence of the 5' splice site in humans (Burge et al., 1999; Lim and Burge, 2001) is shown above the CI cassette 5' splice site; red, cytidine; green, adenosine; blue, guanosine; yellow, thymidine. The height of a specific nucleotide represents the relative frequency of its occurrence. (C) Complementarity of the CI cassette exon 5' splice site to the 5' end of U1 snRNA. Eleven nucleotides each from the CI cassette 5' splice site region and U1 snRNA are shown. The red color indicates the sequence of the 5' splice site complementary to U1 snRNA. The red diamond stands for base pairing between the 5' splice site and the 5' end of U1 snRNA.

which are from +1 to +5 of the intron and the last nucleotide of the exon (Figure 7C). Therefore, the CI cassette 5’ splice site has complementarity sufficient for base pairing with U1 snRNA in the normal range.

2.3.4. G cluster at the 5’ splice site is required for CI cassette exon silencing.

A result from the mutational analysis of the CI cassette exon splicing reporter by a transient expression assay and a gel mobility shift assay (Zhang et al., 2002) suggested that an intronic RNA element acts as a silencer. Considering that deletion of the intronic silencer region or overexpression of NAPOR protein does not completely rescue CI cassette exon inclusion, unidentified silencer elements could mediate regulation of splicing. In addition, exon skipping is dominant in both neuronal and non-neuronal cell lines in the absence of NAPOR, suggesting that silencing mechanism of CI cassette exon splicing is dominant in most cell types. Taken together, the CI cassette exon could be predicted to contain negative *cis*-acting splicing elements in the vicinity of the exon, the splice sites and/or the branch point.

2.3.4.1. U1 snRNP is not stably associated with the 5’ splice site in a biotin-affinity selection assay

A 5' splice site is recognized by U1 snRNA through base pairing, thereby initiating pre-mRNA splicing. The 5’ splice site of an alternatively spliced exon is generally less conserved compared to a constitutively spliced exon. On average, 5’ splice sites of

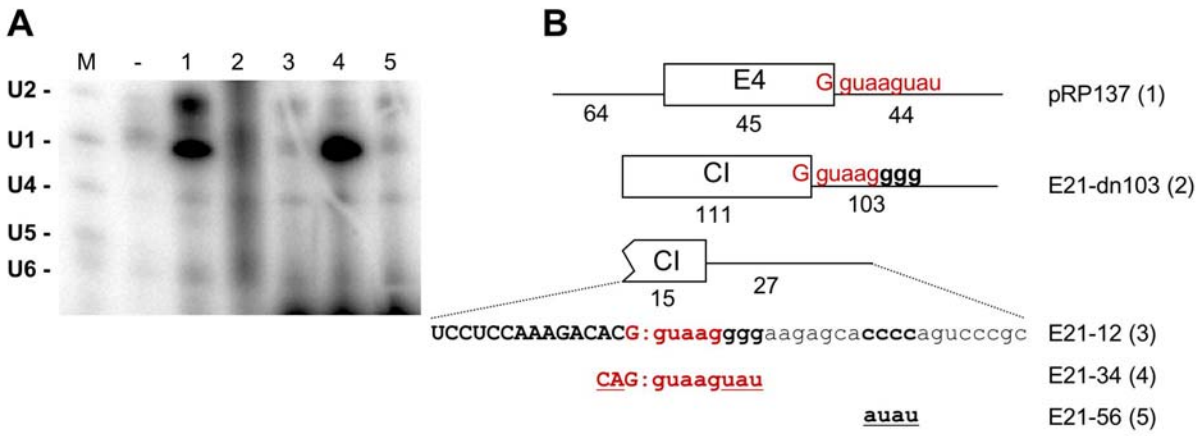


Figure 8. Biotin-affinity selection assay for U1 snRNP binding to the 5' splice site.

RNA substrates were biotinylated at random positions using biotin-16-UTP and T3 RNA polymerase. To assemble splicing complexes, biotinylated RNAs were incubated under splicing reaction conditions with HeLa nuclear extracts. After digestion of bound protein complexes with proteinase K, the 3' end of the RNAs were labeled using 5'-^{32}P-[pCp] (cytidine-3',5'-bis-phosphate) and RNA ligase. (A) The RNA samples were separated on a 7 M urea/10% polyacrylamide gel. Lane M, snRNA marker of HeLa nuclear extract; lane -, non-biotinylated pRP137 control; lane 1, biotinylated pRP137 RNA; lane 2, E21-dn103 RNA; lane 3, E21-12 RNA; lane 4, E21-34 RNA; lane 5, E21-56 RNA. (B) A transcript of pRP137, which is a positive control, is schematically shown. E21-dn103 RNA is an RNA substrate containing CI exon and the 103 nucleotide downstream intron (256 nt). E21-12 RNA contains the wild type 5' splice site including 15 nt exon sequences and 27 nt downstream intron sequences. Mutation at the 5' splice site of E21-34 RNA generates four more base pairs than wild type with U1 snRNA. E21-56 RNA contains a mutation of CCCC to UAUA at position +16 in the intron region downstream from the 5' splice site.

mammals have 6-9 nucleotides that are complementary to the 5' end of U1 snRNA (Burge et al., 1999). The CI cassette exon has six nucleotides at the 5' splice site that are complementary to U1 snRNA (Figure 7 and Figure 8, corresponding sequences are in red). Although the CI cassette exon has an average number of nucleotides to base pair with U1 snRNA, a biotin-labeled CI cassette exon transcript E21-12 fails to bind U1 snRNA (Figure 8). U1 snRNA binding is detected only when the base-pairing is increased to 11 nucleotides by mutation from ACG:GUAAGGGG to CAG:GUAAGUAU at the 5' splice site (E21-34). Failure of its binding to U1 snRNA indicates that the 5' splice site has two possible functions: formation of 1) a secondary structure with complementary nucleotides or 2) a silencing complex with additional *cis*- or *trans*-acting elements.

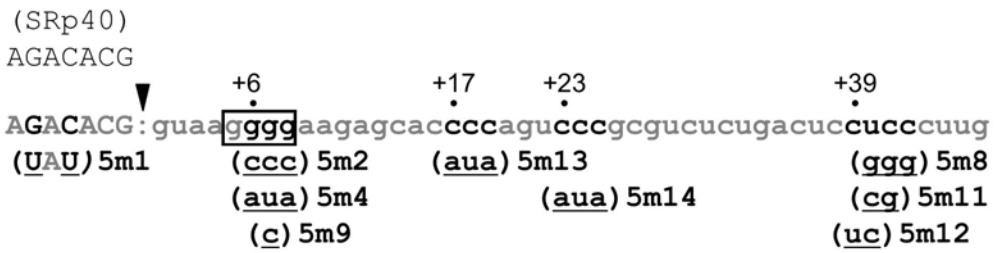
2.3.4.2. An intronic G cluster at the 5' splice site is required for silencing of CI cassette exon splicing.

From the result of the biotin-affinity selection of the 5' splice site (Figure 8), the CI cassette exon does not appear to be a strong exon for splicing although the splice site does have a consensus sequence. The CI cassette 5' splice site contains a G cluster immediately in the downstream intron followed by C-rich sequences (Figure 7). The role of the G cluster in splicing silencing of the CI cassette exon is examined by using site-directed mutations on both the G cluster and the downstream C clusters (Figure 9A). Mutations at positions +6, +7, and +8 of the intron were designed to not disrupt the complementarity to U1 snRNA, in order to exclude the possibility that splicing is affected by changes in the splice site consensus. Mutants were transfected into C2CI2 cells and the splicing patterns were analyzed by RT-PCR from purified total RNAs. Site-directed

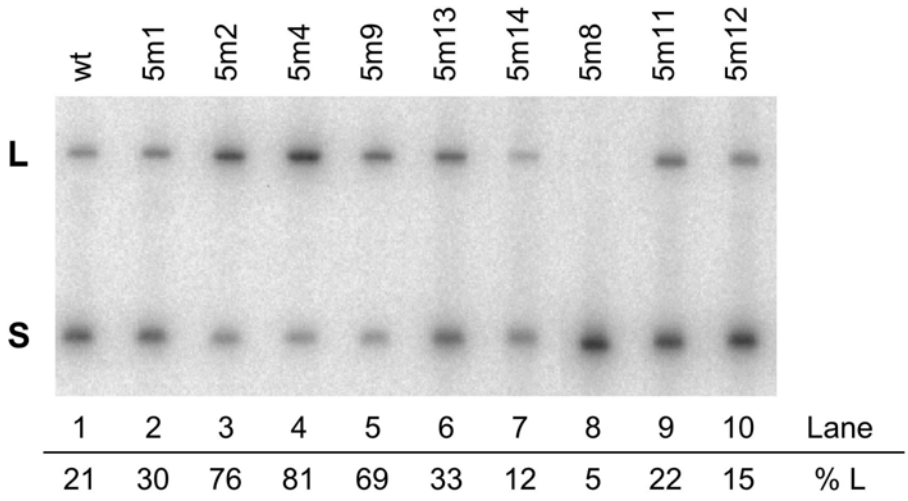
mutations in the G cluster at the 5’ splice site and the downstream intron sequences show that mutations on any nucleotides of the G cluster significantly increase CI cassette exon inclusion three to four fold (Figure 9, 5m2, 5m4 and 5m9). By contrast, mutations in exonic sequences near the 5’ splice site (5m1) and intronic sequences in the downstream C clusters (5m13, 5m14, 5m8, 5m11 and 5m12) show modest changes. Interestingly, replacement of the C cluster at position +40 by a G triplet (5m8) represses CI cassette exon inclusion four-fold, which demonstrates the additive silencing effect of the G cluster. 5m5 is a compensatory mutant of 5m2 that contains one G cluster at position +40 but not at position +6. This induces a slight decrease in exon inclusion compared to 5m4 (Figure 10). Similarly, the wild type construct, which contains a G cluster at position +6 causes a significant decrease in exon inclusion compared to mutants that do not possess the G cluster (5m2, 5m4, and 5m9). Thus, splicing silencing by the G cluster is stronger at position +6 than +40. This suggests that the G cluster is a position-dependent intronic splicing silencer (ISS) showing stronger silencing when it positioned closer to the 5' splice site. Therefore, the position-dependent silencing activity of the G cluster demonstrates that the specific GGG primary sequence is important for splicing silencing.

The silencing mechanism by the G cluster may involve formation of a specific secondary structure rather than a specific primary sequence. Indeed, a stem and loop structure is predicted via base pairing of the G cluster with a downstream C cluster at +39 (Mulfold 2.0, Jaeger et al., 1989; Zuker, 1989). This structure is expected to inhibit binding of positive splicing regulators to the enhancer motifs (as shown below), thereby playing negative roles in recognition of the 5' splice site by U1 snRNA. Whereas

A


(SRp40)
AGACACG
+6
+17
+23
+39
AGACACG:guaaggggaagagcaccccaguccccgcgucucugacuccucccuug
(UAU) 5m1
(ccc) 5m2
(aua) 5m4
(c) 5m9
(aua) 5m13
(aua) 5m14
(ggg) 5m8
(cg) 5m11
(uc) 5m12

B


wt
5m1
5m2
5m4
5m9
5m13
5m14
5m8
5m11
5m12
L
S
1 2 3 4 5 6 7 8 9 10 Lane
21 30 76 81 69 33 12 5 22 15 % L

C

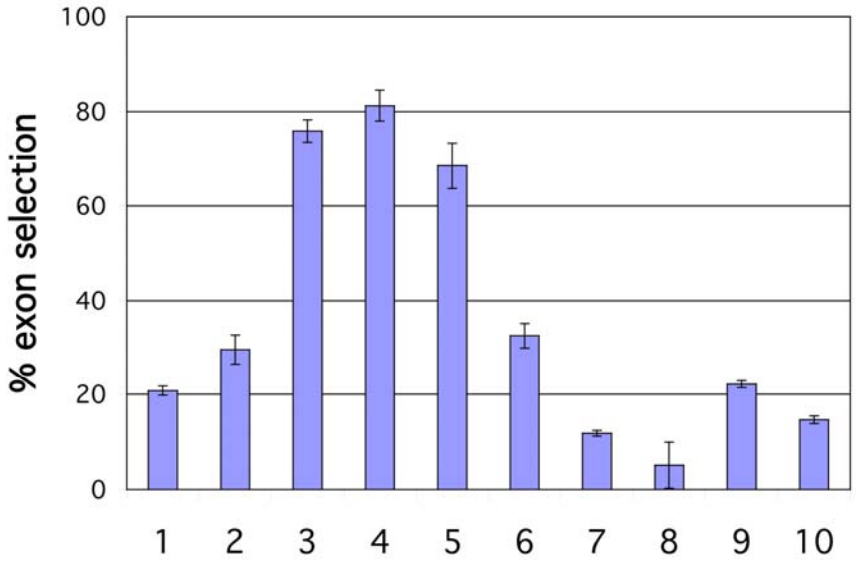

% exon selection
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10

Figure 9. Effect of the G cluster at the 5' splice site on the splicing pattern of the CI cassette exon.

(A) Sequence of the wild type 5' splice site of the C1 cassette exon encompassing an exonic SRp40 enhancer motif and 46 nucleotides of the downstream intron is shown. Exonic sequences, uppercase; intronic sequences, lowercase. Intron positions are represented above the base. A G cluster shown in a box overlaps the 5' splice site consensus sequence at position +5 and +6. The 5' splice site is indicated by an arrowhead. The mutation in 5m1 disrupts the SRp40 motif. Other mutations are designed to disrupt the G cluster and proximal downstream C clusters. Wild type sequences that were mutated, indicated in bold, were replaced with the underlined sequences in parenthesis. (B) RT-PCR analysis was performed using total RNA samples purified from transfectants. The PCR products are shown on a 6% polyacrylamide, 5M urea gel. The percent exon inclusion of the CI cassette exon is quantified at the bottom of the gel. (C) The graph shows the percentage of C1 exon inclusion of wild type and the mutants in order from the 5'- to 3'-end. The percentage of exon inclusion (%L) is calculated as described in Figure 3.

mutations on the G cluster switch the splicing pattern from skipping to inclusion, mutations on C clusters show mild changes. Mutation in the 5m11 mutant that weakens the secondary structure by decreasing one base pair shows a minimal increase in exon inclusion. The CU to UC mutation in the 5m12 mutant strengthens the base pairing by changing GU to GC, whose change is from wobble pairing to Watson-Crick base pairing. As expected, the 5m12 mutant splicing pattern shows more skipping than wild type. This implies that the secondary structure model might be a reasonable model for CI cassette silencing. Another idea to consider is that the C cluster has been known as an enhancer motif by binding to hnRNP K/J in a different system (Matunis et al., 1992; Dreyfuss et al., 1993; Expert-Bezancon et al., 2002). Since the CI cassette exon has an extended cytidine rich (C-rich) region in the downstream intron, it is possible that the GGG sequence could base pair with another nearby C cluster sequence.

2.3.4.3. Secondary structure model lacks strong evidence.

A stem and loop structure composed of 15 nucleotides in the 5' splice site, which is the most stable structure predicted by program Mulfold 2.0 ($\Delta G = -33.4$ kcal), might block U1 RNA from accessing the site (Figure 10A). Eight nucleotides near the 5' splice site base-pairs with an intronic NAPOR binding sequence GUGU in stem I, and stem II is composed of seven base pairing, which contains a U1 snRNA binding sequence UAAG and an hnRNP K binding sequence CCC. This observation gives rise to one hypothesis that the stem and loop structure composed of regulatory elements forms a repressor structure to regulate CI cassette exon splicing. To examine how the stem and loop structure affects splicing, compensatory mutations were created to restore the

secondary structure (Figure 10A and Table 3), which might have been disrupted by the mutations shown in Figure 9. The 5m1 mutant, which disrupts base pairs between the 5' splice site and NAPOR, shows a slight increase in exon inclusion. The 5m2 mutant, which is mutagenized to break stem II, increases CI cassette exon inclusion. The 5m3 mutant, containing the 5m1 and 5m2 mutations, is expected to partially disrupt both stem structures. However, this double mutant lessens the effect of the CCC mutation in the 5m2 mutant. Given the result that disruption of an SRp40 motif in the 5m1 mutant even enhances exon inclusion (Figure 9A), this motif might function as an enhancer element when both stem structures are disrupted. Therefore, the stem structures seem to be stabilized once either stem structure is formed. Mutations in 5m1 and 5m2 disrupting one stem structure result in a partial anti-silencing of the CI cassette. The effect of the latter is more significant due to the loss of one additional G-C base pair (Figure 10B, lanes 2, 4). The 5m3 mutant does not form stem structures as a silencing complex, thus, exon inclusion is reduced compared to 5m2 due to disruption of the SRp40 motif. Other compensatory mutations (5m5, 5m6 and 5m10) are predicted to restore secondary structure disrupted by the single mutations (5m2 or 5m8, 5m4 or 5m7, and 5m9) as shown in Figure 10A (summarized in Table 3). Therefore, 5m5, 5m6 and 5m10 are expected to decrease exon inclusion but exon inclusion increases in all compensatory mutants compared to the single mutations (Figure 10B, lanes 4-11). The mutation in 5m5 decreases exon inclusion, which seems to compensate for the mutation of 5m2. However, the single mutation of 5m8 shows more exon skipping than compensatory 5m5, implying that exon silencing resulted from insertion of a G triplet rather than

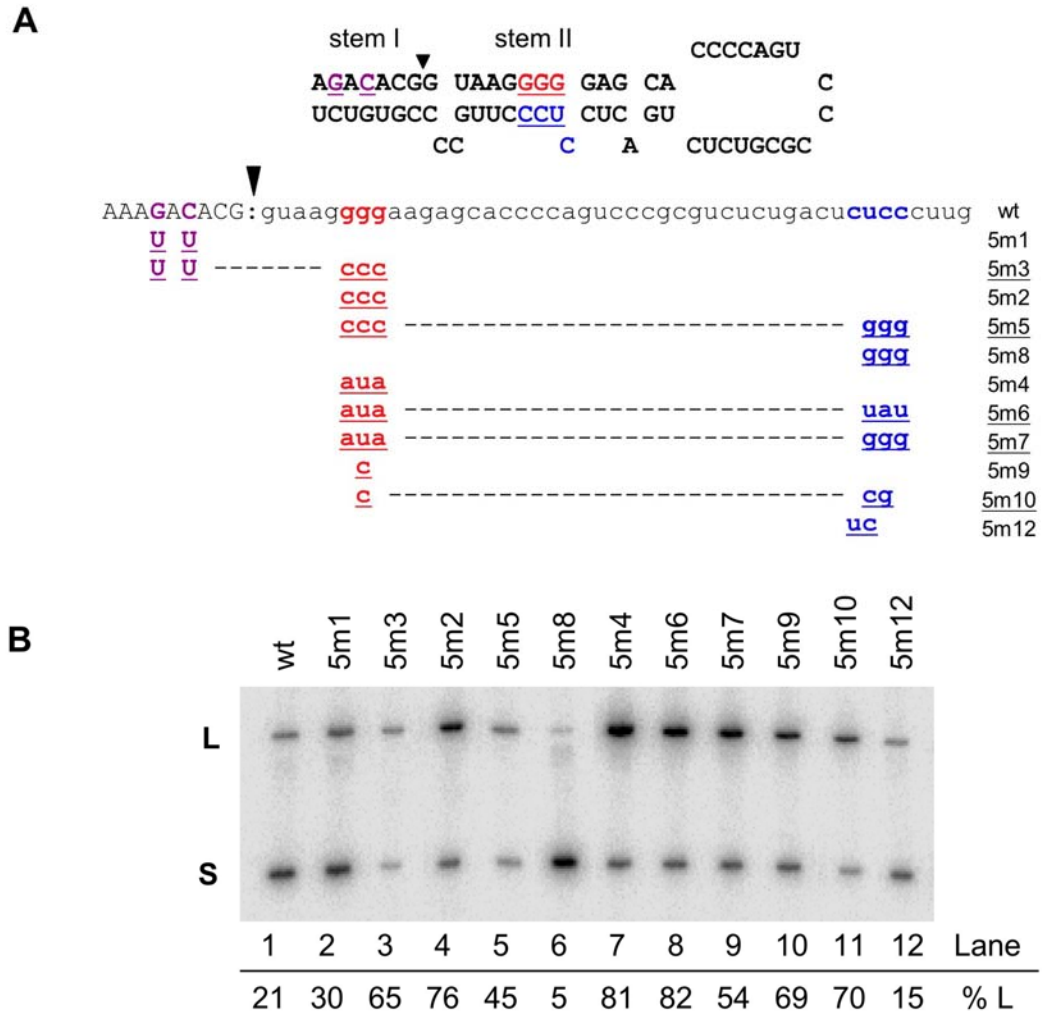

A
stem I
stem II
CCCCAGU
AGACACGG UAAGGGG GAG CA C
UCUGUGCC GUUCCCU CUC GU C
CC C A CUCUGCGC
AAAGACACG:guaaggggaagagcaccccagucccgcgucucugacucucccuug wt
5m1
5m3
5m2
5m5
5m8
5m4
5m6
5m7
5m9
5m10
5m12
B
wt
5m1
5m3
5m2
5m5
5m8
5m4
5m6
5m7
5m9
5m10
5m12
L
S
1 2 3 4 5 6 7 8 9 10 11 12 Lane
21 30 65 76 45 5 81 82 54 69 70 15 % L
C
norm to wild type
0
1
2
3
4
1 2 3 4 5 6 7 8 9 10 11 12
norm 1.0 1.4 3.1 3.7 2.2 0.2 3.9 3.9 2.6 3.3 3.4 0.7

Figure 10. Compensatory mutational analysis of the secondary structure model of the CI cassette exon.

(A) The predicted secondary structure of the CI cassette exon 5' splice site region as determined by the program MULFOLD on top (Zuker, 1989; Jaeger et al., 1989). The 5' splice site, indicated by the arrowhead, is within stem I. Stem 1 contains a binding motif for SRp40 (AGACAG) and a potential binding motif for NAPOR (GUGUCUG). The G cluster is positioned in stem II due to its base pairing with the downstream UCC sequence. The UCC sequence overlaps a binding motif for hnRNP K (CCC). The sequence of stem I chosen for mutagenesis is in purple, the GGG in stem II is in red and its base-pairing partner, CUCC, is in blue. Mutations in the affected sequence are depicted in the same color. Double mutants are underlined and their mutated sequences are linked with a dotted line. The predicted secondary structure and splicing of the mutants are summarized in Table 3. (B) The RT-PCR products are shown on a 6% polyacrylamide, 5M urea gel. (C) Graph shows the percentage of exon inclusion (%L) as calculated in Figure 3.

Table 3. Results of compensatory mutational analysis for the predicted secondary structure.

Splicing reporter/ mutation	Predicted effect on structure model [a]	Predicted change in exon inclusion [b]	Results of exon inclusion [c]
5m1	Disrupted (stem I)	Increased	Increased (1.4)
5m3	Disrupted (stem I & II)	Increased	Increased (3.1)
5m2	Disrupted (stem II)	Increased	Increased (3.7)
5m5	Restored	Decreased	Decreased (2.2)
5m8	Disrupted (stem II)	Increased	Decreased (0.2)
5m4	Disrupted (stem II)	Increased	Increased (3.9)
5m6	Restored	Decreased	Increased (3.9)
5m7	Disrupted (stem II)	Increased	Increased (2.6)
5m9	Disrupted (stem II)	Increased	Increased (3.3)
5m10	Restored	Decreased	Increased (3.4)
5m12	Enforced (stem II)	Decreased	Decreased (0.7)

[a] Disrupted, the stem structure is disrupted; Restored, the stem structure is restored by the second mutation; Enforced, the secondary structure is strengthened by increasing base pair in the stem.

[b] Increased, higher exon inclusion compared to wild type; Decreased, lower exon inclusion compared to wild type.

[c] Change in splicing pattern was determined by comparison of exon inclusion values of test and wild-type splicing reporters. Ratio of % exon inclusion (test)/% exon inclusion (wt) is given in brackets.

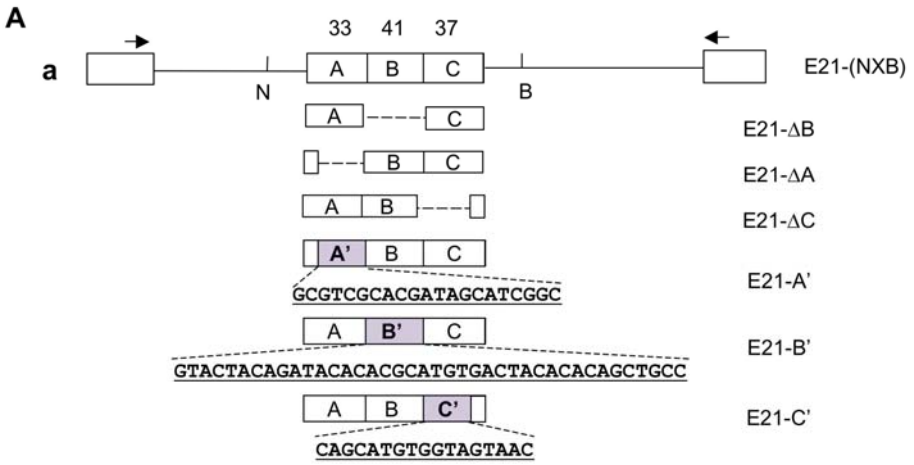
restoration of the stem structure. Mutation of 5m12 is predicted to enforce the stem II structure. The splicing pattern in 5m12 shows a decrease in exon inclusion as expected but the effect is not as significant as in 5m8. Therefore, the data of compensatory mutations do not strongly support the secondary structure model as a silencing mechanism of the CI cassette. However, mutational analysis can't disprove the involvement of RNA secondary structure.

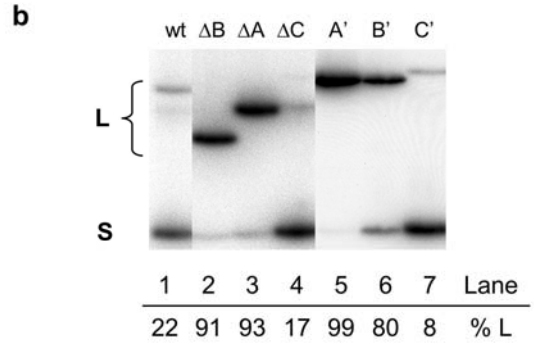
2.3.5. Exonic UAGG motifs are involved in CI cassette exon silencing.

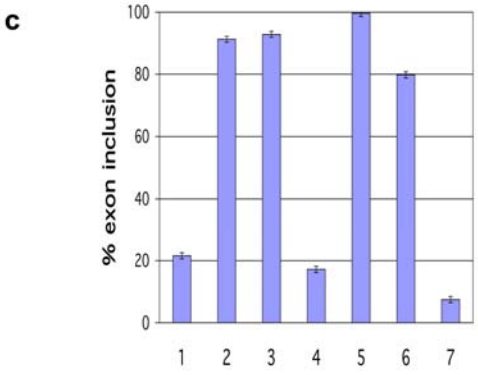
2.3.5.1. Exonic enhancers and silencers regulate CI cassette exon splicing.

To discover exonic regulatory elements, deletion and shuffled mutations were designed (Figure 11A-a, E21-ΔB, ΔA and ΔC). Deletion of regions A and B switches the splicing pattern from exon skipping to inclusion but deletion of region C increases more exon skipping (Figure 11A-b, lane 2-4). To confirm the effect of each region on CI cassette exon splicing, sequences of the region were mutagenized by shuffling and reinserted into that region (E21-A', B', and C'). The splicing patterns of the shuffled mutants correlate with the results of the deletion mutants. Therefore, the sequence of the exon itself influences CI cassette exon splicing, suggesting that it might contain repressor elements in regions A and B as well as enhancers in region C.

However, exonic splicing enhancer motifs generated in shuffled sequences A' and B' could enhance splicing. Region B contains a motif identical to the *tat* silencer, UAGG, which represses K-SAM exon splicing in HeLa cells (Del Gatto et al., 1996) and is


A
a
33 41 37
A B C
N
B
E21-(NXB)
A C
E21-ΔB
B C
E21-ΔA
A B
E21-ΔC
A' B C
E21-A'
GCGTCGCACGATAGCATCGGC
A B' C
E21-B'
GTACTACAGATACACACGCATGTGACTACACACAGCTGCC
A B C'
E21-C'
CAGCATGTGGTAGTAAC


b
wt ΔB ΔA ΔC A' B' C'
L
S
1 2 3 4 5 6 7 Lane
22 91 93 17 99 80 8 % L


c
% exon inclusion
100
80
60
40
20
0
1 2 3 4 5 6 7

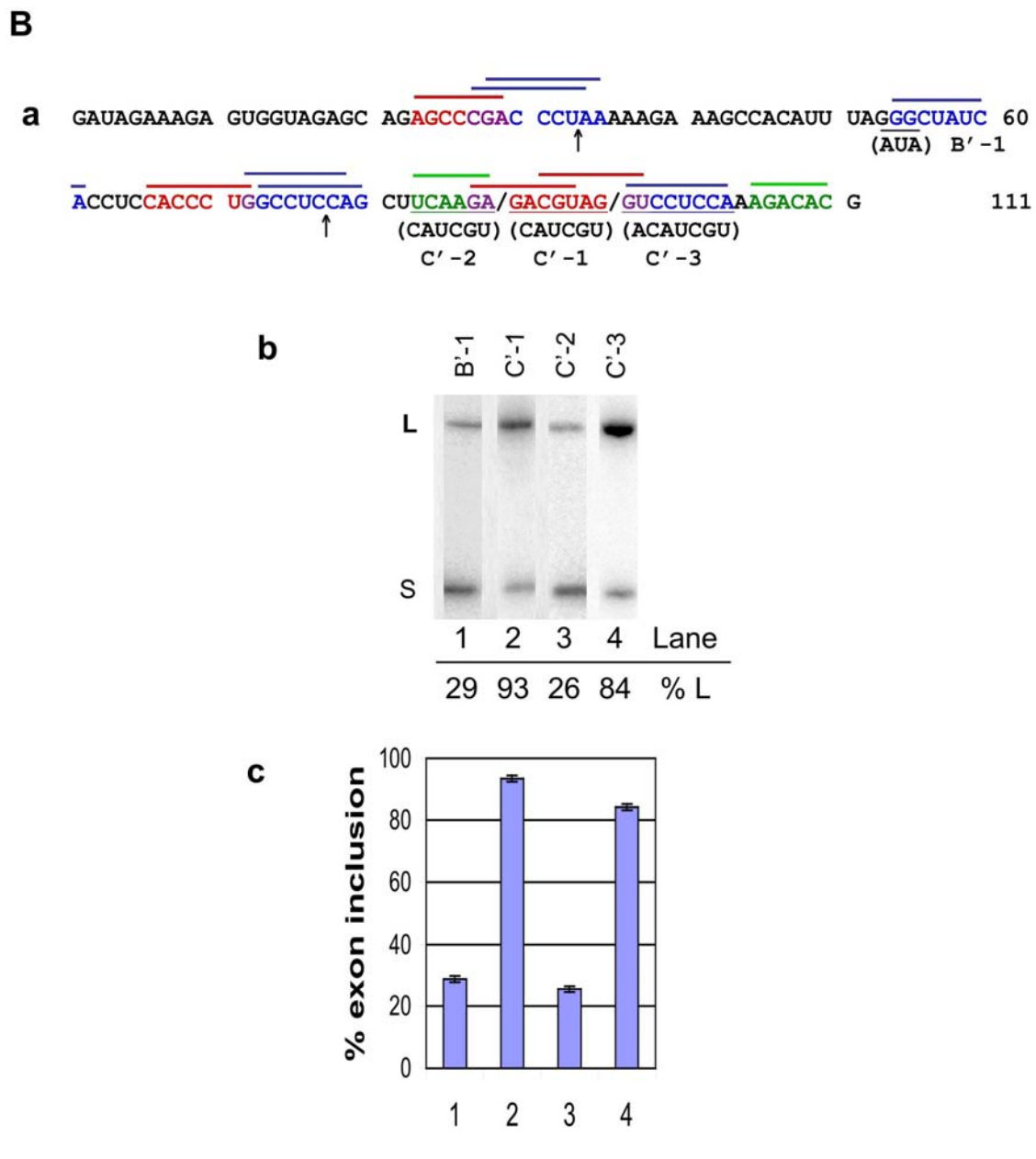

B
a
GAUAGAAAGA GUGGUAGAGC AGAGCCCGAC CCUAAAAAGA AAGCCACAUU UAGGGCUAUC 60
(AUA) B′-1
ACCUCCACCC UGGCCUCCAG CUUCAAGA/GACGUAG/GUCCUCCAAAGACAC G 111
(CAUCGU) (CAUCGU) (ACAUCGU)
C′-2 C′-1 C′-3
b
B'-1
C'-1
C'-2
C'-3
L
S
1 2 3 4 Lane
29 93 26 84 % L
c
% exon inclusion
100
80
60
40
20
0
1 2 3 4

Figure 11. Analysis of exonic sequences for CI cassette exon splicing.

A. Mutations in A and B regions induce CI exon inclusion but mutations in C region increase exon skipping. (a) All the deletion and shuffled mutations are derivatives from the wild type, E21-(NXB). The deleted regions, A (33 nt), B (41 nt), and C (37 nt), in the CI cassette exon are schematically shown. Shuffled sequences are shown below each construct as underlined (A', B' and C'). (b) The splicing patterns of the mutants are shown on a 6% polyacrylamide, 5M urea gel. (c) Graph shows the percentage of exon inclusion (%L) as calculated in Figure 3.

B. C region of the CI cassette exon contains both exonic enhancers and silencers. (a) The sequence of ESE motifs in the CI cassette exon sequence is represented as: SC35 motif, blue; SRp40, green; ASF/SF2, red. The same colored bars indicate the motifs overlapping one another. The replaced sequences are in parenthesis below the underlined wild type sequences. Arrows (↑) indicate the boundary between the regions (A, B and C). Mutations in the B and C region disrupt the corresponding ESE motifs. (b) RT-PCR products are shown on a 6% polyacrylamide, 5M urea gel. (c) Graph shows percentage of exon (%L) as calculated in Figure 3.

homologous to UUAG in Tau exon 10 in the mammalian nervous system (Gao et al., 2000). A sequence 5'-UUUAGGG-3' was mutated to the sequence, 5'-UUUAAUA-3' by site-directed mutagenesis (Figure 11B, B'-1) and transiently expressed in C2C12 cells. RT-PCR analysis shows the same exon-skipping splicing pattern as the wild type. Therefore, the sequence GGG in region B does not act as a splicing silencer. However, we cannot conclude that region B does not contain any splicing silencer since we have not tested all the sequences in that region. Here is a caveat of analysis for deletion mutations. In addition to telling us what region might contain regulatory sequences, it could also generate splicing silencers or enhancers at the junction of the deletion. Moreover, the shortened length of the exonic region could affect splicing regulation. Shuffled mutational analysis would be similar in that splicing regulatory sequences could be generated in between the boundary and in the shuffled sequence itself. Problems derived from the deletion and shuffled mutations call for standards, such as well-deifned regulatory sequences, to rule out those possibilities.

Since mutations in region C could disrupt ESE motifs, each ESE motif in this region was tested to determine which one contributes to CI cassette exon splicing. The online database ESEfinder (*http://exon.cshl.edu/ESE*) was used to identify putative ESE motifs of several splicing regulators, such as ASF/SF2, SC35, SRp40 and SRp55 and to evaluate the mutated sequences by their motif scores (Cartegni et al., 2003). In order to disrupt ASF/SF2, SRp40 or SC35 motif in region C, the hexanucleotide sequence CAUCGU was chosen and replaced (Figure 11Ba, C'-1, C'-2, and C'-3), expecting to decrease exon inclusion. In contrast to the results of deletion and shuffled mutational analysis, C'-1 and C'-3 show significant increases in exon inclusion (Figure 11Bb, lanes

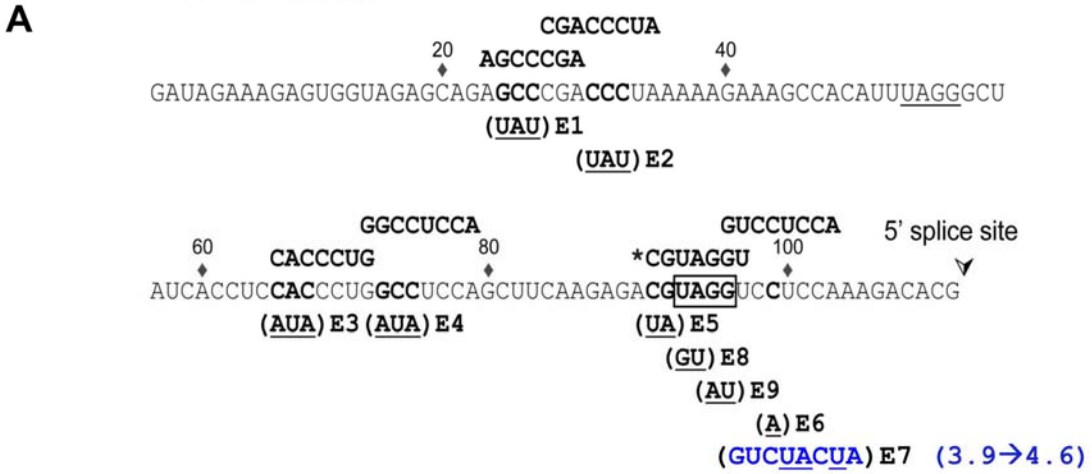
2 and 4) whereas C'-1 shows a slight decrease in exon inclusion as expected (lane 3). Therefore, within in region C, the ASF/SF2 and SC35 motifs, which are ESE motifs, should also contain a silencing activity. From this result, the C' mutant might be explained by silencing activity. The silencing activity could counterbalance one ESE motif but not all of them because the potential silencer motif was disrupted in the C' mutant in which all three ESE motifs were disrupted. Alternatively, mutations might generate a silencer motif or might not disrupt the potential silencer motif in that the C' mutant shows a greater decrease in exon inclusion than the C'-2 mutant. On the other hand, the result of the mutation in B'-1 should be interpreted with caution because the GGG sequence overlapping a SC35 motif could contain both the enhancing and silencing activities. Taken together, this suggests that a potential silencing motif as well as ESE motifs might be clustered in the C region. In addition, approaches using the ESE motifs would be advantageous to investigate silencing motifs in the A and B region. Thus, ESE motifs in the entire exon region will be tested by site-directed mutagenesis in the following section.

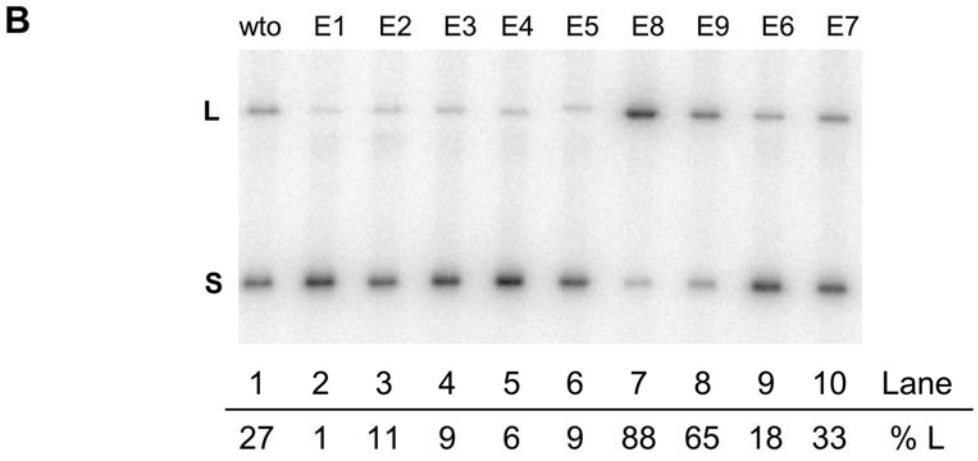
2.3.5.2. An exonic splicing silencer, UAGG, is required for CI cassette exon silencing.

Since the ESE motifs available online are degenerate and short (6-8 nucleotides), the probability of generating ESE motifs is predicted to be high. Given the length of the CI cassette exon (111 bp), 12~18 ESE motifs are expected to occur if they don't overlap each other in the exon. However, the frequency should be variable since regulatory elements are thought to be in the vicinity of the splice signals. In addition, the

occurrence of the ESE motif is predicted to be more frequent in the CI cassette exon than in constitutive exons because exons containing weak splice sites are expected to contain more ESE motifs to compensate for the weak signals (Fairbrother et al., 2002). Four ASF/SF2 motifs, six SC35 motifs, and two SRp40 motifs are found in the CI cassette exonic region (Figure 11B-a and 12A). Most of them overlap one another and are especially clustered in the region from the middle to the 5' splice site as expected.

Mutagenesis of exonic sequences in the CI cassette exon was performed in order to disrupt each ESE motif. Two rules were followed when designing mutagenesis experiments. First, repeated AU sequences were preferred as replacement sequences and contiguous purine sequences were avoided. Second, changed sequences were tested to examine whether or not any new ESE consensus sequences were generated from the mutations using the ESE finder database. Compared to wild type, most exonic mutants show a decrease in CI cassette exon inclusion, indicating that these mutations abolish the enhancing function of predicted ESE motifs (Figure 12B, lanes 2-6 and 9). By contrast, E8 and E9 show a dramatic increase in exon inclusion, indicating a silencing role for the UAGG sequence in agreement with the previous results (C'-1 and C'-3 in Figure 11B). Notably, the UAGG exonic splicing silencer (ESS) overlaps with two ASF/SF2 motifs in the CI cassette exon region, implying that enhancer and silencer proteins may compete for their binding to the same sequence. Thus, CI cassette exon splicing could be regulated by the ratio of those proteins in a specific tissue. Unlike E9, the mutation in E8 reduces the score of the ASF/SF2 motif rather than disrupting the motif, which might result in higher exon inclusion. Alternatively, if UAGG is a


A
CGACCCUA
20
AGCCCGA
40
GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGGCU
(UAU) E1
(UAU) E2
GGCCUCCA
GUCCUCCA
5' splice site
60
CACCCUG
80
*CGUAGGU
100
AUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG
(AUA) E3 (AUA) E4
(UA) E5
(GU) E8
(AU) E9
(A) E6
(GUCUACUA) E7 (3.9→4.6)


B
wto E1 E2 E3 E4 E5 E8 E9 E6 E7
L
S
1 2 3 4 5 6 7 8 9 10 Lane
27 1 11 9 6 9 88 65 18 33 % L

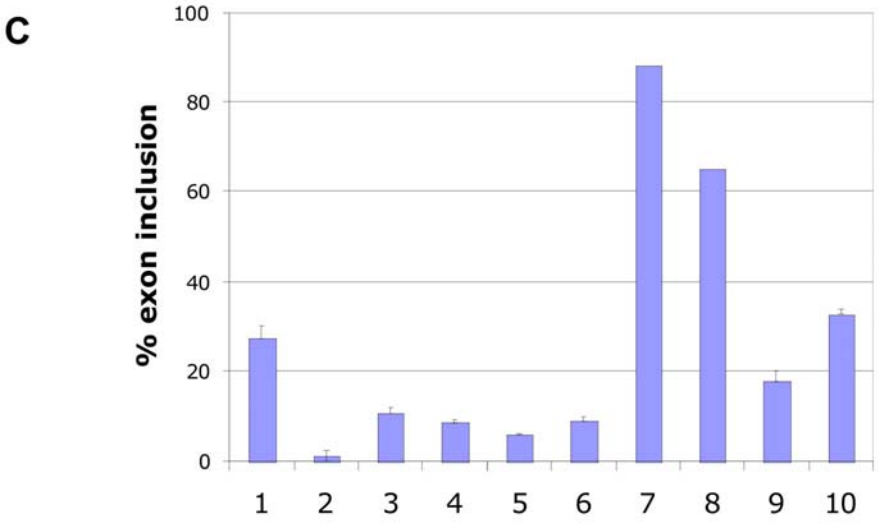

C
% exon inclusion
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10

Figure 12. Mutational analysis of UAGG motif involved in CI cassette splicing silencing.

(A) Sequences of wild type CI cassette exon and site-directed mutations for transient expression are shown. Predicted ESE motifs are represented above the exonic sequences. The mutagenized sequences in the wild type are bold and the sequences of mutations are underlined in parenthesis. Numbers indicate positions of nucleotides in the exon. An asterisk indicates the ASF/SF2 motif overlapping the UAGG motif that is boxed. All mutants were constructed by site-directed mutagenesis to disrupt ESE motifs except E7 and E8. The numbers in a parenthesis represent scores of a SC35 motif at position 96. (B) RT-PCR products are shown on a 6% polyacrylamide, 5M urea gel. (C) Graph shows the percentage of exon inclusion (%L) as calculated in Figure 3.

predominant regulatory element in the CI cassette exon, both the U and A nucleotides of UAGG could be more important for splicing repression than the two G nucleotides. These results indicate that the UAGG sequence is an exonic splicing silencer (boxed in Figure 12A). For a comparison, E7 was designed as a mutation to generate a stronger ESE binding motif for the SR protein, SC35, by increasing the motif score from 3.9 to 4.6. From this, we would expect to increase CI exon inclusion by enhancing the binding of the ESE motif to SC35. Exon inclusion slightly increases from 27 % to 33 %, supporting the results of mutational analysis for ESE motifs. Disruption of the SRp40 motif at position 80 (C'-2 in Figure 11) and at position 95 (5m1) shows no significant influence on CI cassette exon splicing. This result implies that the SRp40 motif is not primarily involved in regulation of CI cassette exon splicing, at least in C2C12 cells.

Individual ESE motifs at each position show somewhat different effects on CI cassette exon inclusion by mutational analysis. The mutated sequences in the exon are not only part of the ESE binding motifs but are also at unique positions from the splice sites. To analyze the function of each ESE motif, the effect of ESE mutations on the same motif at different positions in the exon was investigated. Except for E8 and E9 mutants, ESE mutants are categorized into two groups, an ASF/SF motif and a SC35 motif group. The ASF/SF2 motif group contains the mutants E1, E3 and E5 from the 5' end to 3' end of the exon. The mutated motifs correspond to positions 23, 66 and 89, respectively. E1 mutant (Figure 12B, lane 2), whose mutation is closest to the 3' splice site (motif at position 23), shows the most significant decrease in exon inclusion (1 %) among all the ESE mutants. E3 and E5 (lanes 4 and 6) show a similar decrease in exon inclusion (9 % for both), where the motifs are in the middle of the exon and close to the

5’ splice site. The ASF/SF2 mutation at position 23 close to the 3’ splice site shows the most significant effect on CI cassette exon inclusion, implying that this motif may play a crucial role as a splicing enhancer element by interacting with its positive regulator, ASF/SF2. In E5, the mutation disrupts the ASF/SF2 motif thereby allowing the UAGG silencer to be more accessible for a possible interaction with a negative regulator. On the other hand, the ASF/SF2 motif near the 3' splice site, disrupted in the E1 mutant, could not be involved in recognition of the 5' splice site. Instead, the mutation near the 3’ splice site might have an additive effect by inhibiting 3' splice site recognition, thereby conferring double repression of CI cassette exon splicing at both splice sites. Given the predominant exon skipping of the wild type C1 cassette exon and the significant increase in exon inclusion by mutations in the UAGG sequence, the UAGG silencer is dominant to any ESE motif in the CI cassette exon. Therefore, it is likely that the strongest negative regulation of CI cassette exon splicing in the E1 mutant results from repression of both the 5’ splice site recognition by the UAGG silencer and the 3’ splice site recognition by ESE mutation. The second motif group, SC35, contains the mutants E2 at position 27, E4 at 72, and E6 at 96 (lanes 3, 5, and 9). E2 mutant (lane 3), whose mutation is closest to the 3' splice site, shows a decrease in exon inclusion from 27 % to 11 %. Exon inclusion of the E4 mutant, which lies in the middle of the exon, is 4 % and E6, located close to the 5’ splice site, is 18 %. The strong decrease in exon inclusion seen in E4 could be caused by disruption of two overlapping SC35 motifs in the middle of the exon or due to the importance of the position in CI cassette exon.

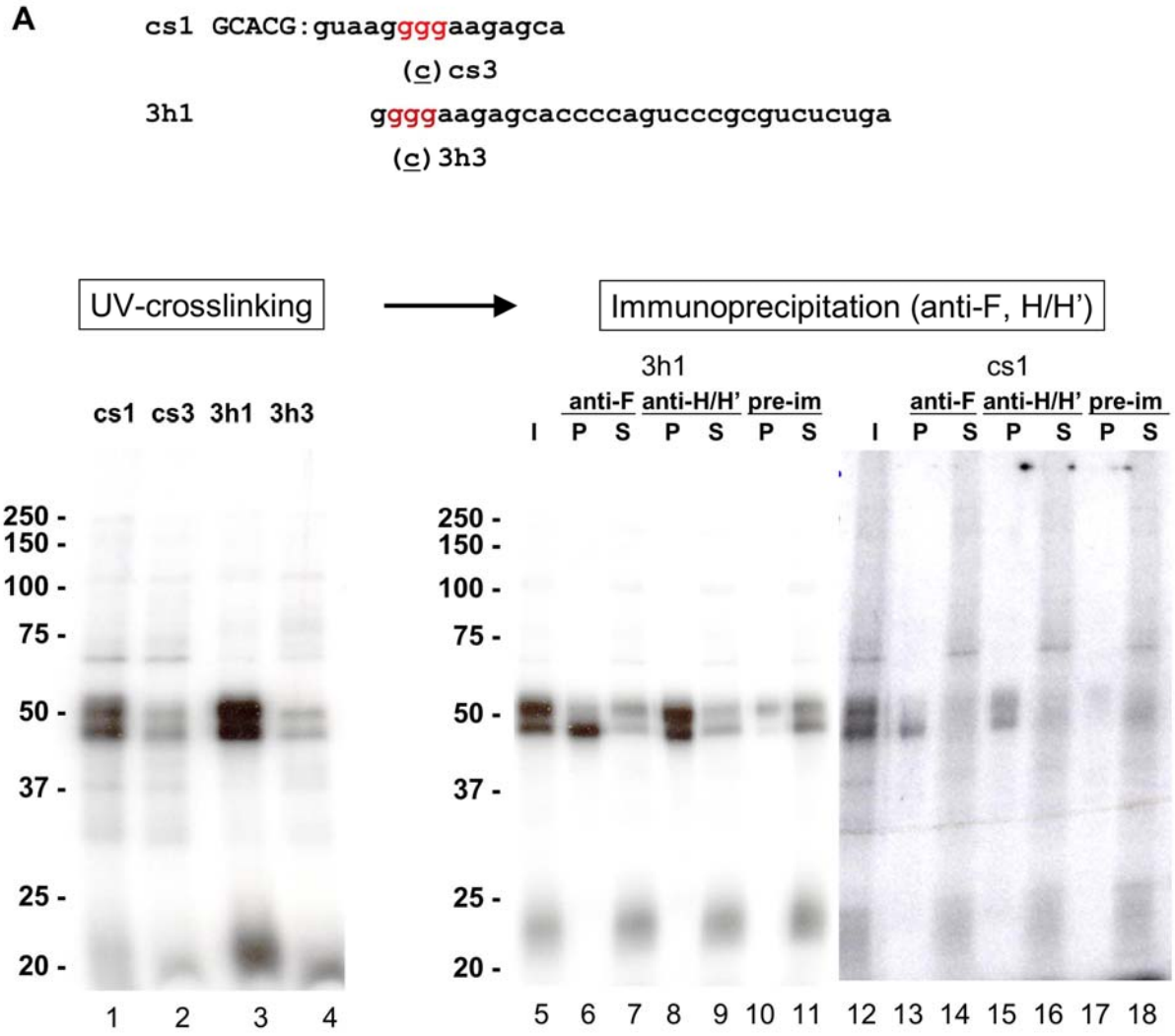
2.3.6. CI cassette exon splicing is regulated by hnRNP proteins.

The *in vivo* study of CI cassette exon splicing by mutational analysis enabled the identification of important splicing silencers, which are a UAGG motif in the exon and a G cluster in the vicinity of the 5' splice site. The next question is whether these two silencers regulate CI cassette splicing individually or in combination. Based on the combinatorial control for alternative splicing wherein regulation is achieved by a specific ratio of splicing regulators, approaches for the identification of RNA binding proteins interacting with the G cluster and the UAGG motif are necessary. In addition, the exonic UAGG sequence interacts with at least two SR proteins since it is embedded in two overlapped ASF/SF2 motifs and a SC35 motif. Given the fact that the splicing machinery is composed of snRNAs and protein complexes, the G cluster at the 5' splice site could be a binding motif for splicing regulators even though the secondary structure model would not be completely ruled out. Thus, *in vitro* UV-crosslinking and immunoprecipitation assays were undertaken to examine the molecular basis of CI cassette exon silencing. Most interactions are weak and transient in spliceosome assembly, thus a method to stabilize the interactions of proteins with a RNA substrate is required for identifying protein factors interacting with RNA sequence elements. UV-crosslinking is a useful tool to stabilize protein-RNA interactions, which induces covalent bonds between amino acid residues and bases of nucleotides in contact with one another. For this reason, this method has identified many protein factors interacting with an RNA substrate involved in splicing reactions. However, it is important to note that the efficiency of UV-crosslinking depends on interacting chemical groups; therefore, the detection of interactions can vary and some interactions might not be detected. In

addition, the nucleotide involved in RNA-protein interactions should be labeled using its specific isotope or by site-specific labeling.

2.3.6.1. Identification of hnRNP F and H/H' proteins involved in binding the 5' splice site G cluster.

In order to identify *trans*-acting factors recognizing the G cluster, guanosines were evenly labeled using α-[^{32}P]-GTP. Internally ^{32}P-labeled transcripts (Figure 13A, cs1) encompassing the 5' splice site of the CI cassette exon were incubated in HeLa nuclear extract under splicing assay conditions and irradiated with UV light. Following the UV irradiation, the samples were incubated with RNase A, resolved by SDS-PAGE and the crosslinked proteins visualized by autoradiography. A protein doublet around 50 kD is crosslinked to the cs1 RNA substrate (Figure 13A, lane 1). The binding specificity of this doublet was assessed by a G-to-C point mutation in the middle of the G cluster (cs3), which is equivalent to the mutation of 5m9 (lane 2). As cs3 does not crosslink these proteins as strongly as cs1, the binding of the G cluster to the two proteins appears to be sequence-specific. This result was confirmed using a different version of RNA substrate, 3h1, containing the G cluster and the downstream intron sequences (Figure 13A). As a comparison, the same point mutation in 3h3 as in cs3 was designed. Like cs1 and cs3, a protein doublet is seen crosslinking to the 3h1 RNA substrate but not to the 3h3 mutant substrate. From these results, it is clear that the G cluster interacts with two proteins at an apparent molecular weight of 50 kD. Unlike most other hnRNP proteins, which display a varying range of affinities for different ribonucleotide homo-


A
cs1 GCACG:guaaggggaagagca
(c) cs3
3h1 ggggaagagcaccccagucccgcgucucuga
(c) 3h3
UV-crosslinking
Immunoprecipitation (anti-F, H/H')
cs1 cs3 3h1 3h3
3h1
cs1
anti-F anti-H/H' pre-im
I P S P S P S
250 -
150 -
100 -
75 -
50 -
37 -
25 -
20 -
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

B

```
w1  AGAGACGUAGGUCCUCCAAAGACACG:guaa*ggggaagagcaccccagucccgcgucucuga
                                          (c)m1
w2      GACGUAGGUCCUCCAAAGACACG:guaaggg*gaagagcaccccagucccgcgucucuga
          (GU)m21                         (cc)m2
```

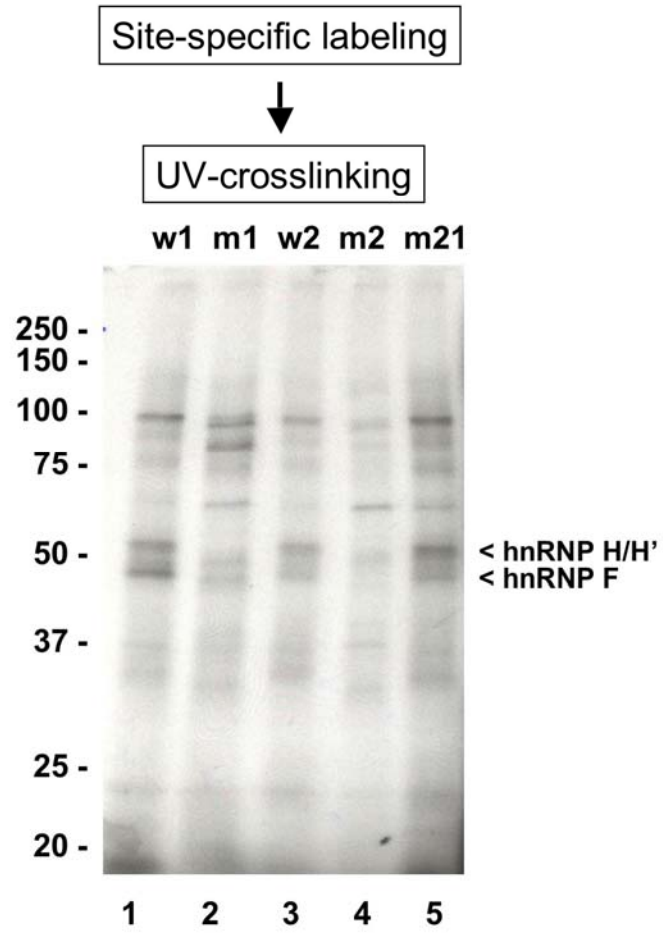


Figure 13. Identification of hnRNP F and H/H' involved in recognition of the G cluster.

(A) UV crosslinked proteins around 50 kD are immunoprecipitated with anti-hnRNP F and H/H' antibodies. In vitro transcripts of internally ^{32}P-labeled wild type (20nt, cs1,3h1) or point mutants (20nt, cs3, 3h3) of G cluster were incubated in HeLa cell nuclear extract and irradiated by UV. Sequences of the RNA substrates are shown on top. A flow chart of the experimental procedure is shown. Following UV irradiation and RNase A digestion, proteins are resolved on a 12.5% SDS-PAGE (lanes 1-4). The crosslinked samples were immunoprecipitated with polyclonal anti-hnRNP F or H/H' antibody (from Milcarek lab, Veraldi et al., 2001) adsorbed on agarose-A (lanes 5-13) and were resolved on a 12.5% SDS-PAGEs. Lane I, 20% of UV crosslinking sample; P, pellet; S, supernatant. Rabbit preimmune serum (pre-im) is a negative control for immunoprecipitation.

(B) Site-specific labeling of G cluster identifies hnRNP F and H/H'. Site-specific labeled RNA substrates contain both the exonic UAGG motif and the 5' splice site G cluster motif identified in the mutational analysis. Asterisks indicate the positions of labeling. The G residue at position 5 (w1, m1) or position 8 (w2, m2 and m21) in the intron was labeled using [γ-^{32}P]-ATP. Point (m1) or dinucleotide (m2) mutations were introduced to compare the crosslinking of the G cluster. The substrate m21 was designed to test the effect of the UAGG silencer on protein binding of the G cluster. The HeLa nuclear extract was used for UV crosslinking. The samples are resolved on a 12.5% SDS-PAGE. Arrowheads indicate the size of identified hnRNP F and H/H' proteins.

polymers, hnRNP F and hnRNP H bind specifically to poly(rG) in vitro (Matunis et al., 1994). Therefore, the apparent molecular weights of these proteins and the binding specificity for the guanosine-rich sequence suggest that the crosslinking proteins may be hnRNP F and H/H'. An immunoprecipitation assay was employed to identify crosslinking proteins using anti-hnRNP F or H/H' antibodies. The crosslinked samples for both the cs1 and 3h1 RNA substrates were incubated with antibodies immobilized on protein A sepharose resins. After washing away unbound proteins, the protein-antibody complexes were eluted from the resin, resolved by SDS-PAGE and visualized by autoradiography (Figure 13A). Two crosslinking protein bands are specifically precipitated with hnRNP F and hnRNP H/H antibodies compared to pre-immune serum, which identifies the upper band as hnRNP H/H' (lanes 8 and 15) and the lower band as hnRNP F (lanes 6 and 13). Wherease the hnRNP F antibody is highly specific, the hnRNP H/H' antibody crossreacts with lower bands at the same position as hnRNP F. This crossreactivity may occur because the antibodies used for immunoprecipitation are polyclonal.

Although evenly labeled RNA substrates show specific crosslinking to hnRNP F and H/H', the background levels still remain in the mutant substrates. To confirm that the G cluster specifically crosslinks those proteins, a site-specific labeling method was utilized for RNA substrates. Two RNA substrates (w1 and w2) were internally labeled on two different sites of the G cluster (Figure 13B). The point mutant (m1) and dinucleotide mutant (m2) RNA substrates were used as comparisons. A doublet around 50 kD is crosslinked to the two wild type substrates (Lanes 1 and 3) but not to the mutants (Lanes 2 and 4), indicating that hnRNP F and H/H' proteins specifically bind to the G

cluster. To examine the effect of the exonic UAGG motif on the interaction of the G cluster with hnRNP F and H/H', a mutation equivalent to E8 (m21) was employed for UV crosslinking. Both of the protein bands are still crosslinked to m21 (Lane 5), indicating that hnRNP F and H/H' bind to the G cluster in a sequence-specific manner. An additional band is detected at the apparent molecular weight of 65 kD, which is crosslinked to m1 and m2 (Lane 2 and 4). This band is not seen in the wild type RNA substrate lanes (lanes 1 and 3), suggesting that the crosslinked band at 65 kD is able to interact with the CI cassette exon when splicing silencing is destroyed by mutations in the UAGG motif. These results indicate that the G cluster interacts with hnRNP F and H/H' proteins.

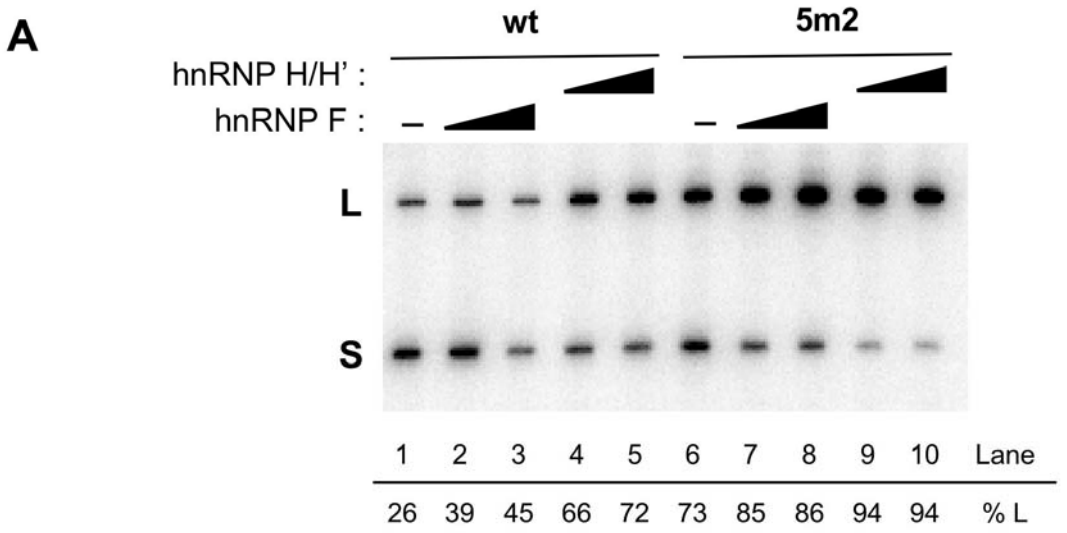
2.3.6.2. HnRNP F and H/H' play positive roles in CI cassette exon splicing.

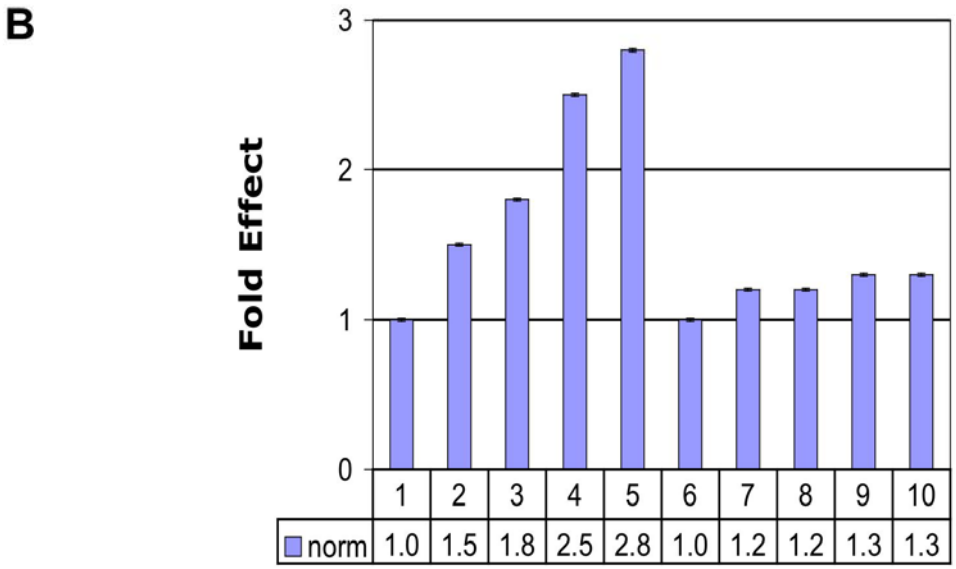
Even though the G cluster acts as a silencer in CI cassette exon splicing, hnRNP F is known to enhance splicing and hnRNP H seems to have a dual role depending on the interacting region (Min et al., 1995; Chou et al., 1999; Chen et al., 1999). In order to test the function of hnRNP F and H/H' proteins in CI cassette splicing, the wild type splicing reporter was co-transfected into C2C12 cells with overexpression plasmid pcDNA4-hnRNP F or pcDNA4-hnRNP H/H'. The 5m2 mutant (Figure 9) was used as a comparison. Transient overexpression of hnRNP F or H/H' induces CI cassette exon inclusion 1.8 and 2.8 fold respectively compared to the vector backbone, respectively (Figure 14A and B, lanes 3 and 5). The effect on exon inclusion after mutation of the G cluster (5m2) is at most 1.2 and 1.3 fold, respectively (Figure 14A and B, lanes 8 and

10). Even the splicing pattern of the 5m2 mutant without hnRNP F or H/H' proteins shows a high percent of exon inclusion (Figure 14A, lane 6). Thus, the G cluster at the 5' splice site is not thought to be the only binding sequence for hnRNP F or H/H' proteins nor to be only the factor enhancing CI cassette exon inclusion. This result indicates that hnRNP F and H/H' play enhancing roles in CI cassette exon splicing. However, it is not likely that the G cluster at the 5' splice site is required for the enhancing function of hnRNP F or H/H' due to an increase in exon inclusion of 5m2. The enhancing mechanism of the two splicing reporters would be different. An increase in exon inclusion of the wild type splicing reporter could be explained by U1 snRNP assembly onto the 5' splice site depending on hnRNP F or H/H', the same mechanism as suggested by Caputi and Zhaler (2000). Exon inclusion of the 5m2 mutant may also be enhanced by binding of hnRNP F or H/H' to G clusters further downstream in the intron. If the enhancing mechanism is the same for both splicing reporters, the G cluster at the 5' splice site should not be involved in splicing but the further downstream region may be involved. Therefore, silencing function of the 5' splice site G cluster is not identified in this section. Instead, these results suggest an enhancing role for hnRNP F and H/H' in CI cassette exon splicing.

2.3.6.3. Identification of hnRNP A1 involved in recognition of the exonic UAGG motif.

An exonic UAGG silencer resembles part of the A1 winner sequence as well as the consensus 3' splice site sequence. Since the consensus high-affinity hnRNP A1


A
wt
5m2
hnRNP H/H' :
hnRNP F :
L
S
1 2 3 4 5 6 7 8 9 10 Lane
26 39 45 66 72 73 85 86 94 94 % L


B
Fold Effect
3
2
1
0
1 2 3 4 5 6 7 8 9 10
norm 1.0 1.5 1.8 2.5 2.8 1.0 1.2 1.2 1.3 1.3

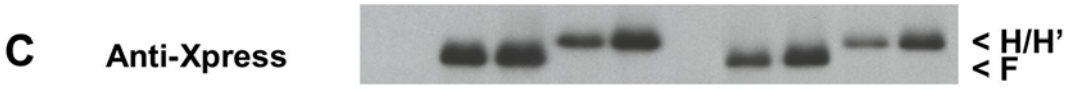

C
Anti-Xpress
< H/H'
< F

Figure 14. Effect of hnRNP F and H/H' overexpression on C1 exon splicing.

(A) The splicing reporter of wild type C1 cassette exon was cotransfected into C2C12 cells with pcDNA4-hnFNPF (cloned from pFlag-F, Chou et al., 1999) or pcDNA4-hnRNP H/H'(cloned from DSEF-1, Bagga et al., 1998). The CCC mutation on the G cluster of 5m2 mutant was tested as a comparison. The ratios of the splicing reporter plasmid to the expression plasmid are 1/4 (lanes 2, 4, 7 and 9) and 1/6 (lanes 3, 5, 8 and 10) each. The splicing products are shown on a 6% polyacrylamide, 5M urea gel and the percentage of exon inclusion is shown below. (B) Graph shows the fold effect of hnRNP F or H/H' expression on the pattern of splicing. Each value in the graph was calculated by dividing the percent exon inclusion (%L) of the splicing reporter co-transfected with hnRNP F or H/H' by the percent exon inclusion (%L) with vector backbone. (C) An immunoblot with anti-Xpress detecting a tag encoded by the expression vector is shown.

binding site, UAGGGA/U, was identified by the SELEX method (Burd and Dreyfuss, 1994b), its binding properties have been best characterized among many hnRNPs. The A1 winner more efficiently crosslinks hnRNP H family members (F and H/H') than hnRNP A1 (Caputi and Zahler, 2001), thus a site-specific labeling method was employed to label RNA substrates. A w3 transcript (Figure 15, 60nt) was site-specifically labeled at the first G of the sequence UA*GG and crosslinked in HeLa nuclear extract (Figure 15). A site-specific labeled w3 transcript crosslinks a 37kD band (Figure 15, lane 5), which is predicted to be hnRNP A1, including protein bands at approximately 65 kD and 50 kD. The A1 winner sequence is shown as a positive control for the hnRNP A1 binding motif (Lanes 1-3). An arrowhead indicates the position of hnRNP A1 in the gel panel. This crosslinked protein is immunoprecipitated with a monoclonal antibody (9H10) specific for hnRNP A1 but not with pre-immune serum, indicating that the labeled G of w3 binds hnRNP A1 (compare between lanes 5 and 7 in Figure 15). A UAGG to GUGG mutation corresponding to the exonic mutation of E8 (Figure 12) was used as a comparison (Figure 15, m3). The RNA substrate, m3, is crosslinked with the band in the vicinity of 37kD but is not precipitated with anti-hnRNP A1 antibody, revealing that the UV crosslinked protein bands with m3 at ~37kD are not hnRNP A1 (Figure 15, lanes 9-11). This result demonstrates that the exonic UAGG motif binds specifically to hnRNP A1 which is known to be involved in splicing repression.

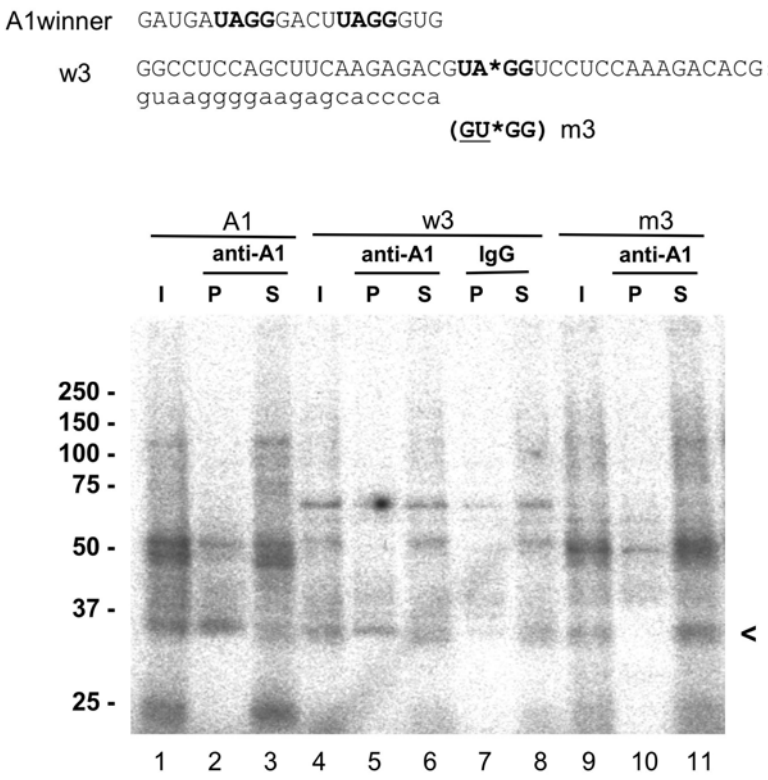
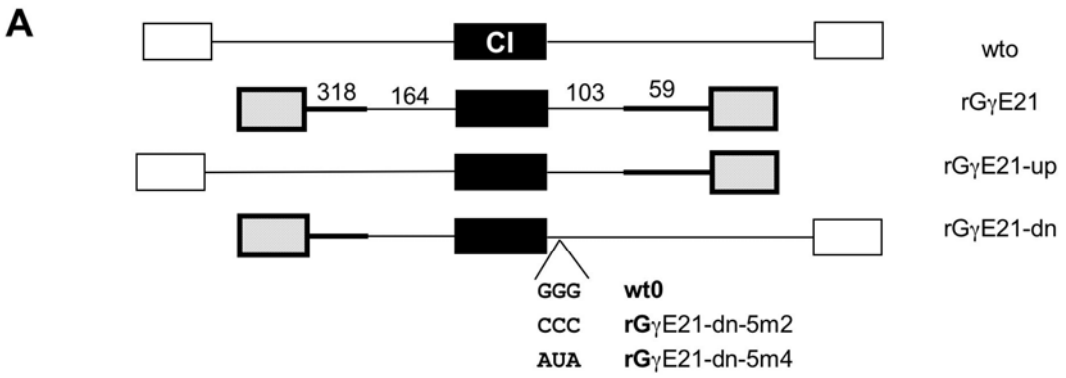


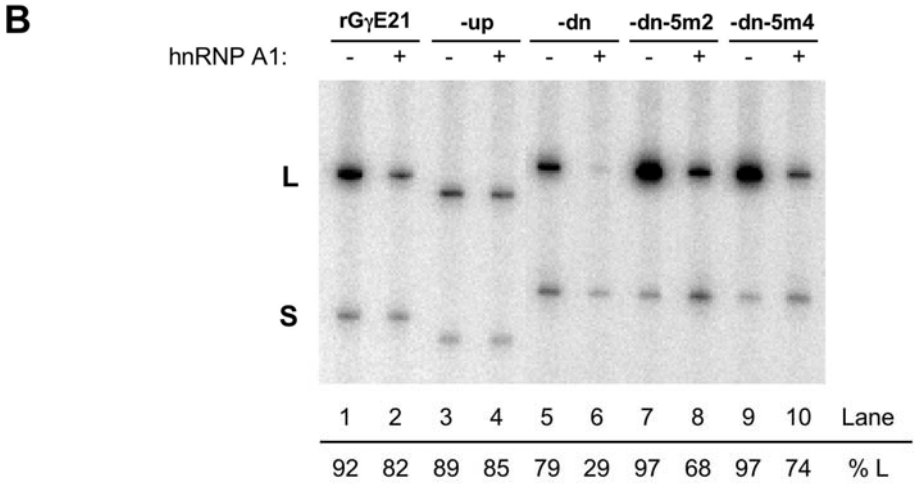
Figure 15. Identification of hnRNP A1 by immunoprecipitation following UV crosslinking.

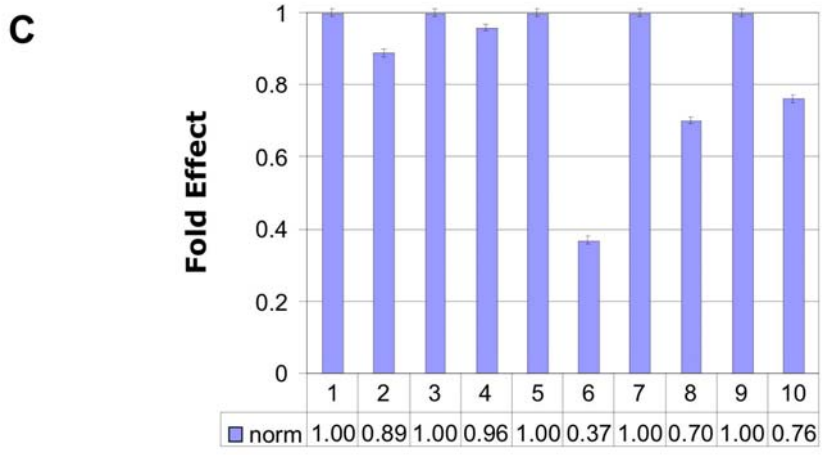
The first G residue of the UAGG silencer (w3 and m3) was labeled by a site-specific labeling method (asterisks). The sequences of the RNA substrates are shown on top. The mutation of m3 corresponds to the mutated sequence of E8. The internally ^{32}P-labeled A1winner sequence (Burd and Dreyfuss, 1994b) was used as a positive control for hnRNP A1 binding. UV crosslinking of RNA substrates was done in HeLa nuclear extract and the samples were precipitated with anti-hnRNP A1 antibody, 9H10 (from Dreyfuss lab). Mouse immunoglobulin G (IgG) was used as a negative control for immunoprecipitation. Abbreviations: I, 20 % input of crosslinking samples; P, pellet; S, supernatant. An arrowhead indicates the molecular size of hnRNP A1.

2.3.6.4. G cluster mediates CI cassette exon silencing by hnRNP A1.

From the results of the effect of PTB on CI cassette splicing, wild type CI cassette splicing reporter is expected not to be an appropriate substrate to test splicing silencing by overexpression. In addition, hnRNP A1 is ubiquitously expressed in cells (Su et al., 2002) and the wild type CI cassette exon splicing reporter is predominantly skipped in C2C12 cells. Given the result that two identified silencer motifs interact with hnRNP proteins that are abundantly expressed, overexpression of hnRNP A1 would not significantly affect CI cassette exon splicing. Thus, the chimeric construct rGγE21 that was utilized for overexpression experiments of PTB (Figure 4) and that contains both the UAGG and G cluster motifs was chosen to test the effect of hnRNP A1 on CI cassette exon splicing. However, chimeric splicing reporters rGγE21 shows a slight decrease in exon inclusion (Lanes 1 and 2). From the previous study, repressor elements in the downstream intron of the CI cassette exon could be required for the silencing function of hnRNP A1 (Zhang et al., 2002). The derivatives of rGγE21, rGγE21-up and rGγE21-dn, were generated by replacing the upstream or downstream intron of γ2 containing its flanking exon with the corresponding regions of the CI cassette exon (Figure 16A). As expected, rGγE21-dn shows a significant decrease in exon inclusion by shifting the major splicing pattern from inclusion to skipping (Figure 16B, lanes 5 and 6). By contrast, rGγE21-up shows a slight decrease in exon inclusion


A
CI
wto
318
164
103
59
rGγE21
rGγE21-up
rGγE21-dn
GGG
wt0
CCC
rGγE21-dn-5m2
AUA
rGγE21-dn-5m4


B
rGγE21
-up
-dn
-dn-5m2
-dn-5m4
hnRNP A1:
- + - + - + - + - +
L
S
1 2 3 4 5 6 7 8 9 10 Lane
92 82 89 85 79 29 97 68 97 74 % L


C
Fold Effect
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10
norm 1.00 0.89 1.00 0.96 1.00 0.37 1.00 0.70 1.00 0.76


D
anti-Xpress

Figure 16. Strong silencing effect of hnRNP A1 on C1 cassette exon splicing via G cluster.

(A) Effect of hnRNP A1 on C1 cassette exon splicing in the context of a chimeric splicing reporter. The chimeric constructs, rGγE21, rGγE21-up and rGγE21-dn, were generated by substituting the flanking region or all of the upstream and/or downstream intron(s) of the γ2 mini-gene construct, rGγ25. Thick lines represent flanking exon and the adjacent intron regions of the γ2 gene and thin lines represent the CI cassette exon, its flanking introns and exons. Intron fragment lengths in the chimeric constructs are marked above the corresponding intron regions. Mutated sequences of 5m2 and 5m4 are shown below the wild type sequence. Each splicing reporter plasmid was cotransfected with pcDNA4-hnRNP A1 at a ratio of 1/2. (B) RT-PCR products are shown on a 6% polyacrylamide, 5M urea gel. The percent exon inclusion is shown below. (C) Graph shows the fold effect of hnRNP A1 on splicing. Each value in graph was calculated by dividing the percent exon inclusion (%L) of the splicing reporter co-transfected with hnRNP A1 by the percent exon inclusion (%L) with vector backbone. (D) An immunoblot with anti-Xpress detecting the expression of transfected pcDNA4-hnRNP A1 is shown.

but the splicing patterns are primarily exon inclusion (Lanes 3 and 4). The G cluster also could be correlated with the hnRNP A1 function based on the result of UV crosslinking of m21 (Figure 13B) and the A1 winner sequence that contains a duplicated GGG sequence. Thus, mutations in G cluster were employed to test whether or not the G cluster is involved in the silencing function of hnRNP A1. Mutations in the G cluster (rGγE21-dn-5m2 and -5m4) in the chimeric construct fail to induce a shift in splicing pattern (lanes 7-10), suggesting that the G cluster acts as a silencer functioning with hnRNP A1. From these results, the intronic G cluster at the 5' splice site, the exonic UAGG and unknown silencer motifs in the further downstream intron are required for silencing of hnRNP A1.

2.3.7. The number of UAGG motifs is important for strong silencing of the CI cassette exon.

The UAGG motif in region B of the CI cassette exon has not been proved as a splicing silencer although the sequence of the motif is identical to the silencing UAGG motif identified in region C. The GGG sequence in region B overlaps with a SC35 motif as well as the UAGG motif, and the GGG to AUA mutation disrupts both the motifs. The disruption of the SC35 motif in the B'-1 mutant might compensate for destruction of the silencing function. Thus, the silencing function of the UAGG motif in region B was tested by a mutation of UA to AU that does not disrupt the SC35 motif. This mutaion results in a significant increase in exon inclusion (Figure 17A, E13). Therefore, these results show that the position of the UAGG motif in region B is a splicing silencer.

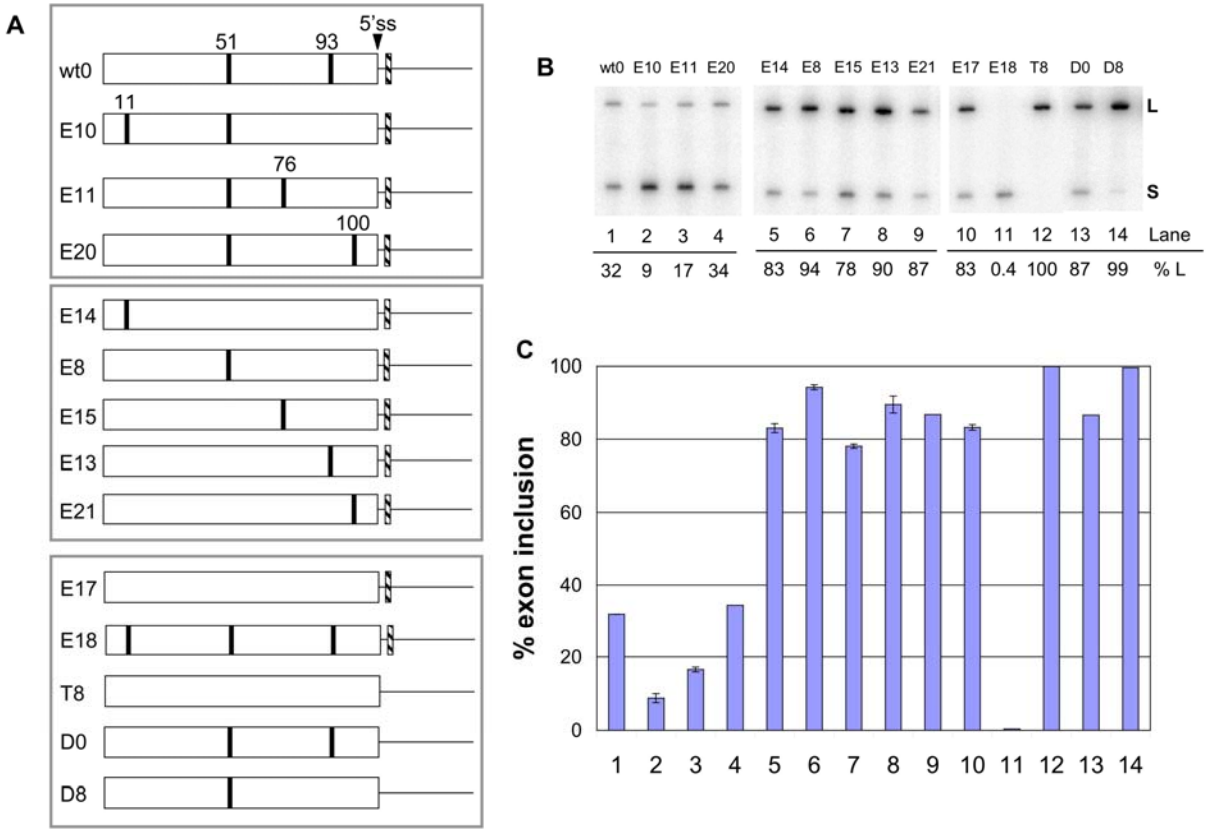


Figure 17. Effect of the number and position of silencers on C1 cassette exon silencing.

(A) The exonic UAGG silencers are shown as black bars and the 5' splice site G cluster is shown as a hatched bar in the schematic. Numbers indicate the positions of the UAGG silencers in the CI cassette exon. Constructs are grouped in boxes according to the number of the silencers. The first group contains two UAGG motifs and the G cluster (E10, E11 and E20), the second group contains one UAGG motif and the G cluster (E14, E15, E13 ad E21) and the third group contains various number of silencers, such as zero (T8), one (E17, D8), two (D0) and four (E18) silencers in combination. (B) RT-PCR products are shown on a 6% polyacrylamide, 5M urea gel. The percent exon inclusion is shown below. (C) Graph shows the percentage of exon inclusion (%L) as calculated in Figure 3.

Here, several questions for the role of the UAGG motifs in CI cassette exon silencing are addressed. Does the UAGG motif contain all the sequences required for silencing? If so, is it sufficient for functional silencing if moved to another position of the exon, like the G cluster? Otherwise, is its position close to the 5' splice site or its overlap with ESE motifs important for silencing? To answer these questions, the UAGG sequence was introduced into the exon at various positions avoiding disruption of known ESE motifs and the splicing patterns compared (Figure 17). When the UAGG motif at position 93 is moved to various positions from the 5' splice site to the 3' splice site (E20, wto, E11 and E10), the silencing function is strong (Figure 17B, lanes 1-4). In addition, the strong silencing function is not destroyed when it is closer to the 5' splice site (E20), which shows primarily exon skipping. Silencing is significantly reduced when either of the two UAGG motifs is removed (E14, E8, E15, E13 and E21, lanes 5-9 in Figure 17B). There is no ESE motif at position 11, where a new UAGG motif was generated by site-directed mutagenesis. Therefore, silencing is unaffected whether or not the UAGG silencer overlaps with ESE motifs. As a result, the silencing function of the UAGG motif is number-dependent, but not position-dependent.

To examine how silencer motifs function in silencing, the three silencers were deleted or added in various combinations to the wild type splicing reporter. All three sites are shown to be required for silencing (Figures 7 and 10). This is confirmed from the result that removing one of them causes de-repression of CI cassette exon silencing (E8, E13 and D0, lanes 6, 8 and 13 in Figure 17B). Although an individual silencer still plays a silencing function, the ibdividual contribution is not significant (E17 and D8, lanes 10 and 14). Strong silencing is obtained only when the three silencers coexist

(wto, E10, E11 and E20). Thus, it is not surprising that CI cassette exon splicing is almost completely silenced when an additional UAGG motif is inserted at position 11 of the exon (E18, lane 11). In the same context, a complete loss of silencing is observed when all three silencers are removed (T8, lane 12), which shows how this alternative exon can be inverted to a constitutive exon by knocking out silencer motifs. Although both UAGG motif and the G cluster possess silencing functions, two UAGG motifs and one G cluster are together required for efficient silencing, suggesting that they function cooperatively.

2.3.8. Search for exons containing the GGGG (G cluster) and UAGG motifs is predictive for alternative splicing in the human and mouse genome

This study identifies two silencers of the CI cassette exon, the exonic UAGG motifs and the 5' splice site G cluster, which are required for efficient splicing silencing. Our collaborators, Gene Yeo and Chris Burge, used the human and mouse genome sequence databases to conduct a bioinformatics search for $\geq$1 UAGG motif in the exon and a GGGG (G cluster) motif in the first 10 nucleotides of the intron. The frequency of a UAGG motif occurring in an exon is higher than a GGGG motif in a 5' splice site. 213 human exons and 200 mouse exons contain both motifs together. Only 19 exons are found in the intersection of the two, which contains the UAGG and GGGG motifs in both human and mouse (Table 4), revealing that the occurrence of the two silencers simultaneously is infrequent. Around 98% of the 213 human exons are coding exons

Table 4. Human and mouse exons containing an exonic UAGG and a 5'splice site GGGG motifs. [a]

Ensembl ID_Exon#: Human (ENSG00000-) Mouse (ENSMUSG00000-)	HUGO ID or Genbank accession #	Exon length (bp)	5' splice site sequence: Human Mouse	cDNA and/or EST evidence for exon skipping [b]	Disease
158195_4 028868_4	*WASF2*	118	AGGgtga**gggg**aa AGGgtga**gggg**aa	No info	Wiscott-Aldrich syndrome
169045_5 007850_5	*HNRPH1*	139	CAGgt**gggg**atgg CAGgt**gggg**atgg	AW579178	
096746_3 020069_2	*HNRPH3* *NM_012207*	139	CAGgt**gggg**atgg CAGgt**gggg**atgg	BE747312	
176884_19 026959_19	*GRIN1* [c]	111	ACGgtaa**gggggg**a ACGgtaa**gggg**aa	See below	
136044_8 020263_8	*NM_018171* *(DIP13BETA)*	147	CAGgta**gggg**agt CAGgta**gggg**atg	No info	
158865_8 030769_9	*NM_052944* *(KST1)*	81	ACAgtaagt**gggg** ACAgtaagt**gggg**	No info	
168453_3 022096_3	*HR* [d]	793	AAGgtaa**gggggg**c GAGgtaa**gggggg**t	BX341278	
068400_2 031153_13	*GRIPAP1*	96	AAGgta**gggg**aac AAGgt**gggg**catc	No info	
136478_8 040548_8	*NM_018469* *NA*	133	AGGgtaa**gggg**ct AGGgtaa**gggg**ct	No info	
108592_17 020706_18	*FTSJ3*	100	CCGgtaaa**gggg**c CCGgtaa**gggg**ca	No info	
152818_5 019820_6	*UTRN*	93	CAGgt**gggg**aaat CAGgt**gggg**acct	No info	
158887_4 005678_4	*MPZ* *NA*	136	CAGgtaa**gggg**cg CAGgtaa**gggg**cg	No info	Charcot-Marie-Tooth neuropathy
181045_5 039908_6	*SCL26A11*	143	CAGgtga**gggg**cc CAGgtga**gggg**ac	No info	
147255_16 031111_15	*IGSF1*	288	CAGgtaa**gggg**aa CTGgtaa**gggg**at	No info	
173957_7 037336_6	*NA* *NA*	91	CAGgtat**gggg**tt CAGgtat**gggggg**t	No info	
106404_2 001739_2	*CLDN15*	165	CCGgtaact**gggg** CTGgtaat**gggggg**	BU164601	
150165_4 021950_5	*ANXA8* *NM_001630*	91	AAGgtaa**gggg**tg AAGgtaa**gggg**tt	BC008813	
179593_2 020891_2	*ALOX15B* *NM_001141*	220	CAGgtga**gggg**cg CAGgtga**gggg**ac	No info	
165816_11 025082_12	*NA* *NA*	552	GAGgtgagt**gggg** TGAgt**gggg**ataa	No info	

[a] The search for exons has been done in collaboration with Gene Yeo in Burge lab at MIT.

[b] Here one representative example is shown.

(*http://genome.ucsc.edu*; *http://www.ncbi.nlm.nih.gov/*; *http://www.ebi.ac.uk/embl/*);

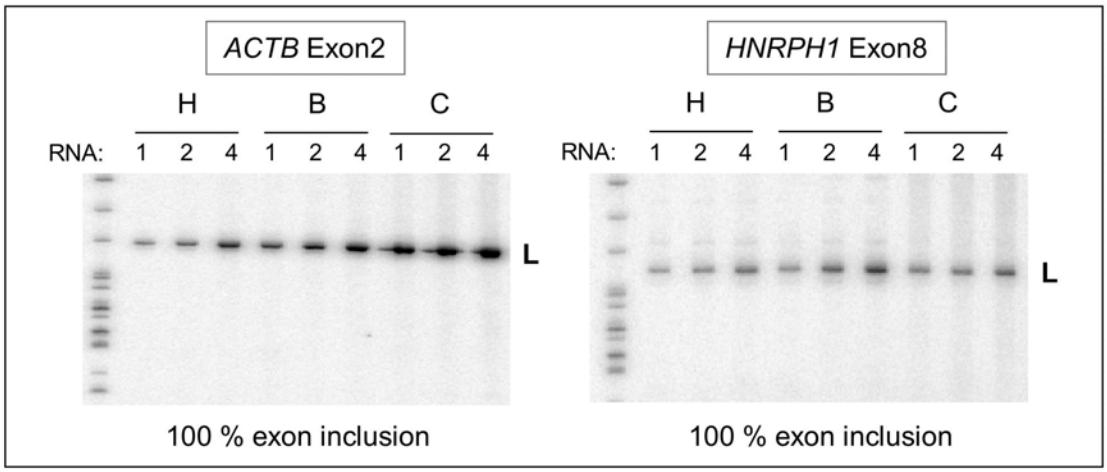
[c] Two UAGG; [d] Three UAGG

and a majority of these are similar to the mean length of human exons (155 nt, Lander et al., 2001), suggesting that they are functional exons encoding proteins, thus generating proteomic diversity. Seven out of twelve exons tested in humans are validated as alternatively spliced cassette exons in a tissue specific manner by RT-PCR analysis (data not shown, in collaboration with G. Yeo), hence the predictive value is greater than 50%. This result satisfies the prediction that the pair of the UAGG and GGGG motifs are silencer motifs in exons of the different gene systems.

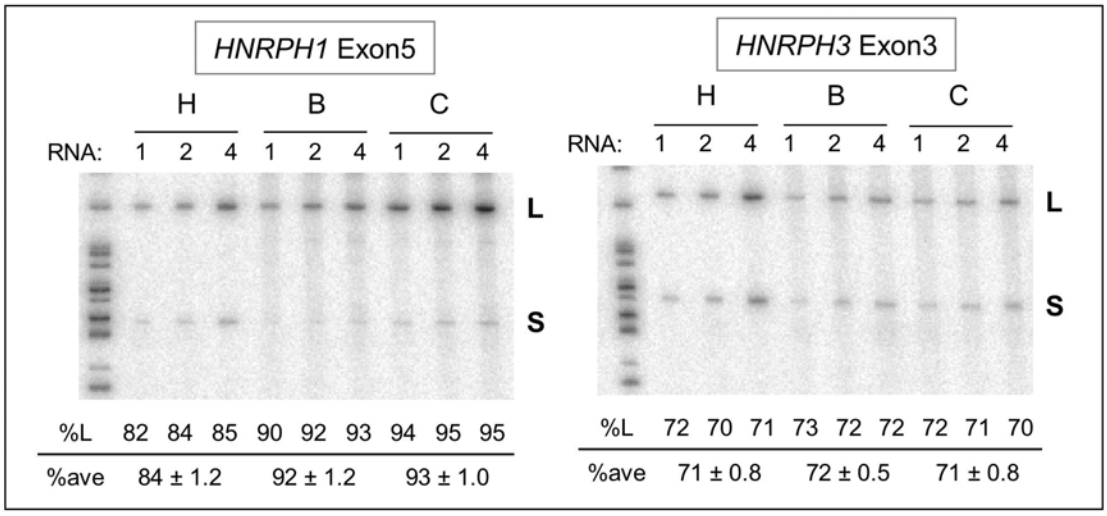
2.3.8.1. Identification of alternative exons of the splicing factor, hnRNP H and 2H9, containing one UAGG and a GGGG motif.

From the search for exons containing one UAGG and a GGGG motif, hnRNP H and 2H9, which are members of the hnRNP H family (HUDO ID, *HNRPH1* and *HNRPH3*), were found (Table 4, in collaboration with G. Yeo). The exon lengths of both *HNRPH1* exon 5 and *HNRPH3* exon 2 are short (139 nt each), which agrees well with the exon definition model (Berget, 1995). The splicing patterns of *HNRPH1* exon 5 and *HNRPH3* exon 2 were examined in mouse heart (H) and brain (B) tissues and C2C12 cells (C) by RT-PCR. For precise quantification, a series of the amount of RNA was utilized for RT reactions, such as 0.1μg, 0.2μg and 0.4μg of polyA$^+$ RNA, and 1μg, 2μg and 4μg of total RNA on the assumption that the amount of polyA$^+$ RNA is approximately equivalent to 10 % of total RNA (Figure 18B, lanes 1, 2 and 4). Exon skipping may not be as prominent as for the CI cassette exon because both have one

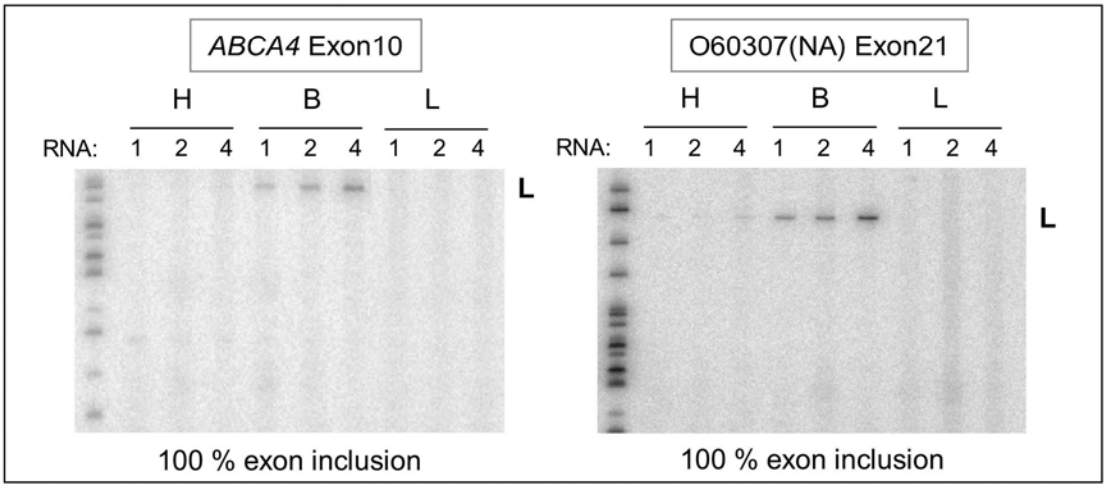
A **0 TAGG, 0 GGGG exons:**



B **1 TAGG + GGGG exons:**



C **2TAGG + GGGG exons, human:**



D **2TAGG + GGGG exons, mouse:**

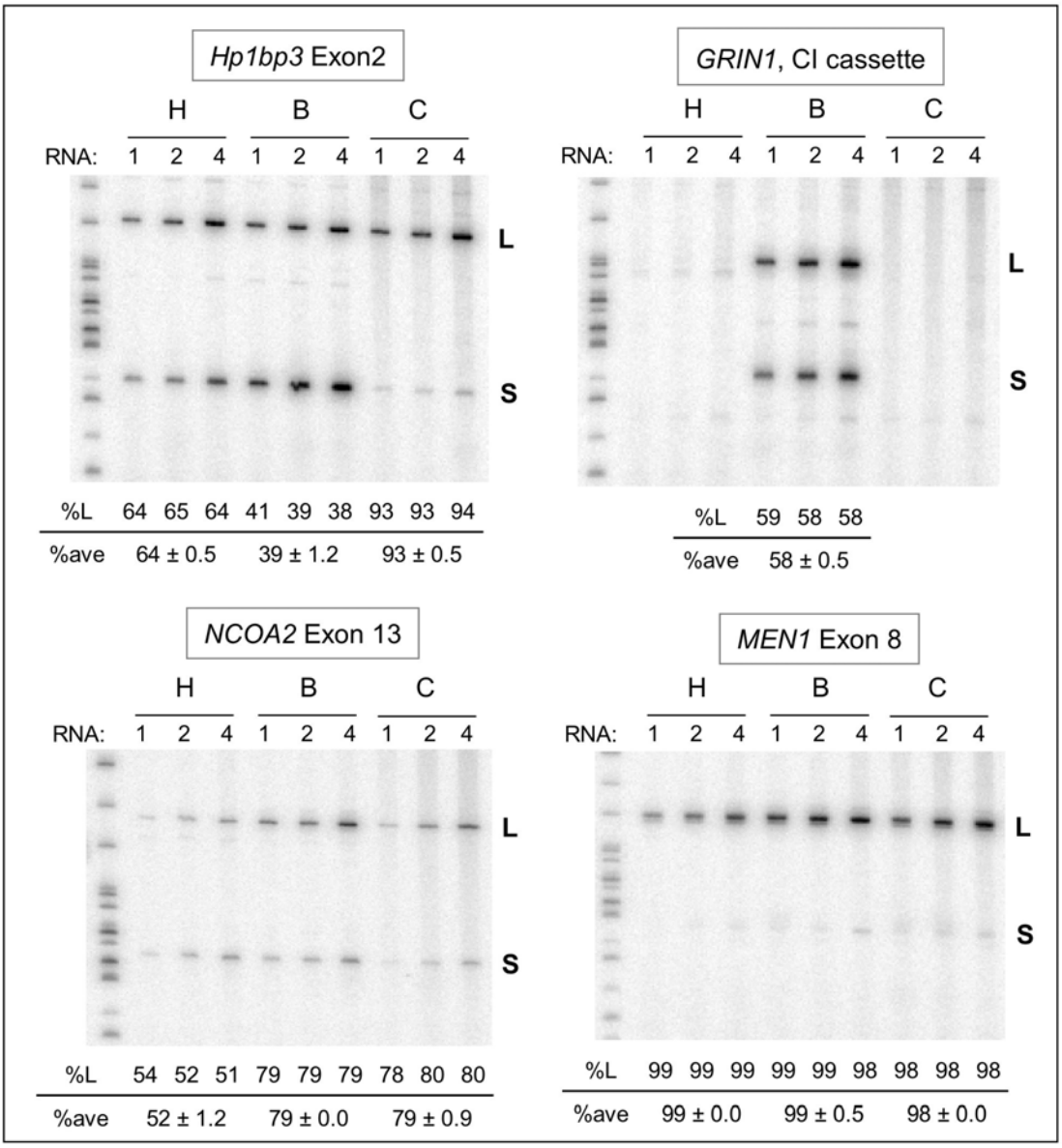


Figure 18. RT-PCR analysis of exons containing UAGG and GGGG motifs.

Exons were chosen from the search (Table 4-6, done by G Yeo and Burge, C) and their splicing patterns were tested by RT-PCR. (A) β-actin (*ACTB*) exon 2 and hnRNP H (*HNRPH1*) exon 8 are shown as control exons that have no UAGG and GGGG motifs. The splicing patterns of exons with one UAGG and one GGGG motifs (B), two UAGG and one GGGG motifs (C for human and D for mouse) are presented in separate boxes. Exon included products (L) and exon skipped products (S) are shown on a 6% polyacrylamide, 5M urea gel. The percent exon inclusion (%L) and the average (% ave) in each tissue are shown below the gel. For panel A, B and D, 22-24 cycles of RT-PCR were performed using 0.1, 0.2 and 0.4 μg (RNA: 1, 2, and 4) of poly$^+$ RNA from mouse heart (H) and brain (B), and 1, 2 and 4μg (RNA: 1, 2, and 4) of total RNA from C2C12 (C) cells. 30 and 24 cycles of RT-PCR were done for *ABCA4* exon 10 and *NA* (unknown gene, O60307) exon 21 in panel C, respectively, using 0.1, 0.2 and 0.4 μg of poly$^+$ RNA from human heart (H) and brain (B), and 1, 2 and 4μg of total RNA from HeLa (L) cells. PCR primers were designed using Primer3 software (*http://frodo.wi.mit.edu/*) and are listed in Appendix A.

UAGG in the exon and a G cluster at the 5' splice site. As expected, RT-PCR results show slight variations in both tested tissues and a cell line where the splicing pattern is predominantly exon inclusion (Figure18B). Exon controls, *ACTB* exon 2 (β-actin) and *HNRPH1* exon 8, which do not contain a pair of the UAGG and GGGG motifs are not skipped (Figure 18A). In this study, hnRNP F and H/H' are shown to function as positive regulators involved in CI cassette exon inclusion, suggesting that the *HNRPH1* exon 5 is regulated through the same silencing mechanism as the CI cassette exon. In addition, an exon 5-skipped transcript of *HNRPH1* produces a premature stop codon, implying that the protein level of hnRNP H is autoregulated by positive feedback through the GGGG motif.

2.3.8.2. Identification of distinct alternative exons containing multiple UAGGs and the G cluster motif.

Using the same dataset, cassette exons containing two or moreUAGG motifs in the exon and a GGGG motif at the 5' splice site (Table 5 and 5, human and mouse, respectively, in collaboration with G. Yeo). The frequency of occurrence is reduced compared to the search for one UAGG in the exon and a single GGGG motif at the 5' splice site. Compared to 19 exons found in the previous search, only *GRIN1* (HUGO ID) CI cassette exon was found in the intersection of human and mouse exons (compare exons between Table 5 and 6). *ABCA4* exon 10, *NCOA2* exon 12 and O60307 (NA) exon 21 (Genebank ID; Ensembl number, ENSG00000099308) were chosen from the list of human exons in Table 5, whose exon lengths are short in pre-mRNA (less than 250 nt), and tested whether the exons are skipped using RT-PCR. As seen in Figure

18C, *ABCA4* exon 10 and O60307 (NA) exon 21 are not skipped in contrast to the prediction. Interestingly, the two exons are specifically expressed in the human brain and the levels are relatively low. *ABCA4* gene encodes a retinal-specific ATP-binding cassette transporter (ABCR) and its mutations have been reported to cause Stargardt disease (Allikmets et al., 1997), This gene is expressed specifically in the retina, which is not an accessible tissue in human subjects. *NCOA2* exon 13 shows exon skipping in both human (data not shown) and mouse and is further described below.

Table 5. Human cassette exons containing a GGGG motif and two or more UAGG motifs. [a]

Ensembl# _Exon# (ENSG00000-)	HUGO, Genbank, or SWISSPROT ID #	Exon length (bp)	5' splice site sequence	RT-PCR evidence for exon skipping	cDNA and/or EST evidence for exon skipping [b]	Disease alleles
176884_19	*GRIN1*	111	ACGgtaa**ggggg**a	Skipped	L13266	
097054_10	*ABCA4*	117	AGAgtaa**gggggg**	negative	No info	Stargardt's disease
140396_13	*NCOA2*	207	CAGgtaa**gggg**tc	Skipped	No info	
099308_21	*O60307*	245	CTGgtaagt**gggg**	negative	No info	
135709_2	*Y513_HUMAN*	501	AGGgtaa**gggg**cc	ND [d]	No info	
165816_11	*NA*	552	GAGgtgagt**gggg**	ND	No info	
130283_7	*LASS1*	637	GCGgtgagt**gggg**	ND	No info	
007565_3	*DAXX*	832	CAGgta**ggggg**tt	ND	- [e]	
185133_2	*PIB5PA*	1166	CCGgtga**ggggg**c	ND	-	
111077_18	*TENC1* [c]	1212	CAGgtga**gggg**ca	ND	-	
142102_4	*Q8TEG9*	1418	CAGgtga**gggg**ac	ND	-	
135835_5	*Q9HCF8*	1556	ATGgtaa**gggg**ct	ND	-	
138080_4	*EMILIN1*	1929	CTGgtga**gggg**ac	ND	-	

[a] The search has been done in collaboration with Gene Yeo in Burge lab at MIT.

[b] Here one representative example is shown. (*http://genome.ucsc.edu*; *http://www.ncbi.nlm.nih.gov/*; *http://www.ebi.ac.uk/embl/*); [c] Three UAGG; [d] Not determined; [e] Dashes (-) indicate exons that are not searched because the exon lengths are long and do not match the exon definition.

More exons found in mouse contain short exon lengths wherase most of Human exons are much larger than normal (500 - 10,000 nt), thus predicting more skipped exons based on the exon definition model. *Hp1bp3* exon 2 has three UAGG motifs in the exon and a GGGG motif at the 5' splice site, which encodes heterochromatin 1 binding protein 3 (Le Douarin et al., 1996). This exon shows a tissue-specific splicing pattern. It is significantly skipped in the brain (39%), is primarily included in C2CI2 (93%) and is a 1:1 ratio of skipped to included mRNA in the heart (Figure 18D). *GRIN1* CI cassette exon is expressed only in the brain and the exon is skipped in a 1:1 ratio of skipped to included bands. By a search for a mouse homolog of *NCOA2* exon 13 found in the human genome, two UAGG motifs and a GGG motif (G triplet) is found in the mouse homologous exon and is listed in Table 6. *NCOA2* exon 13 has two UAGG

Table 6. Mouse cassette exons containing a GGGG motif and two or more UAGG motifs.[a]

Ensembl# _Exon# (ENSMUSG 00000-)	**HUGO, Genbank, or SWISSPROT ID #**	**Exon length (bp)**	**5' splice site sequence**	**RT-PCR evidence for exon skipping**	**cDNA and/or EST evidence for exon skipping [b]**
026959_19	*GRIN1*	111	ACGgtaa**gggg**aa	Skipped	CD363997
023938_18	*No description*	123	GAGgtca**gggg**cc	ND	No info
024947_8	*MEN1_MOUSE*	165	CAGgtgaga**gggg**	skipped	BC036287
026791_8	*GRTR8_MOUSE*	171	CTGgtaaa**gggg**a	ND	BY347810
028759_2	*Hp1bp3* [c]	198	GAGgta**gggg**ctg	skipped	AK075725
005886_13	*NCOA2*	207	CAGgtaa**ggg**ctc	skipped	BC053387
007021_2	*Syngr3*	238	CAGgt**ggg**c**gggg**	negative	No info
015852_2	*NM_030707*	208	AAGgta**gggg**act	ND [d]	No info
024112_9	*CCAH_MOUSE*	440	CAGgta**gggg**tgt	ND	No info
028782_26	*NM_173071*	620	GAGgtga**gggg**ct	ND	No info
022096_4	*HAIR_MOUSE*	781	GAGgtaa**ggggg**t	ND	No info
052325_5	*MAPB_MOUSE*	6490	CAGgtaggt**gggg**	ND	No info

[a] The search has been done in collaboration with Gene Yeo in Burge lab at MIT.

[b] Here one representative example is shown. (*http://genome.ucsc.edu*; *http://www.ncbi.nlm.nih.gov/*; *http://www.ebi.ac.uk/embl/*); [c] Three UAGG; [d] Not determined.

motifs and is also alternatively spliced, showing more skipping in heart than in brain. As mentioned above, this mouse exon contains the GGG motif instead of four Gs and shows a similar extent of exon skipping as the human exon. Thus, a G triplet sequence, not a GGGG motif, appears to be sufficient for a silencing activity in this transcript. *MEN1* exon 8 has two UAGG motifs but the splicing pattern is predominantly inclusion, showing only a trace amount of the skipped exon. This implies that *MEN1* exon 8 could contain enhancers outweighing the silencing function by the UAGG and GGGG motifs. *Syngr3* exon 2 was also chosen to test exon skipping, but skipped PCR products are not detected (Table 6), indicating that this motif search is not 100 % predictable for skipped exons. From these results, a pair of one or two UAGG motifs in the exon and a G cluster at the 5' splice site is a valuable criterion for searching for cassette exons that would be predicted to be skipped in appropriate tissues. In addition, more UAGG motifs in the exon seem to confer a higher probability for tissue-specific splicing.

2.4. Discussion

This study uses the NMDA receptor NR1 subunit (*GRIN1*) CI cassette exon as a model system to understand tissue-specific and developmental-specific splicing control mechanisms. A splicing silencing mechanism is examined since alternative exons tend to be regulated by a loss of splicing depending on environmental conditions. A tissue-specific splicing regulator, NAPOR, has been found to play a positive role in CI cassette exon inclusion (Zhang et al., 2002). Thus, an important question is how splicing of the CI cassette exon is silenced in the absence of NAPOR. Given the functional importance of the cytoplasmic domain encoded by the CI cassette exon, this work focuses on the exon and the region in its vicinity to identify splicing regulatory elements. Using an *in vivo* splicing assay system which utilizes transient expression of a splicing reporter, a G cluster at the 5' splice site and an exonic UAGG motif are identified as *cis*-acting RNA sequence elements involved in CI cassette exon silencing. The *trans*-acting protein factor, hnRNP A1, functioning with these sequence elements is identified. Deleting or inserting each motif addresses effect of an individual motif on the silencing mechanism of CI cassette exon splicing. Finally, prevalence of an exonic UAGG motif and GGGG motif (G cluster) at the 5' splice site in coding regions is investigated by a genome wide search. This pair of silencer motifs occurs infrequently in human and mouse, with the GGGG motif as a limiting factor. Further, the silencing effect is stronger as the number of UAGG motifs increases in the exon, which is consistent with the results of the CI cassette exon. Bioinformatics analysis suggests the possibility that the expression level

of hnRNP H (*HNRPH1*) is autoregulated by the silencing mechanism through the UAGG and GGGG motifs. Importantly, this study addresses a general silencing mechanism directed by hnRNP A1 and hnRNP H proteins rather than a silencing mechanism regulated by tissue-specific regulators. Here, we propose a general silencing mechanism that forms a foundation for a positive regulation model involing tissue-specific splicing regulators.

2.4.1. The position of a G cluster in the vicinity of the 5' splice site is important for splicing silencing.

This study identifies a G cluster at the 5' splice site as a splicing silencer of the CI cassette exon, which mediates the hnRNP A1 silencing function. The silencing complex contains the G cluster as well as hnRNP A1 and its binding UAGG motifs, thus inhibiting the recognition of the 5' splice site by U1 snRNP. Including a G at +6, a position within the consensus sequence of the 5' splice site, the G cluster consists of four Gs creating the GGGG motif. Therefore, the G cluster adjacent to the 5' splice site is thought to be important for silencing activity by masking the splice site from the spliceosome. The silencing activity of the G cluster is weakened (5m7) when moved to the further downstream intron. This silencing effect is not as strong as the effect of the G cluster at the normal position +7, indicating that the inserted GGG sequence at position +40 possesses weaker silencing activity. This suggests that the silencing effect of the G cluster is position-dependent. In addition, the effect is additive as shown in 5m8.

In the rat β-TM gene exon 7, silencing activity of a G cluster within an exon has been reported, which is mediated by hnRNP H (Chen et al., 1999). Unlike the G-rich

ESS in the rat β-TM exon 7, the G cluster at the 5' splice site of the CI cassette is an ISS and its silencing function is not mediated by hnRNP H. The G cluster has been known as a splicing enhancer, especially when present in a downstream intron (Nussinov, 1988; Nussinov, 1989; McCullough and Berget, 1997; McCullough and Berget, 2000). Thus, lack of the GGG motifs in brain-specific exons is likely to result in weak exons (Brudno et al., 2001). In accord with the features of brain-specific exons, there is only one GGG motif in the downstream intron (1-100 nt) of the CI cassette exon. It should be noted that the intronic (A/U)GGG repeat found in human GH intron 3 and chicken β-TM intron 7 is not only a binding motif for hnRNP F and H/H' proteins, but also resembles a binding motif for hnRNP A1 (Sirand-Pugnet et al., 1995; Cogan et al., 1997; Caputi and Zahler, 2001). However, this repeat is a splicing enhancer and is proposed to be important for mRNA export by interacting with hnRNP A1 (Sirand-Pugnet et al., 1995). Therefore, the G cluster identified in this study is a novel intronic splicing silencer in the 5' splice site of the CI cassette exon.

2.4.2. UAGG motifs and the G cluster are involved in CI cassette exon silencing mediated by hnRNP A1.

A UAGG motif is identified in the *GRIN1* CI cassette exon in the present study. Silencer sequences similar to this motif have been found in some alternative exons. However, the UAGG sequence is defined as a functional splicing silencer motif preivously in the *tau* exon 10 (Del Gatto et al., 1996) and in the *GRIN1* CI cassette exon here. Overexpression of the C-terminal domain of hnRNP A1 fused with bacteriophage MS2 coat protein represses Exon 10 splicing (Del Gatto-Konczak et al., 1999).

However, this silencing function is not mediated by the UAGG motif. Therefore, importantly, this study shows the interaction of hnRNP A1 with the UAGG motif for CI cassette exon silencing, and further, the G cluster at the 5' splice site mediates the hnRNP A1 silencing function.

The function of the G cluster can be explained by the "looping out" model (Blanchette and Chabot, 1999). From the data of hnRNP A1 overexpression, a potential intronic silencer would also be required for splicing silencing, supporting this model. An annealing activity of hnRNP A1 that is bound to UAGG motifs in the exon and probably in the intron might induce base pairing of the G cluster with polypyrimidine or C rich regions, thus possibly looping out the 5' splice site. The G cluster mediates the silencing function of hnRNP A1 but their interaction is not detected by immunoprecipitation (data not shown) in this study. This implies that G cluster could be a substrate for annealing activity of hnRNP A1 by base pairing with C-rich sequences in the regions upstream or downstream from the 5' splice site. Therefore, the G cluster might be involved in silencing by the annealing activity of hnRNP A1 when the G cluster is not interacting with hnRNP F or H/H', supporting the secondary structure model. On the other hand, the silencing function of hnRNP A1 could be explained by a "propagation model", in which a high affinity site initiates cooperative binding of multiple consecutive hnRNP A1 proteins on the exon region required for function (Zhu et al., 2001). This model explains the silencing mechanism by a network of weak interactions conferred by multimerization of hnRNP A1. The strong binding sites of hnRNP A1, which is the exonic UAGG motifs in the CI cassette exon, could initiate cooperative binding of hnRNP A1 to the G cluster, which is the relatively low-affinity binding motif, thus masking the 5' splice site. Although

hnRNP proteins have RNA-binding specificity, they are not bound exclusively to the highest-affinity sequence if non-specific RNAs are present in molar excess over the specific binding sites (Dreyfuss et al., 1993; Burd and Dreyfuss, 1994b; Dreyfuss et al., 2002). The (A/U)GGG sequence is not a high-affinity binding motif for hnRNP A1 but it cannot be ruled out that hnRNP A1 could interact with this repeat in a non-specific manner (Dreyfuss et al., 1993; Burd and Dreyfuss, 1994b; Burd and Dreyfuss, 1994a). In particular, the binding of hnRNP A1 to the intron containing the 3' splice site has been shown and is sensitive to mutations in the conserved 3' splice site (Swanson and Dreyfuss, 1988; Buvoli et al., 1990). From the crystal structure of the U1 domain of hnRNP A1, a model combining these two has been proposed that protein-protein interactions promote the specific recognition of pre-mRNA by hnRNP A1 (Ding et al., 1999). Therefore, the function of the G cluster involved in silencing of CI cassette exon splicing by hnRNP A1 could be indirect.

In contrast to hnRNP A1, hnRNP F or H/H' plays an anti-silencing role in CI cassette exon inclusion by binding to the G cluster. In conclusion, the G cluster is a splicing silencer functionally involved in repression by hnRNP A1 and also has enhancing activity by interacting with hnRNP F or H/H'.

2.4.3. A short exonic UAGG motif is involved in exon silencing of the *GRIN1* CI cassette.

The silencing motifs identified here, the exonic UAGG and the 5' splice site G cluster, (G)GGG, are only three or four nucleotides in length, which are short. In contrast to ESEs, ESSs have been unformulated. Individual examples of ESSs are

found by mutational analysis of natural exon sequences, with a premise that RNA sequence elements are required for a silencing activity when the mutation increases exon inclusion. ESS motifs have not been precisely defined by the limitations of mutational analysis unless interacting proteins are identified, such as hnRNP A1, a representative example of splicing repressor (Burd and Dreyfuss, 1994b). Thus, many studies are restricted to the identification of silencing regions by truncation or deletion mutations. For example, although regions required for regulation of splicing of mouse IgM exons M1 and M2, the regulatory sequences are not defined (Kan and Green, 1999). In other studies of diseases, if a point mutation impairs protein function, the regulatory sequence is identified by additional point mutations on the contiguous nucleotides. The UUAG sequence of Tau exon 10 has been identified because of a silent mutation in the genetic disease, FTDP-17 (Gao et al., 2000). It is extremely frequent to generate other or unkown regulatory motifs at the junction of mutated and its contiguous wild-type sequences because of degeneracy of these motifs. This causes the CI cassette exon to be sensitive to any mutations and restriction enzyme sites that are introduced in the exonic region. Although investigation of silencers using deletion or shuffled mutations could generate additional motifs, these results are helpful in narrowing down the region. However, like many other work, it is difficult to define the silencer sequences. Site-directed mutational analysis would be a solution but the effect of the ESE motifs on the mutations is still problematic because the results could be misinterpreted if the mutation disrupts both the silencing and enhancing activities. Therefore, this study takes into consideration of ESE consensus sequences while designing mutations in the CI cassette exon such that the mutated sequence would not

disrupt ESE motifs at the site of the mutation. In conclusion, this study successfully identifies short silencer motifs using known consensus ESE motifs, which suggests that splicing silencers usually overlap enhancer motifs in exons. This study also establishes that the methodology of mutation analysis for identifying exonic silencers using enhancer consensus sequences in an *in vivo* transfection assay.

2.4.4. *Cis*-acting RNA sequence elements and *trans*-acting hnRNP proteins cooperatively control tissue-specific CI cassette exon splicing.

In accord with the suggested model for hnRNP A1 function, which requires cooperative binding of hnRNP A1 to pre-mRNA (Damgaard et al., 2002; Marchand et al., 2002; Zhu et al., 2001), the results of this study implies that splicing silencing of the CI cassette exon could be mediated by an extended silencing network through weak interactions. Using the exonic UAGG motif and the G cluster at the 5' splice site of the CI cassette exon defined in this work, effects of the position and the number of the motifs are investigated. It has been known that the positions of the splicing regulatory sequences are thought to be important for splicing signal specificity (Burge et al., 1999). For example, hnRNP H binding motif is a silencer in the exon but is an enhancer in the intron (Chen et al., 1999; Chou et al., 1999). Hence, a position of the silencer disrupting its specificity would be important for splicing regulation. However, the variation of motif position from the 5' end to the 3' end of the exon does not significantly affect CI cassette exon inclusion. Rather, the number of the silencers is important for CI cassette exon silencing. Therefore, the UAGG motif is a functional exonic silencer and the frequency of occurrence in the exon region is proportional to the strength of the silencing. The G

cluster at the 5' splice site contains one G at position +6, which base pairs with U1 snRNA. Although the G cluster is involved in exon silencing, we cannot rule out the possibility that this result could be from the disruption of the 5' splice signal. Unlike the UAGG motif, the silencing function of the GGG motif is position dependent as well as number-dependent (5m5 and 5m8). In agreement with the model for hnRNP A1 function, both silencers show a stronger silencing effect as the number of binding motifs for hnRNP A1 increases. Although a direct interaction of the G cluster with hnRNP A1 is not detected in the CI cassette, the G cluster is required for strong silencing upon overexpression of hnRNP A1 proteins. The number of G clusters in the intron could contribute to the weak interactions of the silencing network. In the same context, the position of the G cluster closer to the 5' splice site would be important for efficient masking of the 5' splice site, thereby inhibiting the recognition by U1 snRNP. The results of this work are in agreement with a combinatorial control mechanism for alternative splicing. The silencer motifs cooperatively act to bring about strong silencing, indicating that a combinatorial control of these motifs confers variations in splicing patterns. This is further verified by identifying more alternative exons that are spliced in a tissue-specific manner (Figure 18).

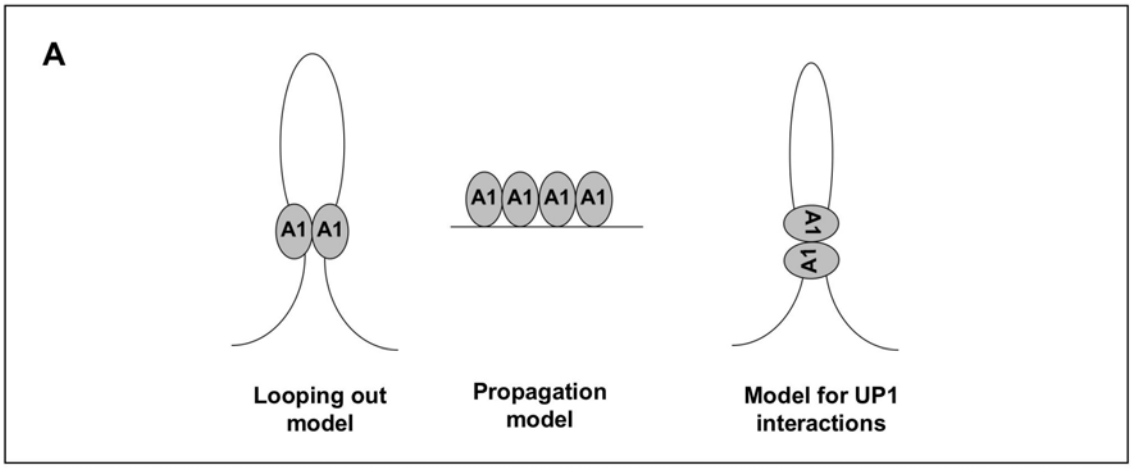
It is presently controversial whether splicing repression of the survival motor neuron 2 (*SMN2*) gene exon 7 is caused by an ESS, or disruption of an ESE. C to T transition in exon 7 creates an hnRNP A1 binding motif which is thought to act as an ESS. An increase in exon inclusion is observed when reducing hnRNP A1 protein levels using RNA interference (Kashima and Manley, 2003). In agreement with this result, splicing of the CI cassette exon containing the hnRNP A1 binding motif is repressed by

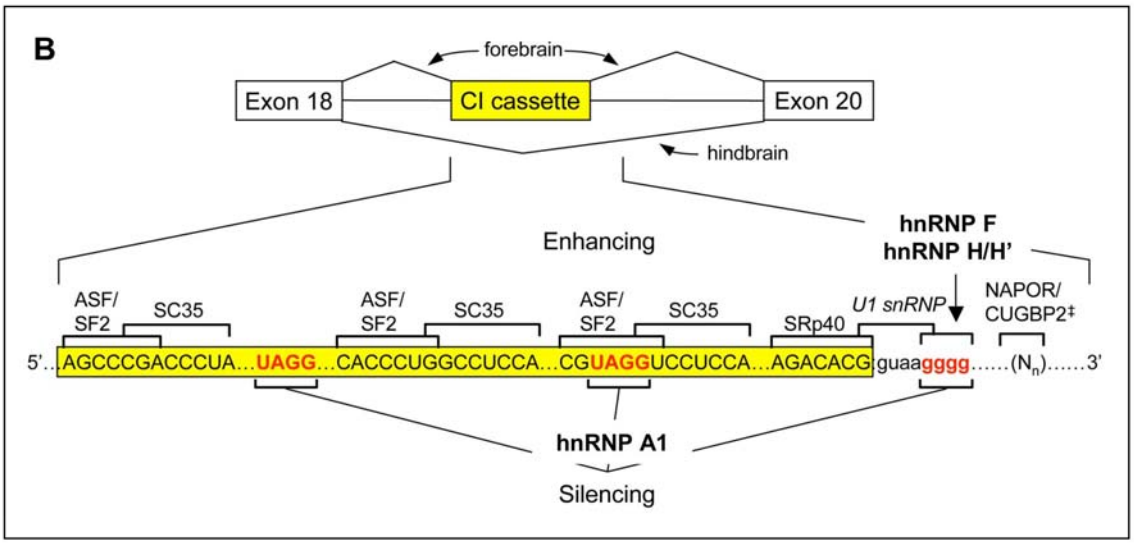
overexpression of hnRNP A1. On the other hand, the transition disrupts an ASF/SF2 motif in the *SMN1* gene and another mutation (C to A mutation) disrupting the motif results in decreasing exon inclusion (Cartegni and Krainer, 2002). Like *SMN2* exon 7, a UAGG motif, which is a binding motif for hnRNP A1, overlaps an ASF/SF2 motif in the CI cassette exon. This work also shows that this ESE motif is required for enhancing of CI cassette splicing (E5). From the results, the complexity of splicing control mechanisms might be explained. Since both hnRNP A1 and ASF/SF2 proteins are relatively abundant and ubiquitously expressed, additional factors should be involved in the control of the *SMN* exon 7 splicing as NAPOR enhances exon inclusion of the CI cassette exon. The G cluster is also involved in the silencing mechanism of CI cassette exon studied here, which suggests that the UAGG motif alon might not be sufficient for the silencing function of hnRNP A1 in some exons like the CI cassette exon. Moreover, additional alternative exons that contain the UAGG and the G cluster are identified in this study. In summary, the silencer motifs identified in the CI cassette exon appear to be involved in alternative splicing in various tissues according to the ratio of splicing factors.

2.4.5. Model for mechanism of CI cassette exon alternative splicing by the G cluster and UAGG motifs, and hnRNP proteins.

From the results of this study, a model for alternative splicing of the CI cassette is proposed (Figure 19A). In the absence of NAPOR, hnRNP A1 mediates exon silencing through two exonic UAGG motifs and a G cluster motif at the 5' splice site of the CI cassette exon (red color). Since the binding of hnRNP A1 to each motif is weak, two

UAGG motifs and a G cluster are required for strong silencing of the CI cassette exon by forming the silencing network through multiple interactions with the silencers. In the CI cassette exon, at least three silencers are required for strong silencing. Additional exonic UAGG motifs or intronic G clusters increase the strength of the silencing effect to almost complete exon silencing. This implies that the splicing signal of the CI cassette exon is not recognized by the splicing machinery when more than four silencers, multiple UAGG and G cluster motifs, occur in the vicinity of the CI cassette exon even though the splice sites and the branch point remain unchanged. The silencing function by hnRNP A1 through the UAGG and G cluster motifs requires additional splicing silencer motifs in the downstream intron, which have not yet been identified. From the results of hnRNP A1 overexpression, additional hnRNP A1 binding motifs or sequence motifs could be involved in hnRNP A1 function, suggesting other UAGG motifs or G clusters in the intron are candidates for intronic silencers. Moreover, the chance of identifying intronic silencers could be increased by looking for conserved UAGG motifs or G clusters among human, mouse and rat, from the alignment of the intronic sequences. The silencing efficiency of CI cassette exon splicing could be constrained by the number of UAGG and G cluster silencer motifs, since hnRNP A1 expression is abundant in most tissues. Further, this suggests that the silencing mechanism of the CI cassette exon by hnRNP A1 is a general exon silencing mechanism rather than a


A
A1 A1
A1 A1 A1 A1
A1
A1
Looping out model
Propagation model
Model for UP1 interactions


B
forebrain
Exon 18
CI cassette
Exon 20
hindbrain
Enhancing
hnRNP F
hnRNP H/H'
ASF/ SF2
SC35
ASF/ SF2
SC35
ASF/ SF2
SC35
SRp40
U1 snRNP
NAPOR/ CUGBP2‡
5'…AGCCCGACCCUA…UAGG…CACCCUGGCCUCCA…CGUAGGUCCUCCA…AGACACG·guaagggg……(Nn)……3'
hnRNP A1
Silencing

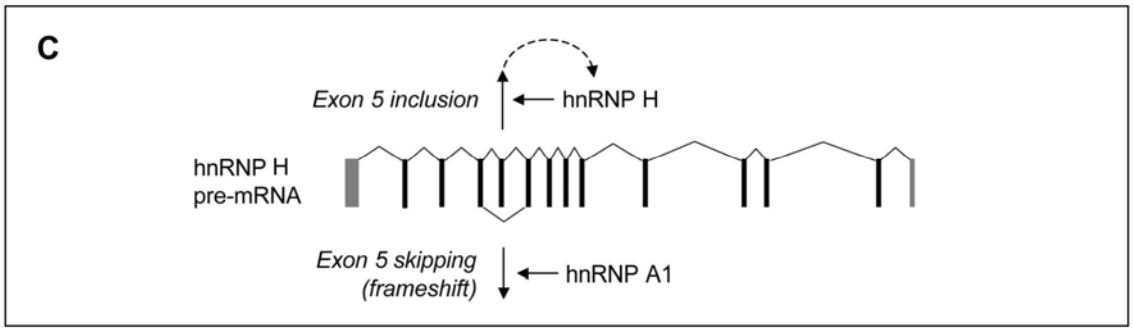

C
Exon 5 inclusion
hnRNP H
hnRNP H
pre-mRNA
Exon 5 skipping
(frameshift)
hnRNP A1

Figure 19. Model for the silencing mechanism of CI cassette exon splicing.

(A) Model for hnRNP A1 interactions. Looping out model (Blanchette and Chabot, 1999). The splice site is excluded by an interaction between hnRNP A1 proteins bound to two motifs, resulting in an increase exon skipping. Propagation model. The initial binding of hnRNP A1 to high-affinity binding motif promotes cooperative binding of consecutive A1 molecules. Model of combining these two models. A UP1 dimer brings the distal 5' splice site into a downstream 3' splice site by looping out the proximal 5' splice site. (B) Splicing silencing model for the CI cassette exon. Pre-mRNA sequences of the CI cassette exon are partially shown with sequence motifs for splicing factors identified in this study. The sequence motifs involved in splicing silencing are in red and unspecified nucleotides in this study are indicated by (N_n), the binding motif of NAPOR. The enhancing role of hnRNP F/H/H' motif is indicated by an arrow. Silencing of the CI cassette exon is mediated by hnRNP A1 through its interaction with two exonic UAGG and a 5' splice site G cluster that are identified in this study. The spliceosomal component, U1 snRNP, is believed to recognize the 5' splice site. Note that the ASF/SF2 motif encompasses the UAGG silencer motif. Taken together, the splicing decision of the CI cassette exon is proposed to occur by the relative ratio of hnRNP A1 to splicing regulators, including hnRNP F/H/H' and NAPOR in various tissues (combinatorial control mechanism). (C) Model for autoregulation of hnRNP H by positive feedback. The black bars indicate internal exons and gray bars indicate 5' and 3' UTRs. Arrows indicate exon inclusion and skipping pathways and related splicing regulators hnRNP A1 and hnRNP H. A Dotted curve with an arrow represents enhancement of hnRNP H protein levels. Skipping of exon 5 results in frameshift of translation and premature translational stop.

specific mechanism in only a certain tissue. How is the tissue-specificity of alternative splicing of the CI cassette exon achieved in a certain tissue? This could be determined by tissue-specific splicing regulators, for example, NAPOR in the forebrain. The splicing components, hnRNP A1, hnRNP F, hnRNP H/H' and SR proteins, show differential expression levels (Su et al., 2002), which could allow differential ratios sufficient for tissue-specific regulation. However, a tissue-specific regulator is a powerful tool for efficient regulation of splicing in a particular tissue. Therefore, for the *GRIN1* CI cassette exon, apparent tissue-specificity is determined by the expression of a tissue-specific positive regulator, NAPOR, in addtion that general exon silencing is driven by the silencer motifs, UAGG motifs in the exon and the G cluster at the 5' splice site.

2.4.6. Prevalence of the UAGG and GGGG motifs in human and mouse.

Consistent with results for the CI cassette exon, exon skipping is dependent on the number of UAGG motifs and strong silencing is not conferred by a single UAGG motif. Some of the exons obtained in the search are identified as alternative exons in mouse heart and brain tissues and C2C12 cells in this study. Thus, the CI cassette exon minigene system is useful to identify splicing regulatory elements in an *in vivo* expression system. In addition, the silencer motifs are predictive to search for alternative exons although the functions of identified genes do not seem to be tightly related. Based on the observations that various genes are regulated by a small number of factors under variable conditions, this pair of motifs might function as a modulator for regulation of gene expression, further conferring tissue-specificity. Interestingly, among

these genes identified, there are transcription factors, splicing factors, membrane receptors and disease-related genes (Table 4-6). Some of them function in regulation of gene expression. This imples that this pair of silencer motifs may be involved in regulation of the expression profiles of cells in response to external stimuli.

2.4.7. Implications of autoregulation of hnRNP by the UAGG and GGGG motifs predicted by genome wide analysis.

From a search for exons containing the UAGG and GGGG motifs, hnRNP H (*HNRPH1*) and hnRNP 2H9 (*HNRPH3*) are predicted to be regulated by this pair of motifs. It also suggests that the level of hnRNP H expression is autoregulated since a premature translation stop codon occurs in skipped transcripts (Figure 19B). This mRNA is predicted to be degraded by the mechanism of nonsense-mediated mRNA decay (Maquat and Carmichael, 2001; Maquat, 2002). Regulation of the expression level could be achieved by encoding truncated proteins, thus reducing the level of functional proteins. Therefore, autoregulation of hnRNP H by a positive feedback loop could be explained by either of the mechanisms. Examples of autoregualation have been reported for splicing regulators by negative feedback, such as PTB and SC35 (Blanchette and Chabot, 1999; Wollerton et al., 2004; Sureau et al., 2001). Therefore, undestanding the regulation mechanism of the CI cassette exon by hnRNP H gives an insight into the splicing regulation mechanism of the exons containing the UAGG and GGGG motif that are regulated by hnRNP H.

2.4.8. Future directions

The effect of hnRNP A1 on exons identified from a genome wide search and biochemical analysis should be tested. However, because of abundant and stable expressions of endogeneous hnRNP A1 and hnRNP H, a method using transient overexpression is difficult to modulate expression levels. Therefore, regulation of these exons including the CI cassette exon should be examined by knock down of hnRNP A1 using stable siRNA expression system.

Given these observations, compensatory mutations were designed to test the secondary structure predicted by Mulfold 2.0 but no strong evidence supporting the secondary structure model has been obtained for the CI cassette exon (Figure 10 and Table 3). Since the secondary structure model proposed by Nussinov (1988, 1989) was not verified using mutational analysis, this could be an experimental limitation. Therefore, other approaches should be employed to test the secondary structure model. Similar to UV crosslinking experiments for RNA-protein binding analysis, RNA-RNA interactions could be tested by examining base pairs within or between RNA substrates. The molecular interactions could be dissected in vitro by RNase mapping or psoralen crosslinking.

Here we utilized known ESE motifs to identify ESS in the CI cassette exon but hexamer motifs found by RESCUE-ESE (Fairbrother et al., 2002) have not been tested. Therefore, addtional ESE motifs that are not identified in this study could exist in the CI cassette exon. Given the observation that CI cassette exon splicing is sensitive even to point mutations, novel ESE or ESS motifs could be identified using the same method that has been used here.

BIBLIOGRAPHY

Abovich, N. and Rosbash, M. (1997). "Cross-Intron Bridging Interactions in the Yeast Commitment Complex Are Conserved in Mammals." Cell 89(3): 403-12.

Adami, G. and Nevins, J. R. (1988). "Splice Site Selection Dominates over Poly(a) Site Choice in Rna Production from Complex Adenovirus Transcription Units." Embo J 7(7): 2107-16.

Adams, M. D., Rudner, D. Z. and Rio, D. C. (1996). "Biochemistry and Regulation of Pre-Mrna Splicing." Curr Opin Cell Biol 8(3): 331-9.

Al-Hallaq, R. A., Yasuda, R. P. and Wolfe, B. B. (2001). "Enrichment of N-Methyl-D-Aspartate Nr1 Splice Variants and Synaptic Proteins in Rat Postsynaptic Densities." J Neurochem 77(1): 110-9.

Allikmets, R.*, et al.* (1997). "A Photoreceptor Cell-Specific Atp-Binding Transporter Gene (Abcr) Is Mutated in Recessive Stargardt Macular Dystrophy." Nat Genet 15(3): 236-46.

Amendt, B. A., Si, Z. H. and Stoltzfus, 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 Cell Biol 15(11): 6480.

Arhin, G. K., Boots, M., Bagga, P. S., Milcarek, C. and Wilusz, J. (2002). "Downstream Sequence Elements with Different Affinities for the Hnrnp H/H' Protein Influence the Processing Efficiency of Mammalian Polyadenylation Signals." Nucleic Acids Res 30(8): 1842-50.

Ashiya, M. and Grabowski, P. J. (1997). "A Neuron-Specific Splicing Switch Mediated by an Array of Pre-Mrna Repressor Sites: Evidence of a Regulatory Role for the Polypyrimidine Tract Binding Protein and a Brain-Specific Ptb Counterpart." Rna 3(9): 996-1015.

Bagga, P. S., Arhin, G. K. and Wilusz, J. (1998). "Dsef-1 Is a Member of the Hnrnp H Family of Rna-Binding Proteins and Stimulates Pre-Mrna Cleavage and Polyadenylation in Vitro." Nucleic Acids Res 26(23): 5343-50.

Barabino, S. M., Blencowe, B. J., Ryder, U., Sproat, B. S. and Lamond, A. I. (1990). "Targeted Snrnp Depletion Reveals an Additional Role for Mammalian U1 Snrnp in Spliceosome Assembly." Cell 63(2): 293-302.

Bennett, M., Michaud, S., Kingston, J. and Reed, R. (1992). "Protein Components Specifically Associated with Prespliceosome and Spliceosome Complexes." Genes Dev 6(10): 1986-2000.

Bentley, D. (1999). "Coupling Rna Polymerase Ii Transcription with Pre-Mrna Processing." Curr Opin Cell Biol 11(3): 347-51.

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 U S A 74(8): 3171-5.

Berget, S. M. (1995). "Exon Recognition in Vertebrate Splicing." J Biol Chem 270(6): 2411-4.

Black, D. L. (1995). "Finding Splice Sites within a Wilderness of Rna." Rna 1(8): 763-71.

Black, D. L. (1998). "Splicing in the Inner Ear: A Familiar Tune, but What Are the Instruments?" Neuron 20(2): 165-8.

Black, D. L. (2000). "Protein Diversity from Alternative Splicing: A Challenge for Bioinformatics and Post-Genome Biology." Cell 103(3): 367-70.
Black, D. L. (2003). "Mechanisms of Alternative Pre-Messenger Rna Splicing." Annu Rev Biochem 27: 27.
Blanchette, M. and Chabot, B. (1999). "Modulation of Exon Skipping by High-Affinity Hnrnp A1-Binding Sites and by Intron Elements That Repress Splice Site Utilization." Embo J 18(7): 1939-52.
Blencowe, B. J. (2000). "Exonic Splicing Enhancers: Mechanism of Action, Diversity and Role in Human Genetic Diseases." Trends Biochem Sci 25(3): 106-10.
Boise, L. H., Gonzalez-Garcia, M., Postema, C. E., Ding, L., Lindsten, T., Turka, L. A., Mao, X., Nunez, G. and Thompson, C. B. (1993). "Bcl-X, a Bcl-2-Related Gene That Functions as a Dominant Regulator of Apoptotic Cell Death." Cell 74(4): 597-608.
Brady, H. A. and Wold, W. S. (1988). "Competition between Splicing and Polyadenylation Reactions Determines Which Adenovirus Region E3 Mrnas Are Synthesized." Mol Cell Biol 8(8): 3291-7.
Brett, D., Hanke, J., Lehmann, G., Haase, S., Delbruck, S., Krueger, S., Reich, J. and Bork, P. (2000). "Est Comparison Indicates 38% of Human Mrnas Contain Possible Alternative Splice Forms." FEBS Lett 474(1): 83-6.
Brudno, M., Gelfand, M. S., Spengler, S., Zorn, M., Dubchak, I. and Conboy, J. G. (2001). "Computational Analysis of Candidate Intron Regulatory Elements for Tissue-Specific Alternative Pre-Mrna Splicing." Nucleic Acids Res 29(11): 2338-48.
Buckanovich, R. J. and Darnell, R. B. (1997). "The Neuronal Rna Binding Protein Nova-1 Recognizes Specific Rna Targets in Vitro and in Vivo." Mol Cell Biol 17(6): 3194-201.
Burd, C. G. and Dreyfuss, G. (1994a). "Conserved Structures and Diversity of Functions of Rna-Binding Proteins." Science 265(5172): 615-21.
Burd, C. G. and Dreyfuss, G. (1994b). "Rna Binding Specificity of Hnrnp A1: Significance of Hnrnp A1 High-Affinity Binding Sites in Pre-Mrna Splicing." Embo J 13(5): 1197-204.
Burge, C. B., Tuschl, T. and Sharp, P. A. (1999). Splicing of Precursors to Mrnas by the Spliceosomes. The Rna World. Gesteland, R. F., Cech, T. R. and Atkins, J. F., Cold Spring Harbor Press: 525-60.
Buvoli, M., Cobianchi, F., Biamonti, G. and Riva, S. (1990). "Recombinant Hnrnp Protein A1 and Its N-Terminal Domain Show Preferential Affinity for Oligodeoxynucleotides Homologous to Intron/Exon Acceptor Sites." Nucleic Acids Res 18(22): 6595-600.
Caputi, M. and Zahler, A. M. (2001). "Determination of the Rna Binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein (Hnrnp) H/H'/F/2h9 Family." J Biol Chem 276(47): 43850-9.
Carlo, T., Sierra, R. and Berget, S. M. (2000). "A 5' Splice Site-Proximal Enhancer Binds Sf1 and Activates Exon Bridging of a Microexon." Mol Cell Biol 20(11): 3988-95.

Cartegni, L., Chew, S. L. and Krainer, A. R. (2002). "Listening to Silence and Understanding Nonsense: Exonic Mutations That Affect Splicing." Nat Rev Genet 3(4): 285-98.
Cartegni, L. and Krainer, A. R. (2002). "Disruption of an Sf2/Asf-Dependent Exonic Splicing Enhancer in Smn2 Causes Spinal Muscular Atrophy in the Absence of Smn1." Nat Genet 30(4): 377-84.
Cartegni, L. and Krainer, A. R. (2003). "Correction of Disease-Associated Exon Skipping by Synthetic Exon-Specific Activators." Nat Struct Biol 10(2): 120-5.
Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. and Krainer, A. R. (2003). "Esefinder: A Web Resource to Identify Exonic Splicing Enhancers." Nucleic Acids Res 31(13): 3568-71.
Chabot, B. (1996). "Directing Alternative Splicing: Cast and Scenarios." Trends Genet 12(11): 472-8.
Chan, R. C. and Black, D. L. (1995). "Conserved Intron Elements Repress Splicing of a Neuron-Specific C-Src Exon in Vitro." Mol Cell Biol 15(11): 6377-85.
Chan, R. C. and Black, D. L. (1997). "The Polypyrimidine Tract Binding Protein Binds Upstream of Neural Cell-Specific C-Src Exon N1 to Repress the Splicing of the Intron Downstream." Mol Cell Biol 17(8): 4667-76.
Charlet, B. N., Logan, P., Singh, G. and Cooper, T. A. (2002). "Dynamic Antagonism between Etr-3 and Ptb Regulates Cell Type-Specific Alternative Splicing." Mol Cell 9(3): 649-58.
Chen, C. D., Kobayashi, R. and Helfman, D. M. (1999). "Binding of Hnrnp H to an Exonic Splicing Silencer Is Involved in the Regulation of Alternative Splicing of the Rat Beta-Tropomyosin Gene." Genes Dev 13(5): 593-606.
Cheng, S. C. and Abelson, J. (1987). "Spliceosome Assembly in Yeast." Genes Dev 1(9): 1014-27.
Chew, S. L., Liu, H. X., Mayeda, A. and Krainer, A. R. (1999). "Evidence for the Function of an Exonic Splicing Enhancer after the First Catalytic Step of Pre-Mrna Splicing." Proc Natl Acad Sci U S A 96(19): 10655-60.
Chiara, M. D. and Reed, R. (1995). "A Two-Step Mechanism for 5' and 3' Splice-Site Pairing." Nature 375(6531): 510-3.
Choi, D. K., Ito, T., Tsukahara, F., Hirai, M. and Sakaki, Y. (1999). "Developmentally-Regulated Expression of Mnapor Encoding an Apoptosis-Induced Elav-Type Rna Binding Protein." Gene 237(1): 135-42.
Chou, M. Y., Rooke, N., Turck, C. W. and Black, D. L. (1999). "Hnrnp H Is a Component of a Splicing Enhancer Complex That Activates a C-Src Alternative Exon in Neuronal Cells." Mol Cell Biol 19(1): 69-77.
Chou, M. Y., Underwood, J. G., Nikolic, J., Luu, M. H. and Black, D. L. (2000). "Multisite Rna Binding and Release of Polypyrimidine Tract Binding Protein During the Regulation of C-Src Neural-Specific Splicing." Mol Cell 5(6): 949-57.
Cogan, J. D., Prince, M. A., Lekhakula, S., Bundey, S., Futrakul, A., McCarthy, E. M. and Phillips, J. A., 3rd (1997). "A Novel Mechanism of Aberrant Pre-Mrna Splicing in Humans." Hum Mol Genet 6(6): 909-12.
Cooper, T. A. and Mattox, W. (1997). "The Regulation of Splice-Site Selection, and Its Role in Human Disease." Am J Hum Genet 61(2): 259-66.

Cull-Candy, S., Brickley, S. and Farrant, M. (2001). "Nmda Receptor Subunits: Diversity, Development and Disease." Curr Opin Neurobiol 11(3): 327-35.
D'Souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V. M., Bird, T. D. and Schellenberg, G. D. (1999). "Missense and Silent Tau Gene Mutations Cause Frontotemporal Dementia with Parkinsonism-Chromosome 17 Type, by Affecting Multiple Alternative Rna Splicing Regulatory Elements." Proc Natl Acad Sci U S A 96(10): 5598-603.
D'Souza, I. and Schellenberg, G. D. (2002). "Tau Exon 10 Expression Involves a Bipartite Intron 10 Regulatory Sequence and Weak 5' and 3' Splice Sites." J Biol Chem 277(29): 26587-99.
Damgaard, C. K., Tange, T. O. and Kjems, J. (2002). "Hnrnp A1 Controls Hiv-1 Mrna Splicing through Cooperative Binding to Intron and Exon Splicing Silencers in the Context of a Conserved Secondary Structure." Rna 8(11): 1401-15.
Del Gatto, F., Gesnel, M. C. and Breathnach, R. (1996). "The Exon Sequence Tagg Can Inhibit Splicing." Nucleic Acids Res 24(11): 2017-21.
Del Gatto-Konczak, F., Olive, M., Gesnel, M. C. and Breathnach, R. (1999). "Hnrnp A1 Recruited to an Exon in Vivo Can Function as an Exon Splicing Silencer." Mol Cell Biol 19(1): 251-60.
Dietrich, R. C., Incorvaia, R. and Padgett, R. A. (1997). "Terminal Intron Dinucleotide Sequences Do Not Distinguish between U2- and U12-Dependent Introns." Mol Cell 1(1): 151-60.
Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R. and Xu, R. M. (1999). "Crystal Structure of the Two-Rrm Domain of Hnrnp A1 (Up1) Complexed with Single-Stranded Telomeric DNA." Genes Dev 13(9): 1102-15.
Dominski, Z. and Kole, R. (1991). "Selection of Splice Sites in Pre-Mrnas with Short Internal Exons." Mol Cell Biol 11(12): 6075-83.
Dominski, Z. and Kole, R. (1992). "Cooperation of Pre-Mrna Sequence Elements in Splice Site Selection." Mol Cell Biol 12(5): 2108-14.
Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. and Burd, C. G. (1993). "Hnrnp Proteins and the Biogenesis of Mrna." Annu Rev Biochem 62: 289-321.
Dreyfuss, G., Kim, V. N. and Kataoka, N. (2002). "Messenger-Rna-Binding Proteins and the Messages They Carry." Nat Rev Mol Cell Biol 3(3): 195-205.
Edery, I. and Sonenberg, N. (1985). "Cap-Dependent Rna Splicing in a Hela Nuclear Extract." Proc Natl Acad Sci U S A 82(22): 7590-4.
Ehlers, M. D., Tingley, W. G. and Huganir, R. L. (1995). "Regulated Subcellular Distribution of the Nr1 Subunit of the Nmda Receptor." Science 269(5231): 1734-7.
Ehlers, M. D., Mammen, A. L., Lau, L. F. and Huganir, R. L. (1996). "Synaptic Targeting of Glutamate Receptors." Curr Opin Cell Biol 8(4): 484-9.
Ehlers, M. D., Fung, E. T., O'Brien, R. J. and Huganir, R. L. (1998). "Splice Variant-Specific Interaction of the Nmda Receptor Subunit Nr1 with Neuronal Intermediate Filaments." J Neurosci 18(2): 720-30.
Expert-Bezancon, A., Le Caer, J. P. and Marie, J. (2002). "Heterogeneous Nuclear Ribonucleoprotein (Hnrnp) K Is a Component of an Intronic Splicing Enhancer Complex That Activates the Splicing of the Alternative Exon 6a from Chicken Beta-Tropomyosin Pre-Mrna." J Biol Chem 277(19): 16614-23.

Fairbrother, W. G., Yeh, R. F., Sharp, P. A. and Burge, C. B. (2002). "Predictive Identification of Exonic Splicing Enhancers in Human Genes." Science 297(5583): 1007-13.
Fogel, B. L. and McNally, M. T. (2000). "A Cellular Protein, Hnrnp H, Binds to the Negative Regulator of Splicing Element from Rous Sarcoma Virus." J Biol Chem 275(41): 32371-8.
Fu, X. D. (1995). "The Superfamily of Arginine/Serine-Rich Splicing Factors." Rna 1(7): 663-80.
Gao, Q. S., Memmott, J., Lafyatis, R., Stamm, S., Screaton, G. and Andreadis, A. (2000). "Complex Regulation of Tau Exon 10, Whose Missplicing Causes Frontotemporal Dementia." J Neurochem 74(2): 490-500.
Garcia-Blanco, M. A., Jamison, S. F. and Sharp, P. A. (1989). "Identification and Purification of a 62,000-Dalton Protein That Binds Specifically to the Polypyrimidine Tract of Introns." Genes Dev 3(12A): 1874-86.
Gatermann, K. B., Hoffmann, A., Rosenberg, G. H. and Kaufer, N. F. (1989). "Introduction of Functional Artificial Introns into the Naturally Intronless Ura4 Gene of Schizosaccharomyces Pombe." Mol Cell Biol 9(4): 1526-35.
Good, P. J. (1997). "The Role of Elav-Like Genes, a Conserved Family Encoding Rna-Binding Proteins, in Growth and Development." Semin Cell Dev Biol 8(6): 577-84.
Grabowski, P. J., Nasim, F. U., Kuo, H. C. and Burch, R. (1991). "Combinatorial Splicing of Exon Pairs by Two-Site Binding of U1 Small Nuclear Ribonucleoprotein Particle." Mol Cell Biol 11(12): 5919-28.
Grabowski, P. J. (1998). "Splicing Regulation in Neurons: Tinkering with Cell-Specific Control." Cell 92(6): 709-12.
Grabowski, P. J. and Black, D. L. (2001). "Alternative Rna Splicing in the Nervous System." Prog Neurobiol 65(3): 289-308.
Graveley, B. R. and Maniatis, T. (1998). "Arginine/Serine-Rich Domains of Sr Proteins Can Function as Activators of Pre-Mrna Splicing." Mol Cell 1(5): 765-71.
Graveley, B. R. (2000). "Sorting out the Complexity of Sr Protein Functions." Rna 6(9): 1197-211.
Graveley, B. R. (2001). "Alternative Splicing: Increasing Diversity in the Proteomic World." Trends Genet 17(2): 100-7.
Graveley, B. R. (2002). "Sex, Agility, and the Regulation of Alternative Splicing." Cell 109(4): 409-12.
Green, R. E., Lewis, B. P., Hillman, R. T., Blanchette, M., Lareau, L. F., Garnett, A. T., Rio, D. C. and Brenner, S. E. (2003). "Widespread Predicted Nonsense-Mediated Mrna Decay of Alternatively-Spliced Transcripts of Human Normal and Disease Genes." Bioinformatics 19(Suppl 1): i118-21.
Grover, A., Houlden, H., Baker, M., Adamson, J., Lewis, J., Prihar, G., Pickering-Brown, S., Duff, K. and Hutton, M. (1999). "5' Splice Site Mutations in Tau Associated with the Inherited Dementia Ftdp-17 Affect a Stem-Loop Structure That Regulates Alternative Splicing of Exon 10." J Biol Chem 274(21): 15134-43.
Guo, W. and Helfman, D. M. (1993). "Cis-Elements Involved in Alternative Splicing in the Rat Beta-Tropomyosin Gene: The 3'-Splice Site of the Skeletal Muscle Exon 7 Is the Major Site of Blockage in Nonmuscle Cells." Nucleic Acids Res 21(20): 4762-8.

Guthrie, C. and Patterson, B. (1988). "Spliceosomal Snrnas." Annu Rev Genet 22: 387-419.
Hall, S. L. and Padgett, R. A. (1994). "Conserved Sequences in a Class of Rare Eukaryotic Nuclear Introns with Non-Consensus Splice Sites." J Mol Biol 239(3): 357-65.
Hall, S. L. and Padgett, R. A. (1996). "Requirement of U12 Snrna for in Vivo Splicing of a Minor Class of Eukaryotic Nuclear Pre-Mrna Introns." Science 271(5256): 1716-8.
Hartmuth, K., Urlaub, H., Vornlocher, H. P., Will, C. L., Gentzel, M., Wilm, M. and Luhrmann, R. (2002). "Protein Composition of Human Prespliceosomes Isolated by a Tobramycin Affinity-Selection Method." Proc Natl Acad Sci U S A 99(26): 16719-24.
Hillman, R. T., Green, R. E. and Brenner, S. E. (2004). "An Unappreciated Role for Rna Surveillance." Genome Biol 5(2): R8. Epub 2004 Feb 2.
Hirose, Y., Tacke, R. and Manley, J. L. (1999). "Phosphorylated Rna Polymerase Ii Stimulates Pre-Mrna Splicing." Genes Dev 13(10): 1234-9.
Hirose, Y. and Manley, J. L. (2000). "Rna Polymerase Ii and the Integration of Nuclear Events." Genes Dev 14(12): 1415-29.
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(12B): 2554-68.
Hollmann, M., Boulter, J., Maron, C., Beasley, L., Sullivan, J., Pecht, G. and Heinemann, S. (1993). "Zinc Potentiates Agonist-Induced Currents at Certain Splice Variants of the Nmda Receptor." Neuron 10(5): 943-54.
Hong, M.*, et al.* (1998). "Mutation-Specific Functional Impairments in Distinct Tau Isoforms of Hereditary Ftdp-17." Science 282(5395): 1914-7.
Horowitz, D. S. and Krainer, A. R. (1997). "A Human Protein Required for the Second Step of Pre-Mrna Splicing Is Functionally Related to a Yeast Splicing Factor." Genes Dev 11(1): 139-51.
Hutton, M.*, et al.* (1998). "Association of Missense and 5'-Splice-Site Mutations in Tau with the Inherited Dementia Ftdp-17." Nature 393(6686): 702-5.
Hwang, D. M., Hwang, W. S. and Liew, C. C. (1994). "Single Pass Sequencing of a Unidirectional Human Fetal Heart Cdna Library to Discover Novel Genes of the Cardiovascular System." J Mol Cell Cardiol 26(10): 1329-33.
Inoue, S., Ito, S., Wakamatsu, K., Jimbow, K. and Fujita, K. (1990). "Mechanism of Growth Inhibition of Melanoma Cells by 4-S-Cysteaminylphenol and Its Analogues." Biochem Pharmacol 39(6): 1077-83.
Jackson, I. J. (1991). "A Reappraisal of Non-Consensus Mrna Splice Sites." Nucleic Acids Res 19(14): 3795-8.
Jacquenet, S., Ropers, D., Bilodeau, P. S., Damier, L., Mougin, A., Stoltzfus, C. M. and Branlant, C. (2001). "Conserved Stem-Loop Structures in the Hiv-1 Rna Region Containing the A3 3' Splice Site and Its Cis-Regulatory Element: Possible Involvement in Rna Splicing." Nucleic Acids Res 29(2): 464-78.
Jaeger, J. A., Turner, D. H. and Zuker, M. (1989). "Improved Predictions of Secondary Structures for Rna." Proc Natl Acad Sci U S A 86(20): 7706-10.

Jensen, K. B., Dredge, B. K., Stefani, G., Zhong, R., Buckanovich, R. J., Okano, H. J., Yang, Y. Y. and Darnell, R. B. (2000). "Nova-1 Regulates Neuron-Specific Alternative Splicing and Is Essential for Neuronal Viability." Neuron 25(2): 359-71.

Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R. and Shoemaker, D. D. (2003). "Genome-Wide Survey of Human Alternative Pre-Mrna Splicing with Exon Junction Microarrays." Science 302(5653): 2141-4.

Jurica, M. S. and Moore, M. J. (2003). "Pre-Mrna Splicing: Awash in a Sea of Proteins." Mol Cell 12(1): 5-14.

Kan, J. L. and Green, M. R. (1999). "Pre-Mrna Splicing of Igm Exons M1 and M2 Is Directed by a Juxtaposed Splicing Enhancer and Inhibitor." Genes Dev 13(4): 462-71.

Kanopka, A., Muhlemann, O. and Akusjarvi, G. (1996). "Inhibition by Sr Proteins of Splicing of a Regulated Adenovirus Pre-Mrna." Nature 381(6582): 535-8.

Kashima, T. and Manley, J. L. (2003). "A Negative Element in Smn2 Exon 7 Inhibits Splicing in Spinal Muscular Atrophy." Nat Genet 34(4): 460-3.

Keene, J. D. (2001). "Ribonucleoprotein Infrastructure Regulating the Flow of Genetic Information between the Genome and the Proteome." Proc Natl Acad Sci U S A 98(13): 7018-24.

Konarska, M. M., Padgett, R. A. and Sharp, P. A. (1984). "Recognition of Cap Structure in Splicing in Vitro of Mrna Precursors." Cell 38(3): 731-6.

Konarska, M. M. and Sharp, P. A. (1987). "Interactions between Small Nuclear Ribonucleoprotein Particles in Formation of Spliceosomes." Cell 49(6): 763-74.

Konig, H., Ponta, H. and Herrlich, P. (1998). "Coupling of Signal Transduction to Alternative Pre-Mrna Splicing by a Composite Splice Regulator." EMBO J 17(10): 2904-13.

Koushika, S. P., Lisbin, M. J. and White, K. (1996). "Elav, a Drosophila Neuron-Specific Protein, Mediates the Generation of an Alternatively Spliced Neural Protein Isoform." Curr Biol 6(12): 1634-41.

Kramer, A., Mulhauser, F., Wersig, C., Groning, K. and Bilbe, G. (1995). "Mammalian Splicing Factor Sf3a120 Represents a New Member of the Surp Family of Proteins and Is Homologous to the Essential Splicing Factor Prp21p of Saccharomyces Cerevisiae." Rna 1(3): 260-72.

Krawczak, M., Reiss, J. and Cooper, D. N. (1992). "The Mutational Spectrum of Single Base-Pair Substitutions in Mrna Splice Junctions of Human Genes: Causes and Consequences." Hum Genet 90(1-2): 41-54.

Kriventseva, E. V., Koch, I., Apweiler, R., Vingron, M., Bork, P., Gelfand, M. S. and Sunyaev, S. (2003). "Increase of Functional Diversity by Alternative Splicing." Trends Genet 19(3): 124-8.

Kuo, H. C., Nasim, F. H. and Grabowski, P. J. (1991). "Control of Alternative Splicing by the Differential Binding of U1 Small Nuclear Ribonucleoprotein Particle." Science 251(4997): 1045-50.

Ladd, A. N., Charlet, N. and Cooper, T. A. (2001). "The Celf Family of Rna Binding Proteins Is Implicated in Cell-Specific and Developmentally Regulated Alternative Splicing." Mol Cell Biol 21(4): 1285-96.

Ladd, A. N. and Cooper, T. A. (2002). "Finding Signals That Regulate Alternative Splicing in the Post-Genomic Era." Genome Biol 3(11): reviews0008.
Ladd, A. N. and Cooper, T. A. (2004). "Multiple Domains Control the Subcellular Localization and Activity of Etr-3, a Regulator of Nuclear and Cytoplasmic Rna Processing Events." J Cell Sci 117(Pt 16): 3519-29. Epub 2004 Jun 29.
Lamba, J. K., Adachi, M., Sun, D., Tammur, J., Schuetz, E. G., Allikmets, R. and Schuetz, J. D. (2003). "Nonsense Mediated Decay Downregulates Conserved Alternatively Spliced Abcc4 Transcripts Bearing Nonsense Codons." Hum Mol Genet 12(2): 99-109.
Lander, E. S.*, et al.* (2001). "Initial Sequencing and Analysis of the Human Genome." Nature 409(6822): 860-921.
Laurie, D. J. and Seeburg, P. H. (1994). "Ligand Affinities at Recombinant N-Methyl-D-Aspartate Receptors Depend on Subunit Composition." Eur J Pharmacol 268(3): 335-45.
Laurie, D. J., Putzke, J., Zieglgansberger, W., Seeburg, P. H. and Tolle, T. R. (1995). "The Distribution of Splice Variants of the Nmdar1 Subunit Mrna in Adult Rat Brain." Brain Res Mol Brain Res 32(1): 94-108.
Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R. and Chambon, P. (1996). "A Possible Involvement of Tif1 Alpha and Tif1 Beta in the Epigenetic Control of Transcription by Nuclear Receptors." Embo J 15(23): 6701-15.
Levitt, N., Briggs, D., Gil, A. and Proudfoot, N. J. (1989). "Definition of an Efficient Synthetic Poly(a) Site." Genes Dev 3(7): 1019-25.
Lewis, B. P., Green, R. E. and Brenner, S. E. (2003). "Evidence for the Widespread Coupling of Alternative Splicing and Nonsense-Mediated Mrna Decay in Humans." Proc Natl Acad Sci U S A 100(1): 189-92. Epub 2002 Dec 26.
Liao, X. C., Colot, H. V., Wang, Y. and Rosbash, M. (1992). "Requirements for U2 Snrnp Addition to Yeast Pre-Mrna." Nucleic Acids Res 20(16): 4237-45.
Lim, L. P. and Burge, C. B. (2001). "A Computational Analysis of Sequence Features Involved in Recognition of Short Introns." Proc Natl Acad Sci U S A 98(20): 11193-8.
Liu, H., Zhang, W., Reed, R. B., Liu, W. and Grabowski, P. J. (2002). "Mutations in Rrm4 Uncouple the Splicing Repression and Rna-Binding Activities of Polypyrimidine Tract Binding Protein." Rna 8(2): 137-49.
Liu, H. X., Zhang, M. and Krainer, A. R. (1998). "Identification of Functional Exonic Splicing Enhancer Motifs Recognized by Individual Sr Proteins." Genes Dev 12(13): 1998-2012.
Liu, H. X., Chew, S. L., Cartegni, L., Zhang, M. Q. and Krainer, A. R. (2000). "Exonic Splicing Enhancer Motif Recognized by Human Sc35 under Splicing Conditions." Mol Cell Biol 20(3): 1063-71.
Lopez, A. J. (1998). "Alternative Splicing of Pre-Mrna: Developmental Consequences and Mechanisms of Regulation." Annu Rev Genet 32: 279-305.
Lynch, K. W. and Maniatis, T. (1996). "Assembly of Specific Sr Protein Complexes on Distinct Regulatory Elements of the Drosophila Doublesex Splicing Enhancer." Genes Dev 10(16): 2089-101.

Manley, J. L. and Tacke, R. (1996). "Sr Proteins and Splicing Control." Genes Dev 10(13): 1569-79.
Maquat, L. E. and Carmichael, G. G. (2001). "Quality Control of Mrna Function." Cell 104(2): 173-6.
Maquat, L. E. (2002). "Nonsense-Mediated Mrna Decay." Curr Biol 12(6): R196-7.
Marchand, V., Mereau, A., Jacquenet, S., Thomas, D., Mougin, A., Gattoni, R., Stevenin, J. and Branlant, C. (2002). "A Janus Splicing Regulatory Element Modulates Hiv-1 Tat and Rev Mrna Production by Coordination of Hnrnp A1 Cooperative Binding." J Mol Biol 323(4): 629-52.
Markovtsov, V., Nikolic, J. M., Goldman, J. A., Turck, C. W., Chou, M. Y. and Black, D. L. (2000). "Cooperative Assembly of an Hnrnp Complex Induced by a Tissue-Specific Homolog of Polypyrimidine Tract Binding Protein." Mol Cell Biol 20(20): 7463-79.
Matunis, M. J., Michael, W. M. and Dreyfuss, G. (1992). "Characterization and Primary Structure of the Poly(C)-Binding Heterogeneous Nuclear Ribonucleoprotein Complex K Protein." Mol Cell Biol 12(1): 164-71.
Matunis, M. J., Xing, J. and Dreyfuss, G. (1994). "The Hnrnp F Protein: Unique Primary Structure, Nucleic Acid-Binding Properties, and Subcellular Localization." Nucleic Acids Res 22(6): 1059-67.
McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D. L. (1997). "5'-Capping Enzymes Are Targeted to Pre-Mrna by Binding to the Phosphorylated Carboxy-Terminal Domain of Rna Polymerase Ii." Genes Dev 11(24): 3306-18.
McCullough, A. J. and Berget, S. M. (1997). "G Triplets Located Throughout a Class of Small Vertebrate Introns Enforce Intron Borders and Regulate Splice Site Selection." Mol Cell Biol 17(8): 4562-71.
McCullough, A. J. and Berget, S. M. (2000). "An Intronic Splicing Enhancer Binds U1 Snrnps to Enhance Splicing and Select 5' Splice Sites." Mol Cell Biol 20(24): 9225-35.
Michaud, S. and Reed, R. (1991). "An Atp-Independent Complex Commits Pre-Mrna to the Mammalian Spliceosome Assembly Pathway." Genes Dev 5(12B): 2534-46.
Miki, T., Bottaro, D. P., Fleming, T. P., Smith, C. L., Burgess, W. H., Chan, A. M. and Aaronson, S. A. (1992). "Determination of Ligand-Binding Specificity by Alternative Splicing: Two Distinct Growth Factor Receptors Encoded by a Single Gene." Proc Natl Acad Sci U S A 89(1): 246-50.
Min, H., Chan, R. C. and Black, D. L. (1995). "The Generally Expressed Hnrnp F Is Involved in a Neural-Specific Pre-Mrna Splicing Event." Genes Dev 9(21): 2659-71.
Modafferi, E. F. and Black, D. L. (1999). "Combinatorial Control of a Neuron-Specific Exon." Rna 5(5): 687-706.
Moore, M. J. and Sharp, P. A. (1992). "Site-Specific Modification of Pre-Mrna: The 2'-Hydroxyl Groups at the Splice Sites." Science 256(5059): 992-7.
Moore, M. J., Query, C. C. and Sharp, P. A. (1993). Splicing of Precursors to Mrnas by the Spliceosome. The Rna World. Gesteland, R. F., Cech, T. R. and Atkins, J. F., Cold Spring Harbor Press**:** 303-357.

Moore, M. J. and Query, C. C. (2000). "Joining of Rnas by Splinted Ligation." Methods Enzymol 317: 109-23.
Moriyoshi, K., Masu, M., Ishii, T., Shigemoto, R., Mizuno, N. and Nakanishi, S. (1991). "Molecular Cloning and Characterization of the Rat Nmda Receptor." Nature 354(6348): 31-7.
Mullen, M. P., Smith, C. W., Patton, J. G. and Nadal-Ginard, B. (1991). "Alpha-Tropomyosin Mutually Exclusive Exon Selection: Competition between Branchpoint/Polypyrimidine Tracts Determines Default Exon Choice." Genes Dev 5(4): 642-55.
Mulligan, G. J., Guo, W., Wormsley, S. and Helfman, D. M. (1992). "Polypyrimidine Tract Binding Protein Interacts with Sequences Involved in Alternative Splicing of Beta-Tropomyosin Pre-Mrna." J Biol Chem 267(35): 25480-7.
Nasim, F. H., Spears, P. A., Hoffmann, H. M., Kuo, H. C. and Grabowski, P. J. (1990). "A Sequential Splicing Mechanism Promotes Selection of an Optimal Exon by Repositioning a Downstream 5' Splice Site in Preprotachykinin Pre-Mrna." Genes Dev 4(7): 1172-84.
Nesic, D. and Maquat, L. E. (1994). "Upstream Introns Influence the Efficiency of Final Intron Removal and Rna 3'-End Formation." Genes Dev 8(3): 363-75.
Niwa, M., Rose, S. D. and Berget, S. M. (1990). "In Vitro Polyadenylation Is Stimulated by the Presence of an Upstream Intron." Genes Dev 4(9): 1552-9.
Niwa, M. and Berget, S. M. (1991). "Mutation of the Aauaaa Polyadenylation Signal Depresses in Vitro Splicing of Proximal but Not Distal Introns." Genes Dev 5(11): 2086-95.
Niwa, M., MacDonald, C. C. and Berget, S. M. (1992). "Are Vertebrate Exons Scanned During Splice-Site Selection?" Nature 360(6401): 277-80.
Nussinov, R. (1988). "Conserved Quartets near 5' Intron Junctions in Primate Nuclear Pre-Mrna." J Theor Biol 133(1): 73-84.
Nussinov, R. (1989). "Conserved Signals around the 5' Splice Sites in Eukaryotic Nuclear Precursor Mrnas: G-Runs Are Frequent in the Introns and C in the Exons near Both 5' and 3' Splice Sites." J Biomol Struct Dyn 6(5): 985-1000.
O'Brien, R. J., Lau, L. F. and Huganir, R. L. (1998). "Molecular Mechanisms of Glutamate Receptor Clustering at Excitatory Synapses." Curr Opin Neurobiol 8(3): 364-9.
O'Day, C. L., Dalbadie-McFarland, G. and Abelson, J. (1996). "The Saccharomyces Cerevisiae Prp5 Protein Has Rna-Dependent Atpase Activity with Specificity for U2 Small Nuclear Rna." J Biol Chem 271(52): 33261-7.
Ohno, M., Sakamoto, H. and Shimura, Y. (1987). "Preferential Excision of the 5' Proximal Intron from Mrna Precursors with Two Introns as Mediated by the Cap Structure." Proc Natl Acad Sci U S A 84(15): 5187-91.
Okabe, S., Miwa, A. and Okado, H. (1999). "Alternative Splicing of the C-Terminal Domain Regulates Cell Surface Expression of the Nmda Receptor Nr1 Subunit." J Neurosci 19(18): 7781-92.
Okano, H. J. and Darnell, R. B. (1997). "A Hierarchy of Hu Rna Binding Proteins in Developing and Adult Neurons." J Neurosci 17(9): 3024-37.

Patton, J. G., Mayer, S. A., Tempst, P. and Nadal-Ginard, B. (1991). "Characterization and Molecular Cloning of Polypyrimidine Tract-Binding Protein: A Component of a Complex Necessary for Pre-Mrna Splicing." Genes Dev 5(7): 1237-51.
Patzelt, E., Thalmann, E., Hartmuth, K., Blaas, D. and Kuechler, E. (1987). "Assembly of Pre-Mrna Splicing Complex Is Cap Dependent." Nucleic Acids Res 15(4): 1387-99.
Philips, A. V., Timchenko, L. T. and Cooper, T. A. (1998). "Disruption of Splicing Regulated by a Cug-Binding Protein in Myotonic Dystrophy." Science 280(5364): 737-41.
Polydorides, A. D., Okano, H. J., Yang, Y. Y., Stefani, G. and Darnell, R. B. (2000). "A Brain-Enriched Polypyrimidine Tract-Binding Protein Antagonizes the Ability of Nova to Regulate Neuron-Specific Alternative Splicing." Proc Natl Acad Sci U S A 97(12): 6350-5.
Query, C. C., McCaw, P. S. and Sharp, P. A. (1997). "A Minimal Spliceosomal Complex a Recognizes the Branch Site and Polypyrimidine Tract." Mol Cell Biol 17(5): 2944-53.
Reed, R. (1996). "Initial Splice-Site Recognition and Pairing During Pre-Mrna Splicing." Curr Opin Genet Dev 6(2): 215-20.
Reed, R. and Palandjian, L. (1997). Spliceosomal Assembly. Eukaryotic Mrna Processing. Krainer, A. R., Oxford**:** 103-129.
Reed, R. (2003). "Coupling Transcription, Splicing and Mrna Export." Curr Opin Cell Biol 15(3): 326-31.
Rio, D. C. (1993). "Splicing of Pre-Mrna: Mechanism, Regulation and Role in Development." Curr Opin Genet Dev 3(4): 574-84.
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(1): 84-94.
Roberts, G. C. and Smith, C. W. (2002). "Alternative Splicing: Combinatorial Output from the Genome." Curr Opin Chem Biol 6(3): 375-83.
Rosbash, M. and Seraphin, B. (1991). "Who's on First? The U1 Snrnp-5' Splice Site Interaction and Splicing." Trends Biochem Sci 16(5): 187-90.
Roscigno, R. F. and Garcia-Blanco, M. A. (1995). "Sr Proteins Escort the U4/U6.U5 Tri-Snrnp to the Spliceosome." Rna 1(7): 692-706.
Roy, J., Kim, K., Maddock, J. R., Anthony, J. G. and Woolford, J. L., Jr. (1995). "The Final Stages of Spliceosome Maturation Require Spp2p That Can Interact with the Deah Box Protein Prp2p and Promote Step 1 of Splicing." Rna 1(4): 375-90.
Ruby, S. W. and Abelson, J. (1988). "An Early Hierarchic Role of U1 Small Nuclear Ribonucleoprotein in Spliceosome Assembly." Science 242(4881): 1028-35.
Ryan, K. J. and Cooper, T. A. (1996). "Muscle-Specific Splicing Enhancers Regulate Inclusion of the Cardiac Troponin T Alternative Exon in Embryonic Skeletal Muscle." Mol Cell Biol 16(8): 4014-23.
Salzman, N. P. (1961). "Animal Cell Cultures." Science 133: 1559-65.
Schaal, T. D. and Maniatis, T. (1999). "Selection and Characterization of Pre-Mrna Splicing Enhancers: Identification of Novel Sr Protein-Specific Enhancer Sequences." Mol Cell Biol 19(3): 1705-19.

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E. and Zipursky, S. L. (2000). "Drosophila Dscam Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity." Cell 101(6): 671-84.
Scott, D. B., Blanpied, T. A., Swanson, G. T., Zhang, C. and Ehlers, M. D. (2001). "An Nmda Receptor Er Retention Signal Regulated by Phosphorylation and Alternative Splicing." J Neurosci 21(9): 3063-72.
Scott, D. B., Blanpied, T. A. and Ehlers, M. D. (2003). "Coordinated Pka and Pkc Phosphorylation Suppresses Rxr-Mediated Er Retention and Regulates the Surface Delivery of Nmda Receptors." Neuropharmacology 45(6): 755-67.
Seraphin, B., Kretzner, L. and Rosbash, M. (1988). "A U1 Snrna:Pre-Mrna Base Pairing Interaction Is Required Early in Yeast Spliceosome Assembly but Does Not Uniquely Define the 5' Cleavage Site." Embo J 7(8): 2533-8.
Seraphin, B. and Rosbash, M. (1990). "Exon Mutations Uncouple 5' Splice Site Selection from U1 Snrna Pairing." Cell 63(3): 619-29.
Service, R. F. (2001). "Proteomics. High-Speed Biologists Search for Gold in Proteins." Science 294(5549): 2074-7.
Sharp, P. A. and Burge, C. B. (1997). "Classification of Introns: U2-Type or U12-Type." Cell 91(7): 875-9.
Shatkin, A. J. (1976). "Capping of Eucaryotic Mrnas." Cell 9(4 PT 2): 645-53.
Si, Z. H., Rauch, D. and Stoltzfus, C. M. (1998). "The Exon Splicing Silencer in Human Immunodeficiency Virus Type 1 Tat Exon 3 Is Bipartite and Acts Early in Spliceosome Assembly." Mol Cell Biol 18(9): 5404-13.
Singh, G., Charlet, B. N., Han, J. and Cooper, T. A. (2004). "Etr-3 and Celf4 Protein Domains Required for Rna Binding and Splicing Activity in Vivo." Nucleic Acids Res 32(3): 1232-41. Print 2004.
Sirand-Pugnet, P., Durosay, P., Brody, E. and Marie, J. (1995). "An Intronic (a/U)Ggg Repeat Enhances the Splicing of an Alternative Intron of the Chicken Beta-Tropomyosin Pre-Mrna." Nucleic Acids Res 23(17): 3501-7.
Smith, C. W. and Valcarcel, J. (2000). "Alternative Pre-Mrna Splicing: The Logic of Combinatorial Control." Trends Biochem Sci 25(8): 381-8.
Staffa, A. and Cochrane, A. (1995). "Identification of Positive and Negative Splicing Regulatory Elements within the Terminal Tat-Rev Exon of Human Immunodeficiency Virus Type 1." Mol Cell Biol 15(8): 4597-605.
Stamm, S. (2002). "Signals and Their Transduction Pathways Regulating Alternative Splicing: A New Dimension of the Human Genome." Hum Mol Genet 11(20): 2409-16.
Standaert, D. G., Testa, C. M., Young, A. B. and Penney, J. B., Jr. (1994). "Organization of N-Methyl-D-Aspartate Glutamate Receptor Gene Expression in the Basal Ganglia of the Rat." J Comp Neurol 343(1): 1-16.
Standley, S., Roche, K. W., McCallum, J., Sans, N. and Wenthold, R. J. (2000). "Pdz Domain Suppression of an Er Retention Signal in Nmda Receptor Nr1 Splice Variants." Neuron 28(3): 887-98.
Su, A. I., Cooke, M. P., Ching, K. A., Hakak, Y., Walker, J. R., Wiltshire, T., Orth, A. P., Vega, R. G., Sapinoso, L. M., Moqrich, A., Patapoutian, A., Hampton, G. M., Schultz, P. G. and Hogenesch, J. B. (2002). "Large-Scale Analysis of the Human

and Mouse Transcriptomes." Proc Natl Acad Sci U S A 99(7): 4465-70. Epub 2002 Mar 19.
Sugihara, H., Moriyoshi, K., Ishii, T., Masu, M. and Nakanishi, S. (1992). "Structures and Properties of Seven Isoforms of the Nmda Receptor Generated by Alternative Splicing." Biochem Biophys Res Commun 185(3): 826-32.
Sureau, A., Gattoni, R., Dooghe, Y., Stevenin, J. and Soret, J. (2001). "Sc35 Autoregulates Its Expression by Promoting Splicing Events That Destabilize Its Mrnas." Embo J 20(7): 1785-96.
Swanson, M. S. and Dreyfuss, G. (1988). "Rna Binding Specificity of Hnrnp Proteins: A Subset Bind to the 3' End of Introns." Embo J 7(11): 3519-29.
Tacke, R. and Manley, J. L. (1999). "Determinants of Sr Protein Specificity." Curr Opin Cell Biol 11(3): 358-62.
Talerico, M. and Berget, S. M. (1990). "Effect of 5' Splice Site Mutations on Splicing of the Preceding Intron." Mol Cell Biol 10(12): 6299-305.
Talerico, M. and Berget, S. M. (1994). "Intron Definition in Splicing of Small Drosophila Introns." Mol Cell Biol 14(5): 3434-45.
Tange, T. O., Damgaard, C. K., Guth, S., Valcarcel, J. and Kjems, J. (2001). "The Hnrnp A1 Protein Regulates Hiv-1 Tat Splicing Via a Novel Intron Silencer Element." Embo J 20(20): 5748-58.
Tarn, W. Y. and Steitz, J. A. (1996a). "Highly Diverged U4 and U6 Small Nuclear Rnas Required for Splicing Rare at-Ac Introns." Science 273(5283): 1824-32.
Tarn, W. Y. and Steitz, J. A. (1996b). "A Novel Spliceosome Containing U11, U12, and U5 Snrnps Excises a Minor Class (at-Ac) Intron in Vitro." Cell 84(5): 801-11.
Teigelkamp, S., McGarvey, M., Plumpton, M. and Beggs, J. D. (1994). "The Splicing Factor Prp2, a Putative Rna Helicase, Interacts Directly with Pre-Mrna." Embo J 13(4): 888-97.
Tingley, W. G., Ehlers, M. D., Kameyama, K., Doherty, C., Ptak, J. B., Riley, C. T. and Huganir, R. L. (1997). "Characterization of Protein Kinase a and Protein Kinase C Phosphorylation of the N-Methyl-D-Aspartate Receptor Nr1 Subunit Using Phosphorylation Site-Specific Antibodies." J Biol Chem 272(8): 5157-66.
Valcarcel, J. and Gebauer, F. (1997). "Post-Transcriptional Regulation: The Dawn of Ptb." Curr Biol 7(11): R705-8.
Veraldi, K. L., Arhin, G. K., Martincic, K., Chung-Ganster, L. H., Wilusz, J. and Milcarek, C. (2001). "Hnrnp F Influences Binding of a 64-Kilodalton Subunit of Cleavage Stimulation Factor to Mrna Precursors in Mouse B Cells." Mol Cell Biol 21(4): 1228-38.
Wang, Z. and Grabowski, P. J. (1996). "Cell- and Stage-Specific Splicing Events Resolved in Specialized Neurons of the Rat Cerebellum." Rna 2(12): 1241-53.
Wassarman, K. M. and Steitz, J. A. (1993). "Association with Terminal Exons in Pre-Mrnas: A New Role for the U1 Snrnp?" Genes Dev 7(4): 647-59.
Watakabe, A., Tanaka, K. and Shimura, Y. (1993). "The Role of Exon Sequences in Splice Site Selection." Genes Dev 7(3): 407-18.
Winkler, A., Mahal, B., Zieglgansberger, W. and Spanagel, R. (1999). "Accurate Quantification of the Mrna of Nmdar1 Splice Variants Measured by Competitive Rt-Pcr." Brain Res Brain Res Protoc 4(1): 69-81.

Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. and Smith, C. W. (2004). "Autoregulation of Polypyrimidine Tract Binding Protein by Alternative Splicing Leading to Nonsense-Mediated Decay." Mol Cell 13(1): 91-100.
Wu, J. Y. and Maniatis, T. (1993). "Specific Interactions between Proteins Implicated in Splice Site Selection and Regulated Alternative Splicing." Cell 75(6): 1061-70.
Wu, Q. and Krainer, A. R. (1996). "U1-Mediated Exon Definition Interactions between at-Ac and Gt-Ag Introns." Science 274(5289): 1005-8.
Xu, R., Teng, J. and Cooper, T. A. (1993). "The Cardiac Troponin T Alternative Exon Contains a Novel Purine-Rich Positive Splicing Element." Mol Cell Biol 13(6): 3660-74.
Zhang, L., Ashiya, M., Sherman, T. G. and Grabowski, P. J. (1996). "Essential Nucleotides Direct Neuron-Specific Splicing of Gamma 2 Pre-Mrna." Rna 2(7): 682-98.
Zhang, L., Liu, W. and Grabowski, P. J. (1999). "Coordinate Repression of a Trio of Neuron-Specific Splicing Events by the Splicing Regulator Ptb." Rna 5(1): 117-30.
Zhang, W., Liu, H., Han, K. and Grabowski, P. J. (2002). "Region-Specific Alternative Splicing in the Nervous System: Implications for Regulation by the Rna-Binding Protein Napor." Rna 8(5): 671-85.
Zheng, Z. M., He, P. and Baker, C. C. (1996). "Selection of the Bovine Papillomavirus Type 1 Nucleotide 3225 3' Splice Site Is Regulated through an Exonic Splicing Enhancer and Its Juxtaposed Exonic Splicing Suppressor." J Virol 70(7): 4691-9.
Zhou, Z. and Reed, R. (1998). "Human Homologs of Yeast Prp16 and Prp17 Reveal Conservation of the Mechanism for Catalytic Step Ii of Pre-Mrna Splicing." Embo J 17(7): 2095-106.
Zhou, Z., Licklider, L. J., Gygi, S. P. and Reed, R. (2002). "Comprehensive Proteomic Analysis of the Human Spliceosome." Nature 419(6903): 182-5.
Zhu, J., Mayeda, A. and Krainer, A. R. (2001). "Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound Hnrnp A1 and Enhancer-Bound Sr Proteins." Mol Cell 8(6): 1351-61.
Zhuang, Y. and Weiner, A. M. (1986). "A Compensatory Base Change in U1 Snrna Suppresses a 5' Splice Site Mutation." Cell 46(6): 827-35.
Zuker, M. (1989). "On Finding All Suboptimal Foldings of an Rna Molecule." Science 244(4900): 48-52.
Zukin, R. S. and Bennett, M. V. (1995). "Alternatively Spliced Isoforms of the Nmdari Receptor Subunit." Trends Neurosci 18(7): 306-13.

APPENDIX A. List of primers

A1. Primers for cloning

Primer	Sequence
E21-1	ATATCCTAGAACGTAGGTTCCTCCAAAGACACGGTAAGGGGAAGAGCA CCCCAGTCCCGCGAATTCAATA
E21-2	TATTGAATTCGCGGGACTGGGGTGCTCTTCCCCTTACCGTGTCTTTGGA GGAACCTACGTTCTAGGATAT
E21-3	ATATCCTAGAACGTAGGTTCCTCCAAAGACACGGTAAGTATAAGAGCAC CCCAGTCCCGCGAATTCAATA
E21-4	TATTGAATTCGCGGGACTGGGGTGCTCTTATACTTACCGTGTCTTTGGA GGAACCTACGTTCTAGGATAT
E21-5	ATATCCTAGAACGTAGGTTCCTCCAAAGACACGGTAAGGGGAAGAGCA ATATAGTCCCGCGAATTCAATA
E21-6	TATTGAATTCGCGGGACTATATTGCTCTTCCCCTTACCGTGTCTTTGGA GGAACCTACGTTCTAGGATAT
3021-RI	ATATATATGAATTCGCCCGTAGGAAGCAGATGC
3021-XbaI	ATATTCTAGATGCCCGTAGGAAGCAGATGC
3255A-RI	ATATATATGAATTCCGTCGCGGCAGCACTGTGTC
3096A-XhoI	AAAATTCTCGAGGCTCTTTCTATCCTAAAATGAATATG
3118A-XhoI	AAATAACTCGAGGGTCGGGCTGGCTCTAC
3183-XhoI	AAATAACTCGAGCCTCCAAAGACACGGTAAGGGG
E21XR	CCCTGGCCTCGAGCTTCAAG
E21XF	CTTGAAGCTCGAGGCCAGGG
E21-U164	ATGCGGCCGCACTCCTTGAAAGTCTG
E21-U164-NotI	ATATTGCGGCCGCGGCCTCTTCTTGACCCGGGG
E21-D103	ATTGGATCCAGGGTTGAAGATG
E21-D103-BamHI	ATGGATCCATCTGTGCCAAGTCCTTCACTC

A2. Primers for site-directed mutagenesis (I)

Mutant	Primer	Sequence
5m1	E21-KHH3	CGTAGGTCCTCCAAATATACGGTAAGGGGAAG
	E21-KHH6	CTTCCCCTTACCGTATATTTGGAGGACCTACG
5m2	E21-KHH4	CCTCCAAAGACACGGTAAGCCCAAGAGCACCCC AGTCCC
	E21-KHH7	GGGACTGGGGTGCTCTTGGGCTTACCGTGTCTTT GGAGG
5m3	E21-KHH5	CGTAGGTCCTCCAAATATACGGTAAGCCCAAG
	E21-KHH8	CTTGGGCTTACCGTATATTTGGAGGACCTACG
5m4	E21-KHH15	CCTCCAAAGACACGGTAAGATAAAGAGCACCCCA GTCCC
	E21-KHH16	GGGACTGGGGTGCTCTTTATCTTACCGTGTCTTT GGAGG
5m5	E21-KHH21	GCAGTCTCTGACTCCGGGCTTGCCCCGTGTCTG
	E21-KHH22	CAGACACGGGGCAAGCCCGGAGTCAGTGACTGC
5m6	E21-KHH17	GCAGTCTCTGACTCCTATCTTGCCCCGTGTCTG
	E21-KHH18	CAGACACGGGGCAAGATAGGAGTCAGTGACTGC
5m9	E21-KHH27	CCTCCAAAGACACGGTAAGGCGAAGAGCACCCC AGTCCC
	E21-KHH28	GGGACTGGGGTGCTCTTCCGCTTACCGTGTCTTT GGAGG
5m11	E21-KHH29	GCAGTCTCTGACTCCCGCCTTGCCCCGTGTCTG
	E21-KHH30	CAGACACGGGGCAAGGGCGGAGTCAGTGACTGC
5m12	E21-KHH31	GCAGTCTCTGACTCUCCCCTTGCCCCGTGTCTG
	E21-KHH32	CAGACACGGGGCAAGGGGAGAGTCAGTGACTGC
5m13	E21-KHH33	GTAAGGGGAAGAGCAATACAGTCCCGCGTCTC
	E21-KHH34	GAGACGCGGGACTGTATTGCTCTTCCCCTTAC
5m14	E21-KHH35	GAAGAGCACCCCAGTATAGCGTCTCTGACTCC
	E21-KHH36	GGAGTCAGAGACGCTATACTGGGGTGCTCTTC

A3. Primers for site-directed mutagenesis (II)

Mutant	Primer	Sequence
Wto	E21-KHH11	CTCCACCCTGGCCTCCAGCTTCAAGAGACG
	E21-KHH12	CGTCTCTTGAAGCTGGAGGCCAGGGTGGAG
B'-1	E21-KHH9	CCCTAAAAAGAAAGCCACATTTAATACTATCACCTCCACCCTGGC
	E21-KHH10	GCCAGGGTGGAGGTGATAGTATTAAATGTGGCTTTCTTTTTAGGG
C'-1	E21-KHH19	CTCGAGCTTCAAGACATCGTGGTCCTCCAAAGAC
	E21-KHH20	GTCTTTGGAGGACCACGATGTCTTGAAGCTCGAG
C'-2	E21-KHH23	CCCTGGCCTCGAGCTCATCGTGACGTAGGTCCTCC
	E21-KHH24	GGAGGACCTACGTCACGATGAGCTCGAGGCCAGGG
C'-3	E21-KHH25	GCTTCAAGAGACGTAGACATCGTAAAGACACGGTAAGGGG
	E21-KHH26	CCCCTTACCGTGTCTTTACGATGTCTACGTCTCTTGAAGC
E1	E21-KHH51	GAGTGGTAGAGCAGATATCGACCCTAAAAAGAAAGCCAC
	E21-KHH52	GTGGCTTTCTTTTTAGGGTCGATATCTGCTCTACCACTC
E2	E21-KHH53	GGTAGAGCAGAGCCCGATATTAAAAAGAAAGCCAC
	E21-KHH54	GTGGCTTTCTTTTTAATATCGGGCTCTGCTCTACC
E3	E21-KHH55	GGGCTATCACCTTATACCTGGCCTCCAGC
	E21-KHH56	GCTGGAGGCCAGGTATAAGGTGATAGCCC
E4	E21-KHH57	CACCTCCACCCTGATATCCAGCTTCAAGAG
	E21-KHH58	CTCTTGAAGCTGGATATCAGGGTGGAGGTG
E5	E21-KHH59	CCAGCTTCAAGAGATATAGGTCCTCCAAAGAC
	E21-KHH60	GTCTTTGGAGGACCTATATCTCTTGAAGCTGG
E6	E21-KHH61	CAAGAGACGTAGGTCATCCAAAGACACGG
	E21-KHH62	CCGTGTCTTTGGATGACCTACGTCTCTTG
E7	E21-KHH63	CAAGAGACGTAGGTCTACTAAAGACACGGTAAGGG
	E21-KHH64	CCCTTACCGTGTCTTTAGTAGACCTACGTCTCTTG
E8	E21-KHH65	CCAGCTTCAAGAGACGGTGGTCCTCCAAAGACACG
	E21-KHH66	CGTGTCTTTGGAGGACCACCGTCTCTTGAAGCTGG
E9	E21-KHH67	GCTTCAAGAGACGTAATTCCTCCAAAGACACGG
	E21-KHH68	CCGTGTCTTTGGAGGAATTACGTCTCTTGAAGC
E10	E21-KHH104	GGATAGAAAGAtaGGTAGAGCAGAGCCCG
	E21-KHH105	CGGGCTCTGCTCTACCTATCTTTCTATCC
E11	E21-KHH106	CCTCCACCCTGGCCTtagGCTTCAAGAGACG
	E21-KHH107	CGTCTCTTGAAGCCTAAGGCCAGGGTGGAGG
E12	E21-KHH108	CAAGAGACGTAGGTCCTCCAAAGACACGgtaaggg
	E21-KHH109	CCCTTACCGTGTCTTTGGAGGACCTACGTCTCTTG
E13	E21-KHH110	GAAAGCCACATTatGGGCTATCACCTCCACCC
	E21-KHH111	GGGTGGAGGTGATAGCCCATAATGTGGCTTTC
E20	E21-KHH129	GAGACGGTGGTCCTAGGCAGACACGGTAAGGGG
	E21-KHH130	CCCCTTACCGTGTCTGCCTAGGACCACCGTCTC

A4. Primers for RNA substrates of UV crosslinking

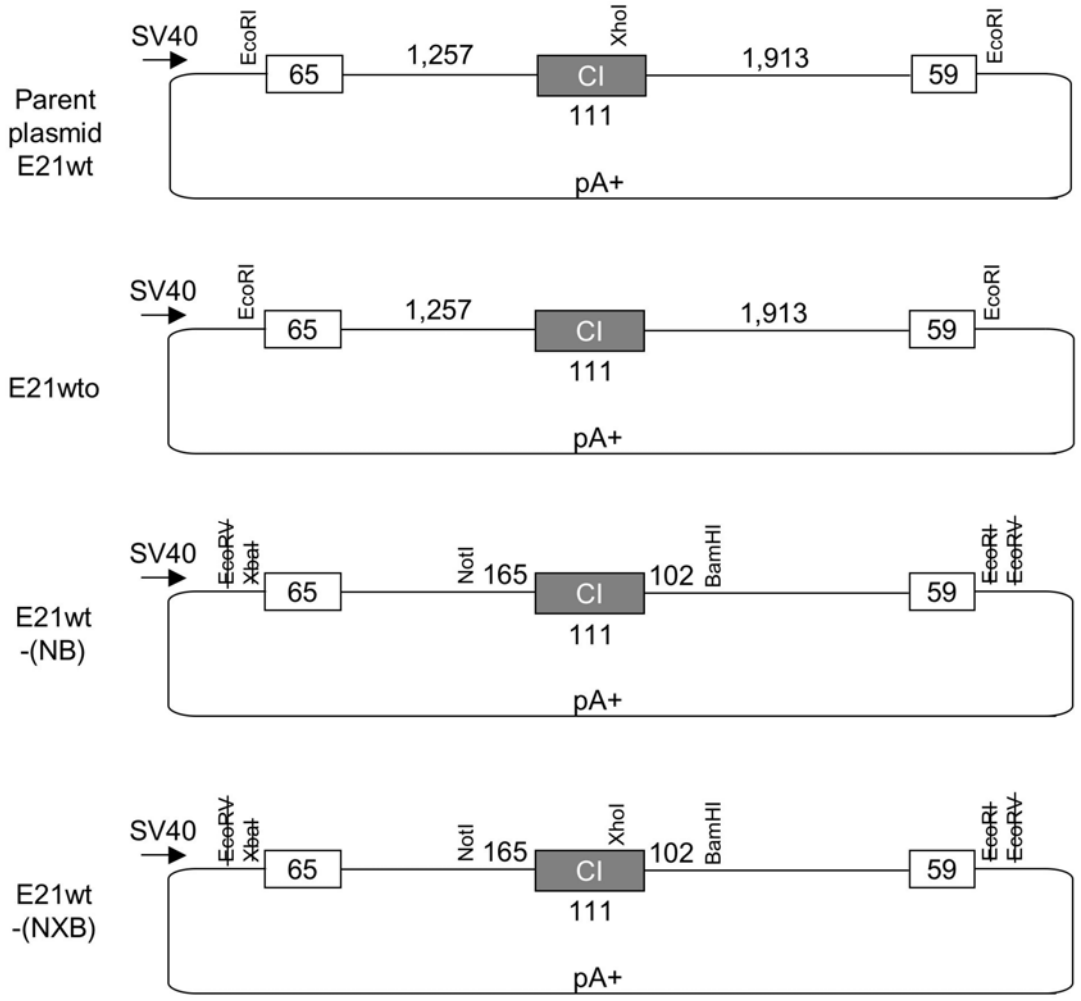
RNA	DNA oligo	DNA sequence
cs1	E21-KHH77	GGATCCTAATACGACTCACTATAGCACGGTAAGGGGAAGAGCA
	E21-KHH78	TGCTCTTCCCCTTACCGTGCTATAGTGAGTCGTATTAGGATCC
cs3	E21-KHH81	GGATCCTAATACGACTCACTATAGCACGGTAAGGCGAAGAGCA
	E21-KHH82	TGCTCTTCGCCTTACCGTGCTATAGTGAGTCGTATTAGGATCC
3h1	E21-KHH47	TAATACGACTCACTATAGGGAAGAGCACCCCAGTCCCGCGTCTCTGA
	E21-KHH48	TCAGAGACGCGGGACTGGGGTGCTCTTCCCTATAGTGAGTCGTATTA
3h3	E21-KHH87	TAATACGACTCACTATAGCGAAGAGCACCCCAGTCCCGCGTCTCTGA
	E21-KHH88	TCAGAGACGCGGGACTGGGGTGCTCTTCGCTATAGTGAGTCGTATTA
5'RNA-w1	E21-KHH45	TAATACGACTCACTATAAGAGACGUAGGUCCUCCAAAGACACGGUAA
	E21-KHH46	TTUCCGTGTCTTTGGUGGUCCTUCGTCTCTTATAGTGAGTCGTATTA
5'RNA-m1	E21-KHH47	TAATACGACTCACTATAGGGAAGAGCACCCCAGUCCCGCGUCUCUGA
	E21-KHH48	TCUGUGUCGCGGGUCTGGGGTGCTCTTCCCTATAGTGAGTCGTATTA
3'RNA	E21-KHH87	TAATACGACTCACTATAGCGAAGAGCACCCCAGUCCCGCGUCUCUGA
	E21-KHH88	TCUGUGUCGCGGGUCTGGGGTGCTCTTCGCTATAGTGAGTCGTATTA
splint w1	E21-KHH49	UGGGGUGCUCUUCCCCUUACCGUGUCUUUG
splint m1	E21-KHH89	UGGGGUGCUCUUCGCCUUACCGUGUCUUUG
5'RNA-w2	E21-KHH69	TAATACGACTCACTATAGACGUAGGUCCUCCAAAGACACGGUAAGGG
	E21-KHH70	CCCTTUCCGTGTCTTTGGUGGUCCTUCGTCTATAGTGAGTCGTATTA
5'RNA-m2	E21-KHH71	TAATACGACTCACTATAGACGUAGGUCCUCCAAAGACACGGUAAGCC
	E21-KHH72	GGCTTUCCGTGTCTTTGGUGGUCCTUCGTCTATAGTGAGTCGTATTA
5'RNA-m21	E21-KHH83	TAATACGACTCACTATAGACGGUGGUCCUCCAAAGACACGGUAAGGG
	E21-KHH84	CCCTTUCCGTGTCTTTGGUGGUCCUCCGTCTATAGTGAGTCGTATTA
3'RNA	E21-KHH73	TAATACGACTCACTATAGAAGAGCACCCCAGUCCCGCGUCUCUGA
	E21-KHH74	TCUGUGUCGCGGGUCTGGGGTGCTCTTCTATAGTGAGTCGTATTA
splint m2	E21-KHH75	UGGGGUGCUCUUCGGCUUACCGUGUCUUUC
5'RNA-w3	E21-KHH96	TAATACGACTCACTATAGGCCTCCAGCTTCAAGAGACGTA
	E21-KHH97	TACGTCTCTTGAAGCTGGAGGCCTATAGTGAGTCGTATTA
5'RNA-m3	E21-KHH98	TAATACGACTCACTATAGGCCTCCAGCTTCAAGAGACGGT
	E21-KHH99	ACCGTCTCTTGAAGCTGGAGGCCTATAGTGAGTCGTATTA
3'RNA	E21-KHH100	TAATACGACTCACTATAGGTCCTCCAAAGACACGGTAAGGGGAAGAGC ACCCCA
	E21-KHH101	TGGGGTGCTCTTCCCCTTACCGTGTCTTTGGAGGACCTATAGTGAGTC GTATTA
splint wt3	E21-KHH102	TGTCTTTGGAGGACCTACGTCTCTTGAAGC
splint mt3	E21-KHH103	TGTCTTTGGAGGACCACCGTCTCTTGAAGC

A5. Primers for RT-PCR

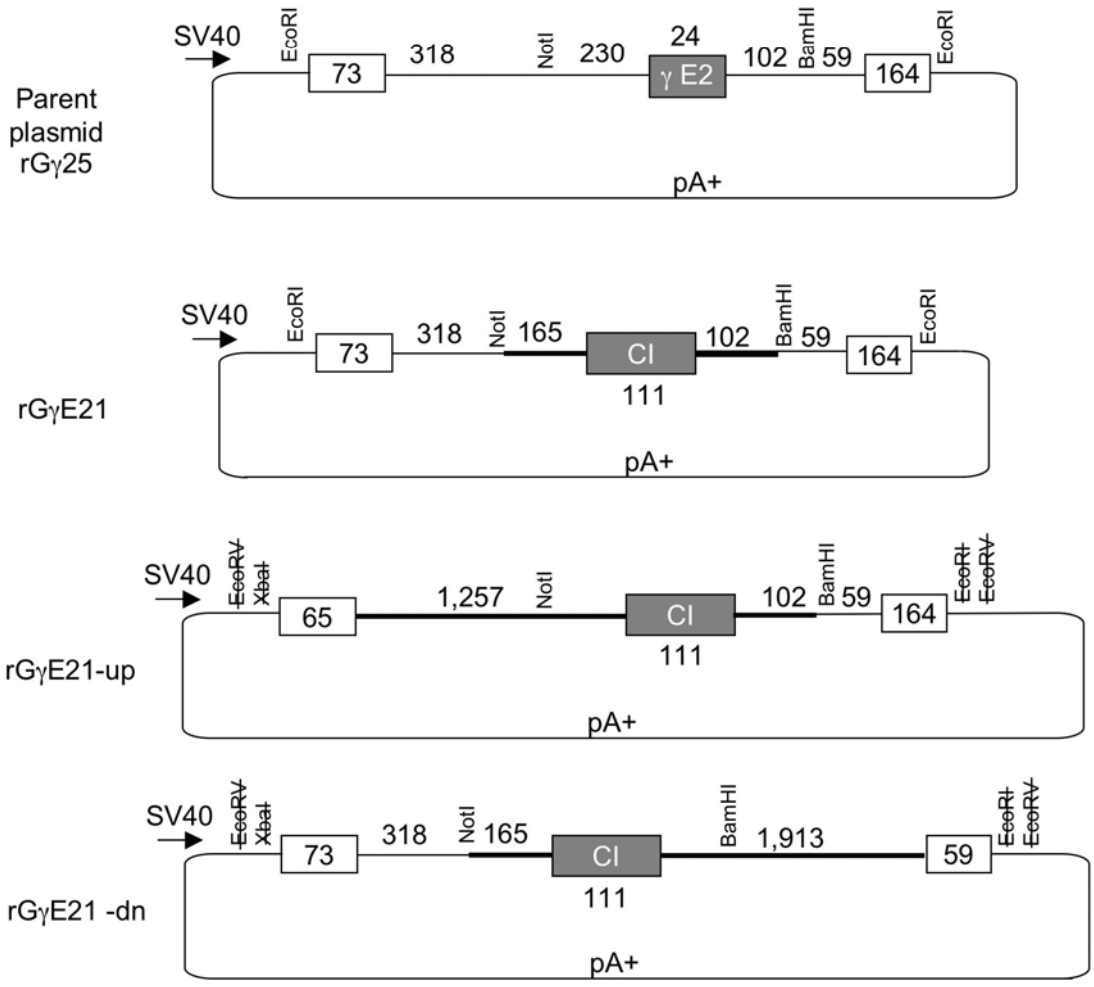
Gene ID and exon #	Primer name	Primer sequence
GRIN1 CI exon	NMDA 3021	ATGCCCGTAGGAAGCAGATGC
	NMDA 3255A	CGTCGCGGCAGCACTGTGTC
γ2 exon 2	rp1	GTATGGTACCCTGCACTATTTTGTG
	rp2	ACCAGTTACTCTGCCATC
ACTB exon 2	mb-actin1	AGCTTCTTTGCAGCTCCTTC
	mb-actin2	TCTTCTCCATGTCGTCCCAG
HNRPH1 exon 8	E21-KHH126	GCTATGGAGGCTATGATG
	E21-KHH127	ACCAGTTACTCTGCCATC
HNRPH1 exon 5	E21-KHH116	GGCTTTGTACGGCTTAGAG
	E21-KHH117	CCTGGCCGCTGCATGGCC
HNRPH3 exon 2	E21-KHH118	AATGGTCCAAATGACGCTAG
	E21-KHH119	CCAGCAATCTTCTTGGTGG
Hp1bp3 exon 2	Hp1bp3-mE1	TACGGGAACGCGGGCG
	Hp1bp3-E3R	TTCCTCCCCTTCAGCAAG
GRIN1 exon 21	NMDA3021	ATGCCCGTAGGAAGCAGATGC
	NMDA3255A	CGTCGCGGCAGCACTGTGTC
NCOA2 exon 13	NCOA2-E12	CAAGCCATCATCAATGACCTC
	NCOA2-mE14R	CTCACAGCTGGACTCTGTGG
MEN1 exon 8	men1-mE7	TCTTTGAAGTGGCCAATGAC
	men1-mE9R	CTGCTTCCCGGCTAACTATG
ABCA4 exon 10	ABCA4-E9	GATCCAGAGCCTGGAGTC
	ABCA4-E11R	AATATGTCCCTCCAGTC
MEN1 exon 8	#3-E20	GCTGAAAGGAGCTTCAGTG
	#3-E22R	AGAGGCTGACTTGGGCAC

APPENDIX B. Plasmid maps

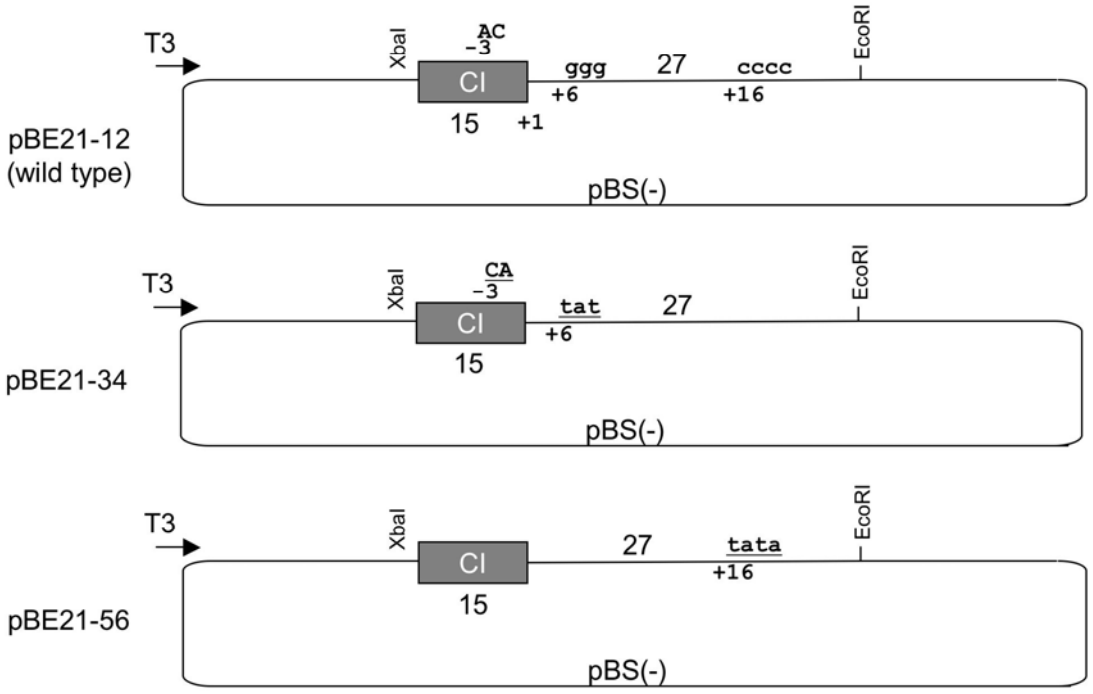
B1. Splicing reporters: wild type derivatives



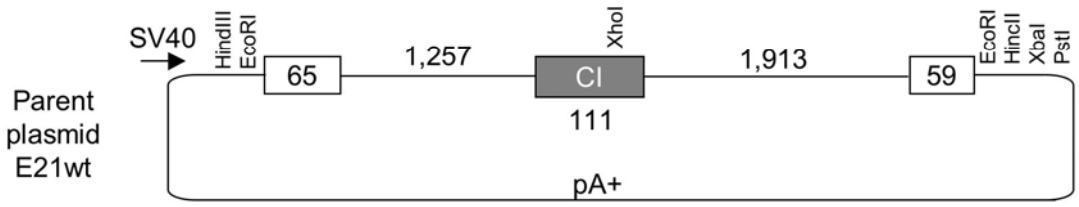
B2. Splicing reporters: chimeric plasmids



B3. Transcription templates: biotin affinity selection



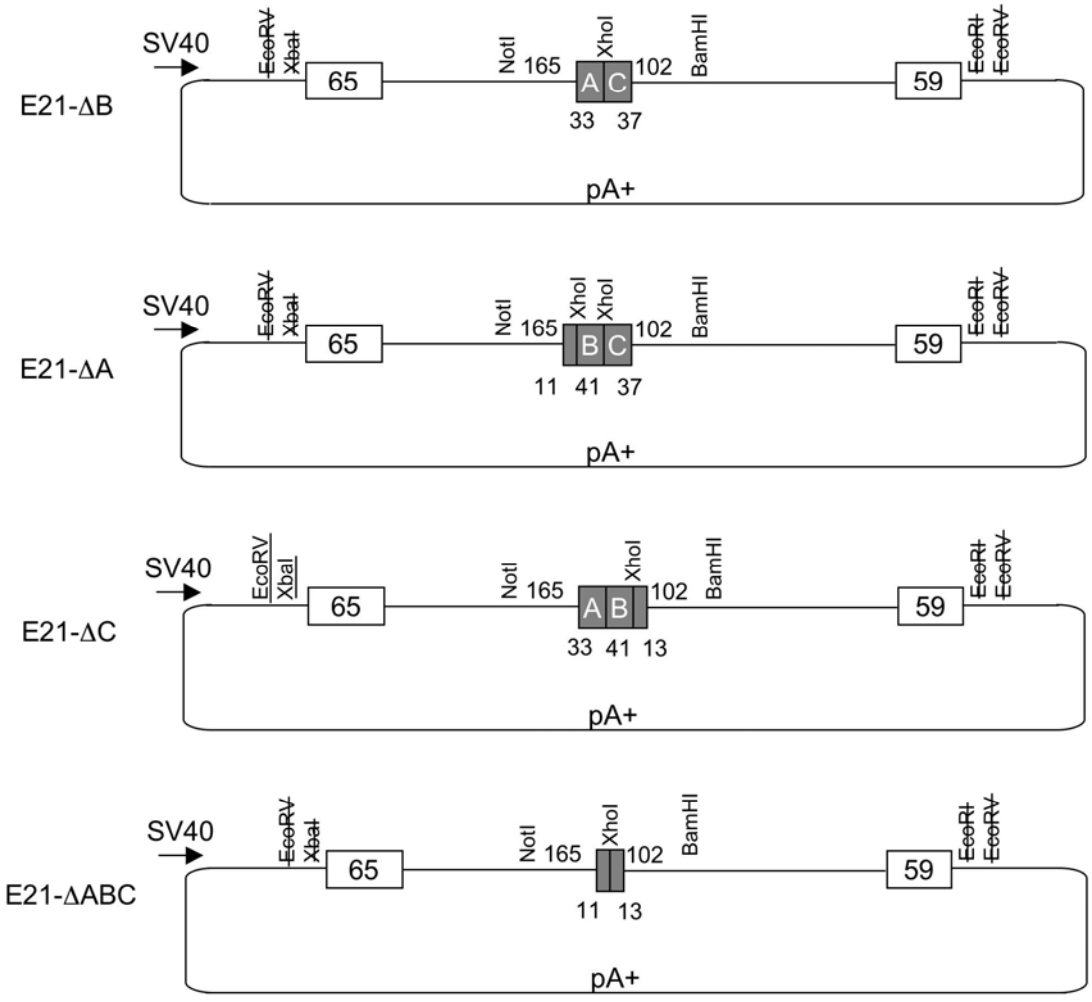
B4. Splicing reporters: 5' splice site mutants



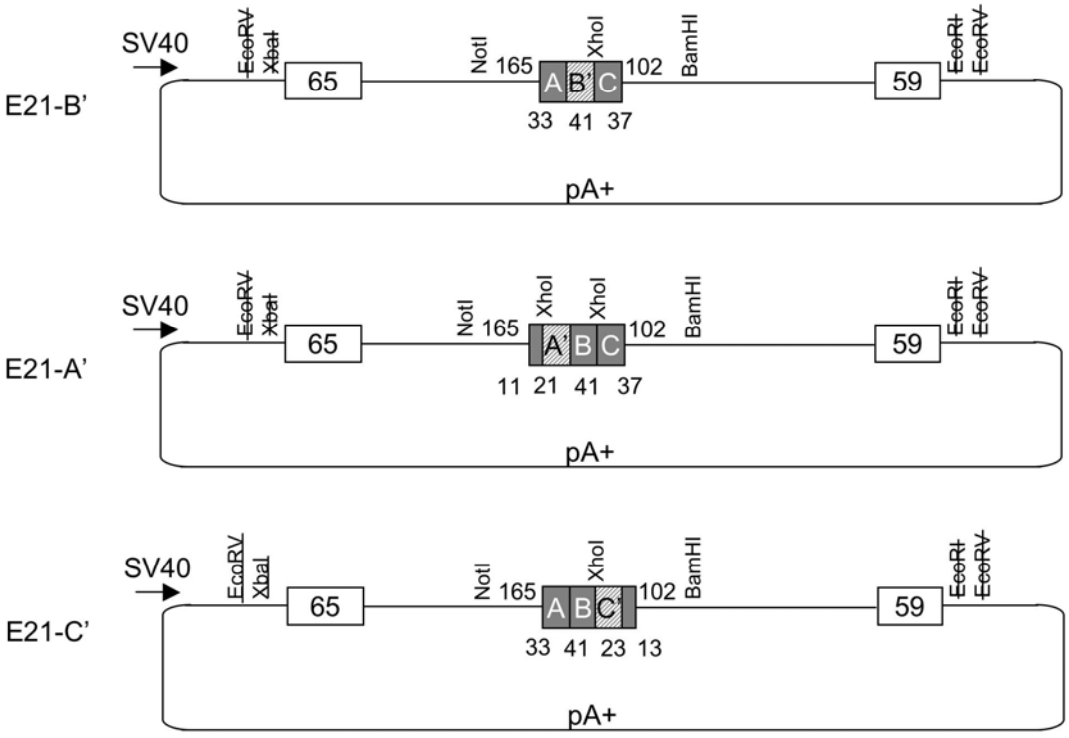
Sequence change of mutants

Wt	AAA**G**A**C**ACG/guaag**ggg**aagagcaccccagucccgcgucucugacucc**ucc**cuug
5m1	AAA**U**A**U**ACG/guaaggggaagagcaccccagucccgcgucucugacuccucccuug
5m2	AAAGACACG/guaag**ccc**aagagcaccccagucccgcgucucugacuccucccuug
5m3	AAA**U**A**U**ACG/guaag**ccc**aagagcaccccagucccgcgucucugacuccucccuug
5m4	AAAGACACG/guaag**aua**aagagcaccccagucccgcgucucugacuccucccuug
5m5	AAAGACACG/guaag**ccc**aagagcaccccagucccgcgucucugacucc**ggg**cuug
6m6	AAAGACACG/guaag**aua**aagagcaccccagucccgcgucucugacucc**uau**cuug
5m7	AAAGACACG/guaag**aua**aagagcaccccagucccgcgucucugacucc**ggg**cuug
5m8	AAAGACACG/guaaggggaagagcaccccagucccgcgucucugacucc**ggg**cuug
5m9	AAAGACACG/guaag**gcg**aagagcaccccagucccgcgucucugacuccucccuug
5m10	AAAGACACG/guaag**gcg**aagagcaccccagucccgcgucucugacucc**cgc**cuug
5m11	AAAGACACG/guaaggggaagagcaccccagucccgcgucucugacucc**cgc**cuug
5m12	AAAGACACG/guaaggggaagagcaccccagucccgcgucucugacuc**ucc**cuug
5m13	AAAGACACG/guaaggggaagagcac**aua**agucccgcgucucugacuccucccuug
5m14	AAAGACACG/guaaggggaagagcaccccagu**aua**gcgucucugacuccucccuug

B5. Splicing reporters: deletion mutants



B6. Splicing reporters: shuffled mutants



> mutated sequences (underlined)

A region

E21-wt GGTAGAGCAGAGCCCGACCCT
E21-A' GCGTCGCACGATAGCATCGGC

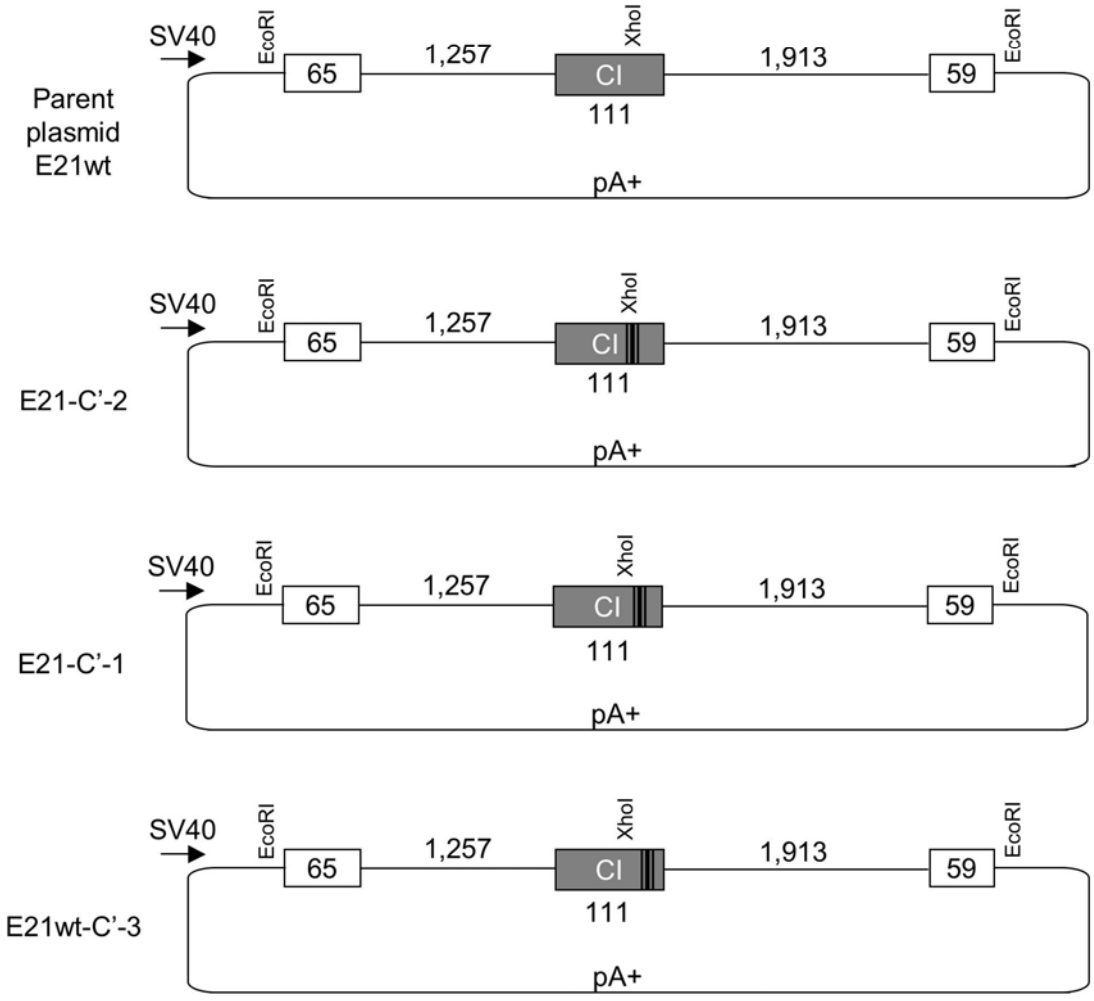
B region

E21-wt(B) AAAAAGAAAGCCACATTTAGGGCTATCACCTCCACCCTGG
E21-B' GTACTACAGATACACACGCATGTGACTACACACAGCTGCC

C region

E21-wt(C) CTCGAGCTTCAAGAGACGTAGGT
E21-C' CTCGAGCAGCATGTGGTAGTAAC

B7. Splicing reporters: replacement mutants (hexamer linker)



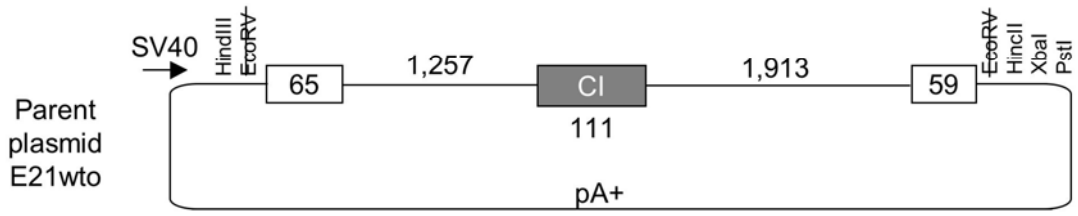
> mutated sequences (underlined)

C'-2 **TCAAGA → CATCGT (position 82)**

C'-1 **AGACGT → CATCGT (position 88)**

C'-3 **GTCCTCC → ACATCGT (position 96)**

B8. Splicing reporters: ESE mutants



> Sequence change of mutants

clone	sequence	template plasmid
E21-wt	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCGAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-up
E21-wto*	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACAC G	E21-wto
E21-E1	GAUAGAAAGAGUGGUAGAGCAGAUAUCGACCCUAAAAAGAAAGCCACAUUUAGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E2	GAUAGAAAGAGUGGUAGAGCAGAGCCCGAUAUUAAAAAGAAAGCCACAUUUAGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E3	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCAUACCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E4	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGAUAUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E5	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGAUAUAGGUCCUCCAAAGACACG	E21-wto
E21-E6	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCAUCCAAAGACACG	E21-wto
E21-E7	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCUACUAAAGACACG	E21-wto
E21-E8	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-wto
E21-E9	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAAUUCCUCCAAAGACACG	E21-wto
E21-E10	GAUAGAAAGAUAGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-E8
E21-E11	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUUAGGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-E8
E21-E12	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUAGGAAGACACG	E21-E8

*The sequence is the original CI cassette sequence, named as wto.

Continued.. Sequence change of ESE mutants

clone	sequence	template plasmid
E21-E13	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E14	GAUAGAAAGAUAGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-E10
E21-E15	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUUAGGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-E11
E21-E16	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUAGGAAGACACG	E21-E12
E21-E17	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUCCAAAGACACG	E21-E8
E21-E18	GAUAGAAAGAUAGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACACG	E21-wto
E21-E20	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUAGGCAGACACG	E21-E12
E21-E21	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUAUGGGC UAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUAGGCAGACACG	E21-E16
E21-T8	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACAC G GUAAGAUA	E21-D8
E21-D0	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGUAGGUCCUCCAAAGACAC G GUAAGAUA	E21-wto
E21-D8	GAUAGAAAGAGUGGUAGAGCAGAGCCCGACCCUAAAAAGAAAGCCACAUUUAGGG CUAUCACCUCCACCCUGGCCUCCAGCUUCAAGAGACGGUGGUCCUCCAAAGACACG GUAAGAUA	E21-D0

B9. Splicing reporters: Overexpression