SPLICING SILENCING BY THE CUGBP2 SPLICING FACTOR: MECHANISM OF ACTION AND COMBINATORIAL CODE FOR SPLICING SILENCING WITH IMPLICATIONS FOR AUTOREGULATION

by

Jill A. Dembowski

Bachelor of Science, Clarion University of Pennsylvania, 2004

Submitted to the Graduate Faculty of

the School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Jill A. Dembowski

It was defended on

September 17, 2009

and approved by

Dr. Jeffrey L. Brodsky, Professor, Department of Biological Sciences

Dr. Craig L. Peebles, Professor, Department of Biological Sciences

Dr. James M. Pipas, Professor, Department of Biological Sciences

Dr. A. Javier Lopez, Associate Professor, Department of Biological Sciences, Carnegie Mellon University

Thesis Advisor: Dr. Paula J. Grabowski, Professor, Department of Biological Sciences

Much of this work was published in PLoS Genetics (Dembowski and Grabowski, 2009) under the terms of the Creative Commons Attribution License, which allows anyone to download, reuse, reprint, modify, distribute, and/or copy articles in PLoS journals, so long as the original authors and sources are cited.

SPLICING SILENCING BY THE CUGBP2 SPLICING FACTOR: MECHANISM OF ACTION AND COMBINATORIAL CODE FOR SPLICING SILENCING WITH IMPLICATIONS FOR AUTOREGULATION

Jill A. Dembowski, PhD

University of Pittsburgh, 2009

Alternative pre-mRNA splicing adjusts the transcriptional output of the genome by generating related mRNAs from a single primary transcript, thereby expanding protein diversity. A fundamental unanswered question is how splicing factors achieve specificity in the selection of target substrates despite the recognition of information-poor sequence motifs. The CUGBP2 splicing regulator plays a key role in the brain region-specific silencing of the NI exon of the NMDA R1 receptor. However, the sequence motifs utilized by this factor for specific target exon selection and the mechanism of splicing silencing are not understood. Here, I use chemical modification footprinting to map the contact sites of CUGBP2 to GU-rich motifs closely positioned at the boundaries of the branch sites of the NI exon, and demonstrate a mechanistic role for this specific arrangement of motifs for the regulation of branchpoint formation. General support for a branch site-perimeter-binding model is indicated by the identification of a group of novel target exons with a similar configuration of motifs that are silenced by CUGBP2. These results reveal an autoregulatory role for CUGBP2 as indicated by its direct interaction with functionally significant RNA motifs surrounding the branch sites upstream of exon 6 of the CUGBP2 transcript, itself. The perimeter-binding model explains how CUGBP2 can effectively embrace the branch site region to achieve the specificity needed for the selection of exon targets and the fine-tuning of alternative splicing patterns.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would sincerely like to thank my mentor and role model, Paula Grabowski, for her guidance and continued support throughout my graduate career, my committee members for their insightful advice, and the Grabowski lab members past and present for thoughtful discussions and assistance with experiments. I would especially like to thank my family for their encouragement and support.

1.0 INTRODUCTION

Maturation of many eukaryotic pre-mRNAs involves the removal of intervening intron sequences and the splicing of exon sequences to generate a mature mRNA transcript. The process of splicing involves a two-step chemical mechanism and step-wise assembly of the spliceosome, the macromolecular machinery that catalyzes the splicing reaction. Splicing is a precisely controlled process but requires a great amount of inherent flexibility. An individual exon can be included or skipped in the mature mRNA by the process of alternative splicing. The strength of individual splice sites as well as the effects of trans-acting factors interacting with cis-regulatory RNA elements adjust spliceosome assembly to fine-tune splicing. Splicing factors shape alternative splicing patterns to adjust modular protein properties in a tissue- and developmental-specific manner. In this way, alternative splicing greatly expands protein diversity given a fixed number of genes. Splicing patterns are sensitive to environmental cues and in the nervous system respond to extracellular stimuli such as cell depolarization. Although alternative splicing is fundamental to diverse cellular processes such as cell differentiation, development, gene expression, and maintenance of cellular homeostasis, we are only just beginning to understand the mechanisms of regulation.

1.1 PRE-MRNA SPLICING: CIS-RNA ELEMENTS AND CHEMICAL MECHANISM

Three RNA elements are integral to the two-step chemical mechanism of splicing. These include the 5′ splice site, the 3′ splice site, and the branch site (Figure 1). These RNA elements are recognized by the splicing machinery to define the exon to be spliced and to guide intron removal and exon joining. Several studies have focused on identifying the consensus nucleotide content of each of these sites. Often times these cis-RNA elements are targets of alternative splicing control and are altered in intronic small nuclear polymorphisms (SNPs) that contribute to splicing defects associated with human disease.

1.1.1 Evolutionarily conserved cis-RNA elements

For the major class of U2 introns, the metazoan 5′ splice site has historically been classified by the consensus sequence, (C/A)AG:GU(A/G)AGU (colon, exon-intron junction) (Mount, 1982). Here the first GU dinucleotide is invariant and the G at position 8 is highly conserved. U1 small nuclear RNA (snRNA) of the U1 small nuclear ribonucleoprotein particle (snRNP) base pairs with this RNA element to define the 3′ end of the exon to be spliced (Zhuang and Weiner, 1986). The strength of an individual 5′ splice site is therefore a direct reflection of its match to the consensus sequence and ability to base pair with U1 snRNA. Recently it has been demonstrated that a subset of atypical 5′ splice sites contain the consensus sequence, ACA:GTTAAG, which is recognized by U1 snRNA in an atypical register, shifted by one nucleotide (Roca and Krainer, 2009). This observation suggests that splice site recognition is more flexible than previously thought. Weak 5′ splice sites are generally associated with exon skipping and alternative splicing control (Kuo et al., 1991). 5′ splice site mutations that decrease splice site strength can cause exon skipping or activate cryptic 5′ splice sites causing intron

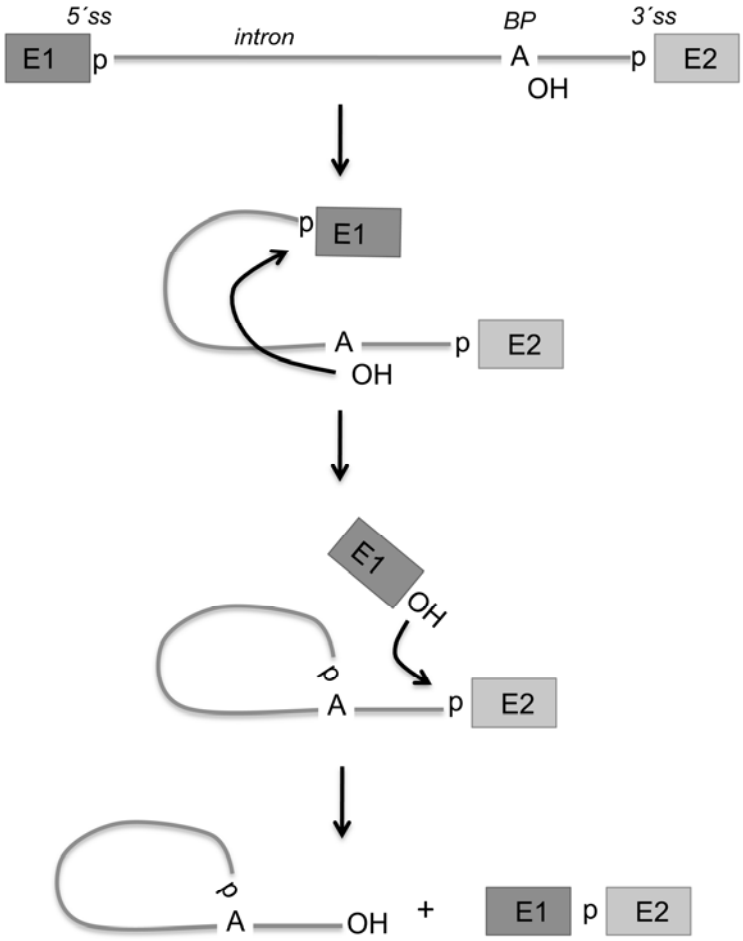


Figure 1. Cis-acting RNA elements are intimately involved in the two-step chemical mechanism of pre-mRNA splicing. A generic pre-mRNA containing two exons (rectangles: E1, E2) flanking an intron sequence (line) with the positions of the 5´ splice site (ss), 3´ ss, and branchpoint (BP) adenosine (A) is shown (top). The phosphate (p) and hydroxyl (OH) groups involved in catalysis of splicing are also indicated. For splicing, two consecutive chemical rearrangements occur resulting in the formation of the spliced E1-E2 and intron lariat products (bottom).

retention (Treisman et al., 1983; O'Neill et al., 1998). Furthermore, mutations that increase 5′ splice site strength have been associated with increased exon inclusion and human disease (Hutton, 2001). In yeast, the 5′ splice site is less degenerate and matches the consensus, AAG:GUAUGU (Long et al., 1997).

The metazoan 3′ splice site consensus sequence is Y_nNYAG:G (Y, pyrimidine; N, any nucleotide; colon, intron-exon junction) (Mount, 1982). Y_n is referred to as the polypyrimidine-tract and the AG dinucleotide at the intron-exon junction is invariant. In fact, 98% of mammalian U2 introns contain a GU at the 5′ splice site and an AG at the 3′ splice site (Sheth et al., 2006). Mutation of the AG dinucleotide at the 3′ splice site can activate cryptic 3′ splice sites, resulting in intron retention or partial exon exclusion, often resulting in a shift in the reading frame (O'Neill et al., 1998). The yeast 3′ splice site is also less degenerate in that the polypyrimidine-tract contains mostly Us rather than Us and Cs as for the metazoan 3′ splice site (Long et al., 1997).

The branch site resides just upstream of the polypyrimidine-tract and has been historically defined by the consensus sequence, UACUAAC (Zhuang et al., 1989). Recently, it has been demonstrated that the branch site consensus is more degenerate than previously thought, YUNAY (Y, pyrimidine; N, any nucleotide; A, the branchpoint) (Gao et al., 2008). The branch site was first identified in yeast where the invariant UACUAAC sequence is recognized (Langford and Gallwitz, 1983). U2 snRNA base pairs with the branch site region to specify the branchpoint adenosine as discussed below. Furthermore, the branch site is recognized sequentially by several factors throughout spliceosome assembly (MacMillan et al., 1994). Mutations at the branch site have been associated with activation of cryptic branchpoints (Reed and Maniatis, 1985) or exon skipping associated with human disease (Kralovicova et al., 2004). It has also been shown that the strength of the 5′ splice site partly controls branch site (Grabowski et al., 1991) and 3′ splice site selection (Hoffman and Grabowski, 1992) across the exon.

The degenerate nature of these splice sites, as well as the large size of mammalian introns (~1000 nts) makes it difficult to accurately predict all exons of a gene by sequence inspection alone. Furthermore, additional RNA elements, including splicing enhancers, are often necessary for precise exon inclusion. Mutations that alter splice site strength change splicing patterns and often manifest in disease states (Novoyatleva et al., 2006). Moreover, relaxation of splice site strength may have provided a platform for the evolution of alternative splicing events (Ast, 2004; Kol et al., 2005).

1.1.2 Chemical mechanism of pre-mRNA splicing

Intervening intron sequences and the concept of "splicing" were originally described for adenovirus 2 (Ad2) and simian virus 40 (SV40) precursor mRNAs (Berget et al., 1977; Chow et al., 1977; Berk and Sharp, 1978a, b). Chemical mechanisms of splicing were initially dissected using the enzymatically spliced yeast pre-tRNA and the Tetrahymena self-splicing group I introns of pre-rRNA as model systems (Cech et al., 1981; Grabowski et al., 1981; Kruger et al., 1982; Greer et al., 1983; Peebles et al., 1983). Later, splicing of adenovirus pre-mRNA was studied in splicing active HeLa nuclear extracts (Padgett et al., 1983), which led to the discovery of the spliceosome. Pre-mRNA splicing follows a two-step chemical mechanism as described in Figure 1. First, the free 2´ hydroxyl group at the branchpoint adenosine attacks the phosphodiester bond that joins E1 to the intron, resulting in the formation of the lariat intron-E2 intermediate and a free hydroxyl group at the 3´ end of E1 (Grabowski et al., 1981; Padgett et al., 1984; Ruskin et al., 1984; Konarska et al., 1985). Next, the free hydroxyl group at E1 attacks the phosphodiester bond at the 3´ splice site that joins the intron to E2, resulting in the formation of the spliced E1-E2 and intron lariat products. This transesterification reaction is evolutionarily conserved across the self-splicing group II introns of Tetrahymena and the spliceosome-regulated introns of yeast and metazoan pre-mRNA. Tetrahymena self-splicing group I introns

follow a similar chemical mechanism, however, guanosine is used as the branchpoint nucleophile rather than adenosine. Taken together, the chemical mechanism of splicing is highly conserved and can be catalyzed by evolutionarily diverged processes. For the purpose of this study, I will focus on spliceosome-mediated splicing.

1.2 THE SPLICEOSOME

The spliceosome is the macromolecular machinery that carries out the chemical mechanism of pre-mRNA splicing. The spliceosome was simultaneously identified in yeast as a 40S complex intimately associated with pre-mRNAs undergoing splicing (Brody and Abelson, 1985) and as a 60S complex assembled on Ad2 pre-mRNA in splicing active HeLa nuclear extracts (Grabowski et al., 1985). Purification of spliceosomal and pre-spliceosomal complexes and intermediates has revealed the vast nature of its RNA and protein components (Grabowski and Sharp, 1986; Konarska and Sharp, 1987; Bennett et al., 1992). The spliceosome consists of U1, U2, U4, U5, and U6 snRNPs and hundreds of accessory factors (reviewed by (Padgett et al., 1986; Luhrmann et al., 1990; Reed, 1996)). Individual snRNPs are formed from the association of both common and unique snRNP associated proteins (SAPs) with U1, U2, U4, U5, and U6 snRNAs.

The spliceosome is integral to the splicing process and shares several common features with the ribosome (reviewed by (Staley and Guthrie, 1998; Staley and Woolford, 2009)). They are both catalytically active ribonucleoprotein (RNP) machines consisting of dynamic RNA-RNA, RNA-protein, and protein-protein interactions. They are both highly efficient and accurate and able to accommodate diverse RNA substrates with proofreading ability. Furthermore, step-wise assembly of the spliceosome and ribosome involve RNA cleavage and ligation reactions, RNA and protein rearrangements, and hundreds of accessory factors. These accessory factors

include RNA binding proteins that stabilize RNA-RNA interactions, helicases that function to reconfigure RNA-RNA and RNA-protein interactions, and GTPases and ATPases that function to generate energy needed for conformational transitions. Spliceosome assembly occurs by the step-wise association of snRNPs and accessory factors and proceeds through complexes H, E, A, B, and C as discussed below and illustrated in Figure 2. Although there is a parallel splicing pathway in yeast, the metazoan spliceosome is discussed in detail below.

1.2.1 Complex H: non-specific RNA binding

Immediately upon addition of pre-mRNA to an *in vitro* splicing reaction, complex H forms. This complex does not depend on ATP or functional 5´ and 3´ splice sites (Konarska and Sharp, 1986). It is associated with non-specific binding by mostly heterogeneous nuclear ribonucleoproteins (hnRNPs) and is not considered to be part of the functional spliceosome (Michaud and Reed, 1991). Unique sets of hnRNP proteins associate with individual pre-mRNAs, which likely depends on the sequence content and size of the pre-mRNA tested (Bennett et al., 1992). Importantly, complex H is not resistant to high salt (250 mM) like the spliceosome (Reed, 1990).

1.2.2 Complex E: splice site recognition and exon definition

Early (E) complex is the first committed step in the splicing process and forms independently of ATP (Michaud and Reed, 1991). Here, the 5´ splice site is recognized and defined by U1 snRNP (Michaud and Reed, 1991) by direct nucleotide interactions between the 5´ end of U1 snRNA and the pre-mRNA (Zhuang and Weiner, 1986). U2 snRNP auxiliary factor (U2AF) recognizes and defines the 3´ splice site (Reed, 1990). U2AF is composed of a 65 kDa subunit, $U2AF^{65}$, and a 35 kDa subunit, $U2AF^{35}$. $U2AF^{65}$ binds to the polypyrimidine-tract through direct

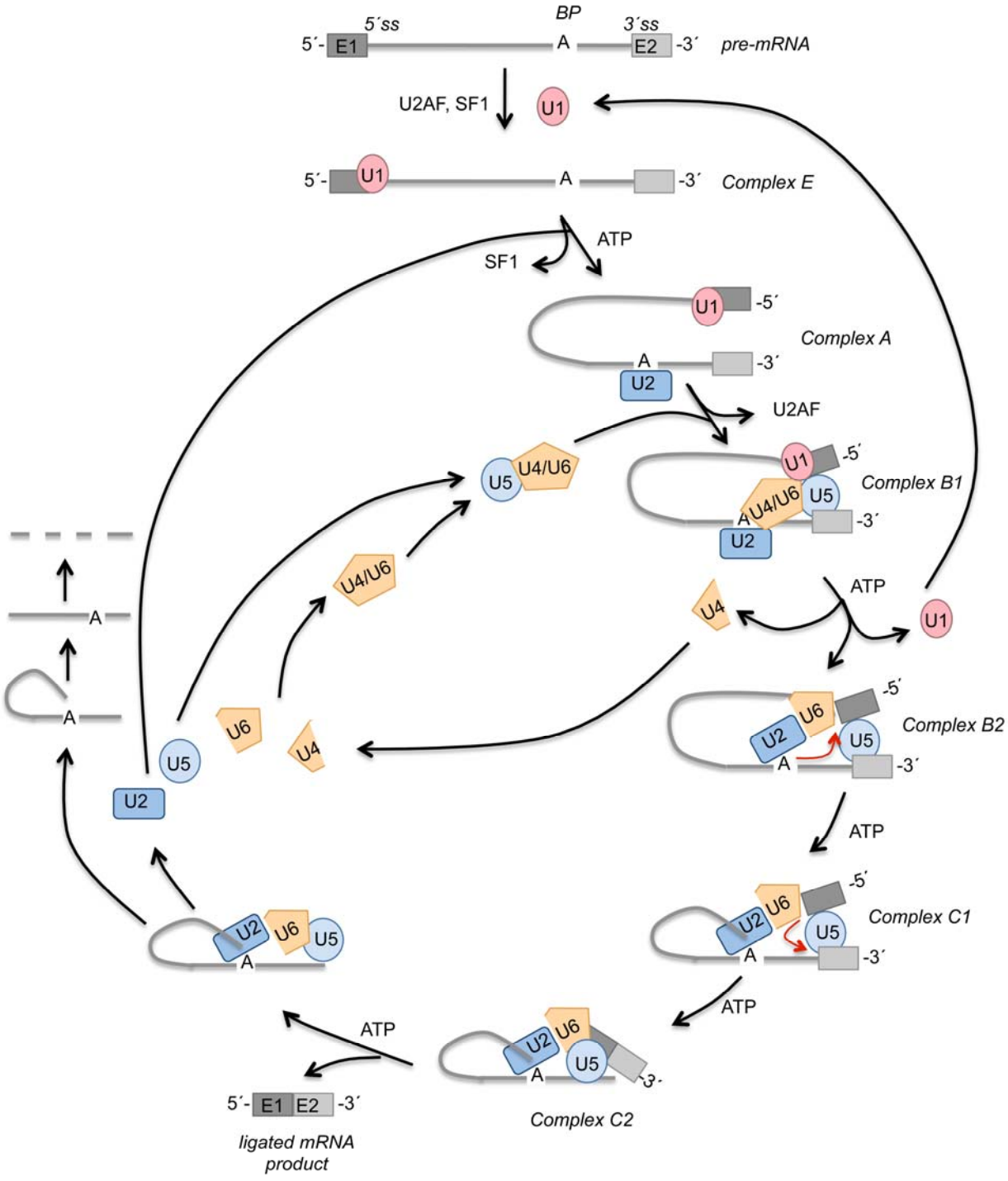


Figure 2. Step-wise assembly of the spliceosome guides RNA rearrangements and catalysis for precise intron removal and exon joining. A generic pre-mRNA containing two exons (E1, E2) flanking an intron of variable length is shown in the 5´ to 3´ direction (top). The

relative positions of the 5′ splice site (ss), branchpoint (BP) adenosine (A), and 3′ ss are indicated. For Complex E assembly, U1 snRNP, U2AF, and SF1 bind to the 5′ splice site, 3′ splice site, and branch site, respectively. In Complex A, U2 snRNP replaces SF1 at the branch site to designate the adenosine to be used as the branchpoint. Complex B1 involves the association of U4:U6 and U5 snRNPs and removal of U2AF. In Complex B2, U1 and U4 snRNP leave, the spliceosome rearranges, the U2:U6 snRNP complex forms, and the branchpoint adenosine attacks the 5′ splice site of the upstream exon. In Complex C1, the 5′ splice site of the upstream exon attacks the 3′ splice site of the downstream exon for the formation of the lariat intron intermediate and ligated mRNA product in Complex C2. After splicing, the lariat intron intermediate is debranched and degraded and all the components of the spliceosome are recycled. ATP-dependent steps are indicated.

interactions with its RNA binding domain (RBD) (Zamore et al., 1992) and $U2AF^{35}$ associates with the 3′ splice site through interactions with $U2AF^{65}$ (Zhang et al., 1992). Splicing factor 1 (SF1) makes direct contacts with the branch site region and buries the branchpoint adenosine in a hydrophobic cleft (Abovich and Rosbash, 1997; Liu et al., 2001). The U1 snRNP-specific proteins U1 A, U1 C, and U1 70K, the snRNP core proteins B and B′, and the 115, 92, 88, 72, and 42 kDa SAPs also associate with the pre-mRNA during complex E assembly (Reed, 1990). Furthermore, members of the serine and arginine (SR)-rich family of protein factors, which are characterized by a domain of repeating serine/arginine dipeptides (termed RS domain), also associate during this step. SR factors are important for mediating protein-protein interactions through their RS domains and have essential roles in general splicing as well as diverse roles in specifying alternative splicing patterns (Fu, 1995).

Binding by U1 snRNP and U2AF to the 5′ and 3′ splice sites, respectively, defines the exon as the unit of RNA to be spliced as indicated by the exon definition model first described by Robberson et al. (1990). In support of this model, SR protein factors interact across the exon to promote exon definition (Kohtz et al., 1994; Staknis and Reed, 1994a). For example, the SR protein, ASF/SF2 (alternative splicing factor/splicing factor 2), has been shown to interact simultaneously with the U1 70K component of U1 snRNP (Wu and Maniatis, 1993; Kohtz et al., 1994) and the RS domain of $U2AF^{35}$ (Wu and Maniatis, 1993). The SR protein, SC35, was shown to carryout similar cross exon interactions (Wu and Maniatis, 1993). Furthermore, because it is the first committed step in the splicing process, exon definition is a particularly sensitive mechanism to specify alternative splicing patterns (Robberson et al., 1990; Kuo et al., 1991).

1.2.3 Complex A: branchpoint selection

During complex A assembly, U2 snRNP interacts with the branch site in an ATP-dependent manner (Konarska and Sharp, 1987; Jamison and Garcia-Blanco, 1992). U2 snRNA base pairs with the branch site region to form a bulged duplex, with the branchpoint adenosine unpaired, thereby presenting the adenosine to be used as the nucleophile in the first step of the splicing reaction (Query et al., 1994). Interaction of $U2AF^{65}$ with the polypyrimidine tract, positions the basic residues of its RS domain at the branchpoint (Valcarcel et al., 1996). The positive nature of this domain neutralizes the negative charge of the phosphate backbone of the snRNA-pre-mRNA complex and therefore facilitates the U2 snRNA-branch site interaction. Furthermore, SAP155 of U2 snRNP interacts with the RS domain of $U2AF^{65}$ and positions U2 snRNP to the branch site by interacting with nucleotides flanking the branchpoint (Gozani et al., 1998). In fact, SAP155 replace SF1 (Berglund et al., 1998), thereby allowing the branchpoint to be accessible by the splicing machinery.

In addition to SAP155, SAPs 145, 130, 114, 62, 61, 49, and 33 and U2 snRNP-specific proteins A´ and B´´ associate in complex A (Bennett et al., 1992). SAP114, 62, and 61 are considered SF3a and 155, 145, 130, and 49 are considered SF3b. These proteins bind approximately 20 nucleotides upstream of the branch site in a non-specific manner and also function to tether U2 snRNP to the branch site (reviewed by (Will and Luhrmann, 1997)). Moreover, the SR protein, SC35, has been shown to mediate interactions between U1 and U2 snRNPs for cross intron communication of the spliceosome (Fu and Maniatis, 1992). Note that sometime after complex A assembly, U2AF disassociates from the pre-mRNA (Bennett et al., 1992).

1.2.4 Complex B: chemical step 1

During complex B, the tri-snRNP particle, composed of U5 and U4/U6 snRNPs associates with the spliceosome (Bindereif and Green, 1987; Konarska and Sharp, 1987; Jamison et al., 1992). At this time, the catalytically inactive pre-spliceosome is converted to the catalytically active spliceosome. Furthermore, this step is associated with the lag time observed in an *in vitro* splicing reaction. U4/U6 snRNP enters as a duplex RNA with base pairing between the U4 and U6 snRNAs. RNA-RNA rearrangements of the spliceosome lead to the disruption of the U4/U6 duplex and the release of U4 snRNP (Lamond et al., 1988). This process likely involves the activity of an RNA helicase or "unwindase" (reviewed by (Staley and Guthrie, 1998)). At this time, the 3´ end of U6 snRNA base pairs with the 5´ end of U2 snRNA, forming a short RNA helix (Hausner et al., 1990; Datta and Weiner, 1991; Wu and Manley, 1991) closely positioned at the catalytic center of the spliceosome (Madhani and Guthrie, 1992; Sun and Manley, 1995). U6 snRNP interaction with the 5´ splice site and U2 interaction with the branch site positions the spliceosome to initiate the first catalytic step in splicing (Sun and Manley, 1995). Furthermore, U5 snRNP associates with the pre-mRNA at the 5´ and 3´ splice sites to align the pre-mRNA substrate for splicing (Wassarman and Steitz, 1992; Wyatt et al., 1992; Sontheimer and Steitz, 1993). Removal of U4 snRNP occurs at a time concomitant with the formation of the lariat-3´ exon intermediate and removal of U1 snRNP (Lamond et al., 1988). In fact, U1 snRNP may dissociate at an earlier step after U2 snRNP association or may be less stably bound, because it is not detected in purified spliceosomes after U2 snRNP association.

During complex B assembly, SF4 also associates to help generate an active spliceosome (Utans and Kramer, 1990). The 200, 116, 112, 110, and 40 kDa U5 snRNP-specific SAPs and the 102, 90, 82, 68, 60, 57, and 55 kDa SAPs also associate during this step (Bennett et al., 1992). As for the earlier steps in spliceosome assembly, SR proteins have been implicated in this step of splicing. For example, SR proteins facilitate the incorporation of the tri-

snRNP into the spliceosome (Roscigno and Garcia-Blanco, 1995) and have also been demonstrated to stimulate base pairing between U2 and U6 snRNAs (Tarn and Steitz, 1995). Moreover, several of the tri-snRNP associated proteins contain RS domains (Fetzer et al., 1997), which may be involved in protein-protein interactions and structural transitions of the spliceosome.

1.2.5 Complex C: chemical step 2

Complex C is designated by the formation of the 5′ exon and lariat intron-3′ exon intermediates from complex B. After the first step of splicing, a second conformational change occurs, which results in the release of the lariat intron and ligated exon products. U2, U5, and U6 snRNAs remain associated with the lariat intron product (Konarska and Sharp, 1987; Lamond et al., 1988; Wassarman and Steitz, 1992). *In vivo,* the lariat product is debranched and degraded and the snRNP components are recycled to carryout additional rounds of splicing (Ruskin and Green, 1985).

1.2.6 The catalytic center of the spliceosome

U6 snRNP has been shown to be closely associated with the 5′ splice site during the first step of splicing and the 3′ splice site during the second step, which implicates it in catalysis of the splicing reaction (Sawa and Shimura, 1992). In fact, the U2/U6 snRNP has been demonstrated to carryout a splicing-like reaction in the absence of protein (Valadkhan and Manley, 2001; Valadkhan et al., 2007), suggesting that the spliceosome may itself be a ribozyme like the self-splicing introns of Tetrahymena pre-rRNA. Conversely, it has been suggested that the highly conserved U5 snRNP-associated protein, Prp8, is the catalytic center of the spliceosome (reviewed in (Wachtel and Manley, 2009)). Prp8 physically interacts with the 5′ splice site, 3′

splice site, and branch site during various steps of splicing, appears to be involved in the transition between the two catalytic steps, and is essential for the spliceosome-catalyzed splicing process. This protein also contains an RNase H-like fold that would be ideal for facilitating RNA rearrangements for splicing.

Although the spliceosome has been extensively studied in the past, several questions remain. How can the spliceosome recognize such diverse RNA substrates with little sequence information and still carryout the precise process of splicing? Is the spliceosome a ribozyme? If so, why are so many proteins involved? Do these proteins control the specificity and timing of splicing? Furthermore, how do alternative splicing factors interact with the spliceosome to specify alternative splicing patterns?

1.3 MINOR CLASS INTRONS

So far, our discussion of the spliceosome has been about the major class U2 spliceosome, which is associated with the splicing of the major class GT-AG introns. Here, I briefly discuss the minor class of AT-AC introns and the U12-containing spliceosome that carries out their splicing. The observation that some genes do not contain canonical splice sites and the idea that an alternate group of splice sites exist was first discussed by Jackson (1991). This rare class of introns contains the 5´ splice site consensus, AUAUCCUU (beginning with the first nucleotide in the intron), the 3´ splice site consensus, YCCAC (Y, pyrimidine; ending in the last nucleotide of the intron), and the branch site consensus, UCCUUAAC (second A, branchpoint) (Hall and Padgett, 1994; Sharp and Burge, 1997).

The consensus 5´ splice site is complementary to U11 snRNA and the consensus branch site is complementary to U12 snRNA. In fact, U11 and U12 snRNAs are the minor class counterparts of U1 and U2 snRNAs, respectively (Hall and Padgett, 1996; Tarn and Steitz,

1996b). Furthermore, the highly diverged U4 and U6 snRNAs, U4atac and U6atac, are associated with this minor class spliceosome (Tarn and Steitz, 1996a). The only major class snRNA that is involved in splicing of minor class introns is U5 snRNA (Tarn and Steitz, 1996b). As for the major class U6 and U2 snRNAs, U6atac and U12 snRNAs form a duplex RNA (Tarn and Steitz, 1996a). The minor class spliceosome catalyzes splicing by the same chemical mechanism as the major class spliceosome. The discovery of this minor class spliceosome reflects the fundamental requirements for the construction of a catalytically active macromolecular machine and provides support for the RNA interactions identified for the major class spliceosome.

Some pre-mRNAs contain multiple U2-type introns but a single U12-type intron. In fact, there is approximately one U12-type intron for every 300 U2-type introns, which is consistent with the lower abundance of the minor class spliceosome (Sheth et al., 2006). It has been demonstrated that removal of U12-type introns occurs slower than removal of U2-type introns and is therefore the rate-limiting step in the maturation of a pre-mRNA containing both classes of introns (Patel et al., 2002). Therefore, U12-type introns may have evolved as a mechanism to regulate timing of gene expression.

1.4 ALTERNATIVE SPLICING

Given the degenerate nature of metazoan consensus splice sites, how does the spliceosome achieve the specificity needed to accurately define a set of exons and precisely remove an intervening intron sequence while maintaining the fidelity of the reading frame? Consensus splice sites only contain half of the information needed for proper splice site recognition (Lim and Burge, 2001). Furthermore, numerous “decoy” exons exist within large metazoan introns that resemble real exons in size and splice site strength (Sun and Chasin, 2000). However,

these exons are rarely spliced. Therefore, accessory factors are necessary for proper splicing in metazoans. These factors include splicing enhancers that function to aid in the assembly of the spliceosome and splicing silencers that function to inhibit spliceosome assembly, often times by blocking access to the 5′ and 3′ splice sites. In yeast, the introns are much smaller in size and consensus splice sites are far more conserved. Therefore, the yeast spliceosome is sufficient for accurate splicing, consistent with the limited use of splicing enhancers and silencers.

The use of splicing enhancers and silencers to adjust spliceosome assembly and exon selection is the process of alternative splicing. In metazoans, spliceosome assembly can proceed through multiple pathways leading to silencing (skipping) or enhancement (inclusion) of an alternatively spliced exon (cassette exon), or may lead to other types of splicing decisions such as mutually exclusive exon selection, alternative 5′ and 3′ splice site selection, or intron retention. The tissue- and developmental-specific expression of splicing enhancers and silencers tunes splicing patterns in a spatial and temporal manner. The protein diversity generated by alternative splicing, in part, provides an explanation for the evolution of complex organisms with specialized cell types given a fixed number of genes. In fact, it is estimated that 92-94% of all genes expressed in human tissues undergo alternative splicing events (Wang and Burge, 2008). The *Drosophila* Down syndrome cell adhesion molecule (Dscam) transcript is one extreme example of molecular diversity generated from a single gene by splicing (Schmucker et al., 2000). This transcript has the potential to produce 38,016 unique protein isoforms that function to generate neuronal circuits during brain development. In simpler cases, one or two exons are alternatively spliced to modulate protein properties in specific cell types.

1.4.1 Cis-regulatory RNA elements

Exonic splicing enhancer (ESE) and silencer (ESS) and intronic splicing enhancer (ISE) and silencer (ISS) elements are RNA sequence motifs that are recognized by alternative splicing

factors to tune splicing patterns (reviewed by (Ladd and Cooper, 2002; Wang and Burge, 2008)). Collectively, these motifs are referred to as splicing regulatory elements (SREs). ESE and ESS elements are located within the exons that they regulate and ISE and ISS elements are located within the flanking intron sequences in close proximity to the exons that they regulate. SREs are typically short and degenerate, present in multiple copies, and are usually single-stranded, although secondary structure has been implicated in the function of a few elements. These elements provide additional information to the splicing machinery and are necessary to specify alternative splicing patterns. The current goal in the splicing field is to generate a parts list of regulatory elements that can be used to predict the splicing pattern of an exon by sequence inspection alone.

Large-scale randomization and selection approaches have led to the identification of numerous ESE/S and ISE/S sequences that functionally affect splicing patterns. The results of these selections have been used to generate ESEFinder (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi) (Cartegni et al., 2003), RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) (Fairbrother et al., 2002), and FAS-ESS (http://genes.mit.edu/fas-ess/) (Wang et al., 2004) web servers, which offer databases of splicing motifs with search capabilities. These global approaches have generated a catalogue of splicing elements that can be useful to predict regulation of individual exons of a gene. However, the roles of these regulatory sequence motifs in natural pre-mRNAs is likely influenced by the strength of the flanking splice sites, the sequence context, and by the expression levels of the regulatory factors, which may vary in different cells or during development. Another issue is the copy number of the regulatory motifs. A single motif that has a weak effect on splicing can have a strong effect in the presence of additional copies of the same or functionally similar motifs.

1.4.2 Splicing codes

Splicing codes are combinations of ESE/S and ISE/S motifs that tune splicing patterns through specific interactions with cognate RNA binding proteins (reviewed by (Wang and Burge, 2008)). Three criteria can be used to experimentally verify a splicing code for a SRE. (1) Changes in the copy number and/or arrangement of the RNA motifs should tune the splicing pattern to a significant degree. (2) The splicing code should be functionally transferable to a new splicing substrate by gain-of-function. (3) Genome-wide searches for the splicing code should identify additional regulatory targets at frequencies greater than random chance. Testing such candidates can further refine the code, and may suggest interesting new biological targets.

1.4.3 Trans-acting protein factors

RNA binding proteins, including members of the SR, hnRNP, and hnRNP K-homology (KH) families, have been shown to regulate alternative splicing events at the level of spliceosome function. These splicing factors contact their RNA targets directly by recognition of ESE/S or ISE/S elements. Observed splicing patterns are often a result of the dynamic balance of enhancing and silencing functions.

SR splicing factors, such as ASF/SF2, generally enhance exon inclusion and contain an RBD together with a characteristic RS domain (reviewed by (Graveley, 2000; Maniatis and Tasic, 2002)). These factors are involved in alternative splicing as well as constitutive exon inclusion as discussed above for the spliceosome. Upon binding to ESE motifs, SR proteins recruit factors such as U1 snRNP or U2AF, or antagonize the effects of splicing silencers to facilitate exon definition.

Proteins in the hnRNP family, such as polypyrimidine-tract binding protein (PTB) or hnRNPA1, contain multiple RNA recognition motifs (RRMs) and often promote exon skipping by

antagonizing the roles of SR proteins or snRNPs (Choi et al., 1986; Burd et al., 1989; Ladd and Cooper, 2002). Nonetheless, some hnRNP family members enhance exon inclusion, such as Fox1 and Fox2 (Underwood et al., 2005). Others have dual roles capable of both enhancement and silencing, such as members of the CUG binding protein (CUGBP) and ETR3-like family, termed CELF factors (Ladd et al., 2001; Zhang et al., 2002). In the KH domain family, the Nova1 and Nova2 splicing factors have been characterized as dual-functional regulators of alternative splicing events for transcripts encoding synaptic functions (Ule et al., 2003; Ule et al., 2005).

Proteins from the SR and hnRNP families of RNA binding proteins are phosphorylated and dephosphorylated to adjust their functions (reviewed by (Stamm, 2007)). Phosphorylation can modify the ability of RNA binding proteins to interact with RNA and other proteins, can alter subcellular localization, and can influence intranuclear localization. Consequently, such adjustments can affect the ability of splicing enhancers and silencers to regulate alternative splicing events or may couple splicing to signal transduction pathways. Furthermore, these modifications may serve important roles in mechanisms that integrate splicing responses and extracellular signaling. The findings that many alternative splicing factors are also involved in transcription, polyadenylation, 3´ end formation, nuclear export of mRNAs, nonsense-mediated mRNA decay (NMD), and translation indicates that splicing is integrated with other pathways of gene expression (reviewed by (Goldstrohm et al., 2001; Long and Caceres, 2009)).

1.4.4 Mechanisms of alternative splicing regulation

Although mechanisms of alternative splicing regulation are diverse and largely uncharacterized, PTB, Nova, and Fox splicing factors have been extensively studied and serve as good examples for regulation. PTB is a widely expressed RNA binding protein that has been well studied for its functions to silence neural- and muscle-specific exons in a variety of cell types (Zhang et al., 1999). PTB and PTB isoforms bind to pyrimidine-rich motifs related to the core

sequence, UCUU, through four RRMs to silence alternative exons (Wagner and Garcia-Blanco, 2001; Spellman and Smith, 2006). PTB is best characterized for its role in silencing the neural-specific NI exon of the tyrosine kinase, c-src, in nonneuronal cells (Black and Grabowski, 2003). PTB silences the NI exon by binding to motifs in the flanking introns to loop out the exon and block access of the splicing machinery to the splice sites. PTB can also bind near the 3´ splice site of an alternative exon to antagonize $U2AF^{65}$ binding (Singh et al., 1995) or to sequester the branch site (Ashiya and Grabowski, 1997). It can also silence an exon by propagating across the RNA to antagonize binding of splicing enhancers and the general splicing machinery (Wagner and Garcia-Blanco, 2001). Furthermore, two recent studies have demonstrated that PTB can silence alternative exons by inhibiting exon and intron definition events (Izquierdo et al., 2005; Sharma et al., 2005). Thus, PTB is a remarkably versatile splicing silencer.

Mechanisms of splicing silencing and enhancement by Nova proteins have also been studied at the level of spliceosome assembly. Nova1 and Nova2 are brain-specific RNA binding proteins that were first identified as target antigens in the neurological disorder, paraneoplastic opsoclonus-myoclonus ataxia (Buckanovich et al., 1993). Nova proteins have 3 KH-type RBDs, which recognize consensus YCAY motifs (Y, pyrimidine) to enhance or silence splicing of alternative exons (Dredge and Darnell, 2003; Dredge et al., 2005). The interaction of Nova with an ESS inhibits binding of U1 snRNP to the adjacent 5´ splice site (Ule et al., 2006). However, when Nova binds to an ISE downstream of an alternative exon, this interaction promotes assembly with U2 snRNP. It is likely that the positioning of SREs in the pre-mRNA influences the functions of splicing factors (Dredge et al., 2005).

The mammalian Fox proteins are homologs of the *Caenorhabditis elegans* feminizing gene on X protein. Mammalian Fox family members, Fox1 and Fox2, are expressed in neurons of the adult brain and function as enhancers of brain-specific exons that have Fox binding sites in their downstream intron. Fox1 and Fox2 bind to a short but specific consensus motif, UGCAUG, and function to activate splicing of an alternative exon (Underwood et al., 2005).

Fox1 has been characterized for its function as a splicing enhancer of the NI exon of the c-src transcript (Underwood et al., 2005). Fox1 binds downstream of the NI exon at a conserved UGCAUG hexamer in a way that is thought to antagonize the splicing silencer, PTB. Like Nova, Fox proteins can also act as splicing silencers when bound to the intron upstream of an alternatively spliced exon (Fukumura et al., 2007).

Taken together, alternative splicing factors are remarkably versatile proteins that contact SREs to tune splicing patterns in a tissue-specific manner. These factors often have different effects on spliceosome assembly depending on the position and sequence context of their cognate RNA binding sites. Furthermore, competition and cooperation of alternative splicing factors acting on the same pre-mRNA creates a tug-of-war for splicing regulation. Still a common characteristic of these RNA binding proteins is that they recognize short and degenerate sequence motifs, which once again raises the question of how splicing factors can accurately select the correct target exons for regulation.

1.4.5 Regulation of tissue- and developmental-specific expression of splicing factors

In recent years, several alternative splicing factors have been characterized for having unique features that contribute to their tissue- and developmental-specific expression and functions. It is becoming clear that these features are more common than once thought and may actually be characteristic of most alternative splicing factors. These characteristics include alternative splicing within the splicing factor transcript, autoregulation and NMD, cross-regulation of family members, and regulation by micro RNAs (miRNAs). Several examples are described below.

Many alternative splicing factors are themselves alternatively spliced to generate several protein isoforms that carryout distinct functions in specific cell types. For example, mammalian Fox1 and Fox2 transcripts are extensively alternatively spliced to generate multiple protein isoforms (Underwood et al., 2005). Skipping of one exon within Fox transcripts from both genes

deletes essential amino acid residues in the protein to eliminate RNA binding activity (Nakahata and Kawamoto, 2005). In this way, alternative splicing of splicing factor transcripts can be used to regulate their function.

Furthermore, many alternative splicing factors regulate their own expression and activity by controlling alternative splicing of exons within their own transcript. For example, Nova1 has the ability to silence an exon within its own transcript (Dredge et al., 2005). The peptide segment corresponding to this exon is phosphorylated *in vivo* and may serve as a localization signal or to modulate protein function. Furthermore, many factors from the SR family of splicing enhancers have been shown to increase inclusion of an exon within their own transcript to cause a shift in the reading frame revealing a premature termination codon (PTC) that targets the transcript for NMD (Ni et al., 2007). NMD is a quality control process that selectively degrades mRNAs containing PTCs to control for genetic and splicing errors and regulate protein expression (Chang et al., 2007). Conversely, several splicing silencers have been shown to function by the opposite mechanism to cause skipping of an exon within their own transcript to expose a PTC (Wollerton et al., 2004; Baraniak et al., 2006). Therefore, NMD is a process utilized by splicing factors to control their own protein levels in the cell.

Additionally, cross-regulation of closely related family members has been shown for several splicing factors. In the nervous system, cross-regulation between PTB family members drives mutually exclusive isoform expression in different cell types. In the mouse brain, neural PTB (nPTB) is exclusively expressed in neurons, while PTB expression is restricted to glial cells (Boutz et al., 2007b). One explanation for this mutually exclusive pattern of expression is that PTB causes skipping of an exon in the nPTB transcript by a mechanism similar to autoregulation to target the transcript for NMD. PTB and nPTB regulate distinct, yet overlapping sets of exons and their cell-specific expression is associated with the differentiation of neuronal and glial cells during brain development.

The expression of PTB is also regulated through post-transcriptional events involving miRNAs. miRNAs are a well studied group of ~22 nucleotide RNAs that bear complimentarity to sequences in the 3´ untranslated regions (UTRs) of certain mRNAs where their interactions either repress protein synthesis or promote mRNA decay (He and Hannon, 2004). The neuron-specific miRNA, miR-124, inhibits expression of PTB during neuronal differentiation (Makeyev et al., 2007). Decrease in PTB expression leads to an increase of nPTB expression, which is associated with a global switch from non-neuronal to neuronal splicing patterns. Furthermore, a role for miR-133 in the regulation of nPTB expression has been demonstrated in differentiating muscle cells (Boutz et al., 2007a).

Taken together, these characteristics suggest that global patterns of alternative splicing can be controlled in different cell types and during development by several different quality control mechanisms. How other splicing factors are targets of such regulation is still to be investigated.

1.5 THE NMDA R1 RECEPTOR

In the nervous system, alternative splicing greatly expands protein diversity and is necessary for neural development and shaping the synapse for processes such as learning and memory. Here, alternative splicing is generally the rule rather than the exception. The N-methyl-D-aspartate (NMDA) receptor NR1 subunit is just one example. NMDA receptors are calcium permeable glutamate receptors present at the postsynaptic membrane of excitatory synapses in the mammalian central nervous system. These receptors are important for neuronal survival and learning and memory, while abnormalities in receptor activity have been implicated in neurological disorders. These receptors allow calcium influx that regulates signaling pathways involved in synaptic plasticity in response to the binding of glutamate and glycine under

depolarizing conditions. Alternative splicing of the NR1 subunit of this receptor has been shown to modulate receptor function in a brain region-specific manner and in response to cell excitation.

1.5.1 Brain region-specific alternative splicing

Alternative splicing of three exons of the NR1 subunit of the NMDA receptor pre-mRNA (GRIN1 gene) generates eight receptor isoforms with different functional characteristics (reviewed by (Zukin and Bennett, 1995; Cull-Candy et al., 2001)). The NI cassette exon (Ensembl transcript GRIN1-002, exon 3b), which encodes a portion of the extracellular domain, modulates receptor sensitivity to zinc ions, protons, and polyamines. The CI cassette exon (GRIN1-002, exon 19), which encodes an intracellular region, regulates membrane localization of the receptor and intracellular signaling from the receptor (Ehlers et al., 1995; Okabe et al., 1999; Bradley et al., 2006). In addition, mutually exclusive splicing of C2 and C2´ exons adjusts receptor trafficking to the synapse (Mu et al., 2003). Thus, alternative splicing fine-tunes NMDA receptor properties.

The brain region-specific splicing factor, CUGBP2 (also called NAPOR, CELF2, ETR3, or BRUNOL3), silences the NI exon and enhances the CI exon of the GRIN1 transcript (Zhang et al., 2002). These dual functions in splicing regulation, together with the forebrain-enriched and hindbrain-deficient pattern of protein expression, are thought to account for the natural distribution of NR1 isoforms in the brain (Figure 3). That is, the NI exon is mostly skipped and the CI exon is mostly included in the forebrain as directed by the presence of CUGBP2, whereas both of these splicing events switch to the opposite patterns in the hindbrain because of an absence of CUGBP2. The molecular mechanisms that explain the dual functions of CUGBP2 are not yet understood. Although it's RNA binding motifs have been found in the introns downstream from the CI exon and upstream from the NI exon, functionally significant motifs in these regions have yet to be identified.

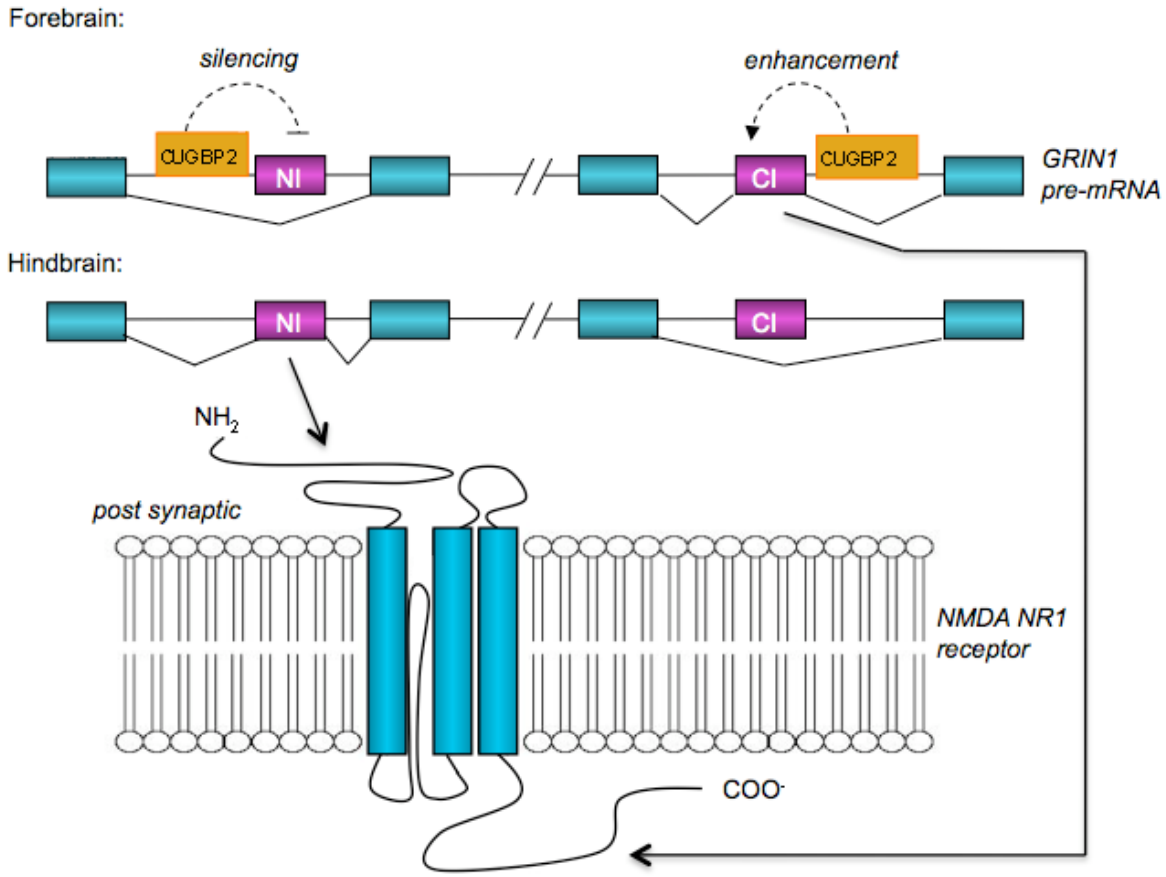


Figure 3. Model for brain region-specific splicing of the NMDA NR1 receptor transcript. The GRIN1 pre-mRNA is shown with the NI and CI cassette exons indicated in pink and the flanking exons in blue. The broken line indicates a gap in the sequence. In the forebrain, tissue-specific CUGBP2 expression drives skipping of the NI exon (silencing) and inclusion of the CI exon (enhancement). In the CUGBP2-deficient hindbrain, the opposite pattern of splicing occurs. Below, the relative positions of the NI and CI peptide cassettes on the NMDA NR1 receptor are shown.

Alternative splicing of the NI and CI exons is under combinatorial control by RNA binding proteins from the SR, hnRNP, and KH families of splicing factors. The dynamic balance of these splicing factors in different brain regions determines whether NI or CI will be predominantly included or skipped in the population of mature GRIN1 transcripts. The CI exon is enhanced by SR family members ASF/SF2 and SC35 (Han et al., 2005), hnRNP H (Han et al., 2005), and the KH domain protein, Nova (Ule et al., 2006) in addition to CUGBP2. It is also silenced by hnRNP A1 (Han et al., 2005). In each of these cases, known cis-regulatory motifs within the exon and downstream intron have been shown to contribute to regulation. Furthermore, the NI exon is silenced by CUGBP2, PTB, and Nova splicing factors (Zhang et al., 2002) and potentially enhanced by Fox family members (unpublished data). The NI and CI cassette exons are representative examples of the complex regulation of an exon by multiple factors, which leads to the fine-tuning of its tissue- and developmental-specific protein activities (Black and Grabowski, 2003).

1.5.2 Alternative splicing in response to neuronal excitation

Recent studies have shown that the CI cassette is an inducible exon in living cells, since its splicing pattern shifts to predominant exon skipping after membrane depolarization or as a consequence of overexpression of a constitutively active form of Ca^{2+}/calmodulin-dependent protein kinase (CaMK) IV (An and Grabowski, 2007; Lee et al., 2007). The CI exon encodes an intracellular peptide region that is believed to mediate intracellular signaling from the cell membrane to the nucleus. In this way, the decrease in CI exon inclusion may be part of a feedback loop that reduces signaling in neurons to stabilize neuronal excitability during chronic excitation. Furthermore, this depolarization-induced skipping pattern is reversed upon removal of membrane depolarization (An and Grabowski, 2007), indicating that splicing patterns are plastic and probably involved in maintaining cellular homeostasis.

The observation that splicing of the CI exon in the nucleus is responsive to cell excitation at the membrane implicates a signaling pathway to transfer the message. Notably, specific antagonists of the NMDA receptor and several cell permeable inhibitors of signal transduction pathways, including Protein Kinase A and CaMKs, were found to attenuate splicing responsiveness in the primary neurons (An and Grabowski, 2007). These findings suggest that one or more signaling pathways from NMDA receptors located at the membrane to factors in the nucleus are responsible for inducing changes at the level of splicing of the CI exon.

hnRNPA1 has been implicated as the depolarization-responsive splicing factor involved in this signaling event (An and Grabowski, 2007). In support of this, additional responsive exons have been shown to contain hnRNPA1 silencing elements. However, it is not clear how the activity of hnRNPA1 is modified in response to cell excitation. Furthermore, two types of CaMK IV-responsive RNA elements (CaRREs) have been found within the CI exon. One element is similar to an element defined previously in the 3´ splice site region of the responsive STREX exon of the BK channel (Xie et al., 2005). Genome-wide searches for additional exons that contain CaRREs have identified CaMK IV-responsive exons in transcripts with functions such as calcium homeostasis, intracellular signaling, and vesicular transport. Still, the protein factors that recognize such elements have yet to be identified.

Together these studies highlight the importance of conserved sequence elements in the regulation of cellular responsiveness of groups of co-regulated exons, but this work is in an early stage. Many questions remain about the nature of the signal transduction pathways involved in conveying information from the cell membrane to splicing factors in the nucleus, and the nature of the landscape of responsive exons that has yet to be fully characterized. The extent to which the induced splicing changes reflect corresponding changes at the protein level, as opposed to biological noise, is also not known. A recent study has described the RNA splicing capacity of neuronal dendrites in hippocampal cultures (Glanzer et al., 2005). Although

speculative, this raises the question of whether certain mRNAs with one or two unspliced introns might be targeted to dendrites for local splicing in response to changes in synaptic activity.

1.6 CUGBP2

CUGBP2 is a member of the CELF family of hnRNP RNA binding proteins that has been characterized for its role in the regulation of tissue-specific splicing events in accordance with brain region- and muscle-specific expression during development. Misregulation of CUGBP2 has been associated with global splicing defects contributing to disease phenotypes of myotonic dystrophy type 1 (DM1). CUGBP2 consists of three RRMs and a divergent domain of unknown function and has been shown to bind to GU-rich sequence motifs to either silence or enhance exon selection. However, the mechanisms of regulation are currently uncharacterized.

1.6.1 The CELF family of RNA binding proteins

CUGBP2 is a member of the larger family of CELF RNA binding proteins. Currently, in mammals, there are six known family members, CUGBP1, CUGBP2, CELF3, CELF4, CELF5, and CELF6 that differ in their unique functions and tissue-specific expression patterns (Good et al., 2000; Ladd et al., 2001). CUGBP1 is ubiquitously expressed while expression of CELF3 and CELF5 is restricted to the brain. CUGBP2 is expressed in heart, muscle, and brain tissue and CELF4 and CELF6 are expressed in a subset of tissues. CUGBP2 is most closely related to the CUGBP1 family member and in fact share many overlapping functions.

CELF proteins have three RRMs that contact RNA in a sequence-specific manner and a variable divergent domain of unknown function (Figure 4). The RRM is a simple but widely used

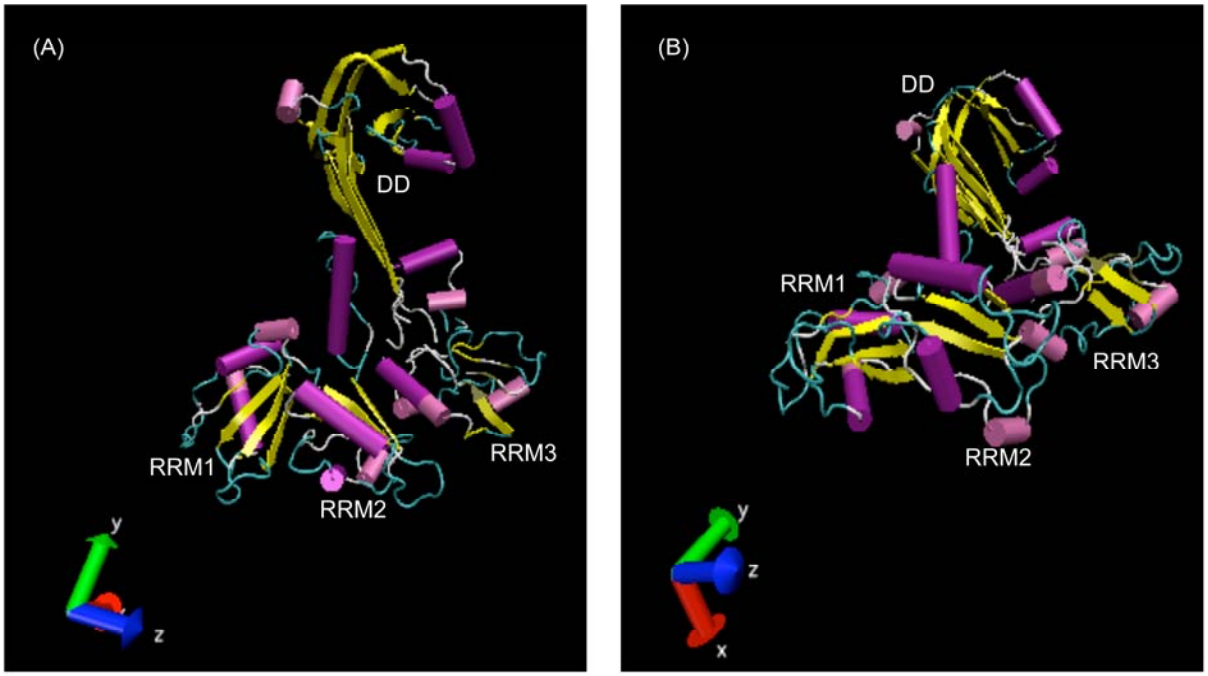


Figure 4. Predicted tertiary structure of CUGBP2. (A) Side-view of the CUGBP2 tertiary structure predicted using the InterProScan software (EMBL-EBI). Snapshots of the PDB structure were taken with the interactive molecular graphics program, VMD. The predicted positions of α-helices (pink), β-sheets (yellow), and unstructured regions (blue) are indicated on RRMs 1, 2, and 3 and the divergent domain (DD). Relative positions of the x-, y-, and z-axes are shown for comparison. (B) Bottom-view of the CUGBP2 structure in (A). Note that this is a predicted structure of CUGBP2 and is likely different than the actual structure of the protein.

domain consisting of a βαββαβ topology that can fold in different conformations to recognize unique RNA or protein substrates (reviewed by (Maris et al., 2005)). The RRM provides high specificity in the selection of target substrates and the use of multiple RRMs increases the specificity and affinity of binding by a single protein. CUGBP1 and 2 RRMs contact GU-rich sequences as indicated by a number of biochemical and functional studies (Takahashi et al., 2000; Suzuki et al., 2002; Faustino and Cooper, 2005; Marquis et al., 2006). However, specific nucleotide motifs recognized by each of these proteins have yet to be identified. It is likely that each of these proteins regulates a distinct, yet overlapping group of target exons by slight variations in their binding abilities.

CUGBP2 protein isoforms differ in their N-terminal amino acid sequence by the use of alternative start codons and in the divergent domain by alternative splicing (Li et al., 2001). Three isoforms are known to exist including NAPOR1, NAPOR2, and NAPOR3. For the purpose of this study, I will focus on the NAPOR1 isoform of CUGBP2. How these isoforms differ in RNA binding or splicing regulation is currently unknown.

1.6.2 Brain region-specific expression in the central nervous system

Brain-specific expression of CUGBP2 is restricted to the forebrain, particularly the cerebral cortex, where CUGBP2 carries out functions associated with neuronal cell differentiation and development of the nervous system in addition to adjusting modular receptor properties at the synapse (Choi et al., 1999; Good et al., 2000; Li et al., 2001; Levers et al., 2002; Zhang et al., 2002; Choi et al., 2003). CUGBP2 expression is enriched in post-mitotic differentiating neurons during embryogenesis and early postnatal life (Choi et al., 1999; Levers et al., 2002). However, expression is deficient in undifferentiated proliferative cells. Therefore, the cell type-specific expression of this splicing factor likely contributes to unique splicing patterns in differentiated neurons. RNA microinjection studies in zebrafish embryos have demonstrated that

overexpression of CUGBP2 results in severe developmental defects in the nervous system (Choi et al., 2003). At the synapse, brain region-specific expression of NMDA receptor NR1 subunit isoforms is closely correlated with CUGBP2 expression and regulation (discussed above). Taken together, these observations suggest that CUGBP2 carries out essential functions in the central nervous system that are tightly controlled by its brain region- and cell type-specific expression patterns.

1.6.3 Regulation of heart- and muscle-specific exons during development

CUGBP2 has also been characterized for its muscle-specific expression and roles in muscle and heart development (Good et al., 2000; Li et al., 2001; Suzuki et al., 2002; Gromak et al., 2003; Ladd et al., 2005; Kalsotra et al., 2008). CUGBP1 and 2 function as splicing enhancers of cardiac troponin T (cTNT) exon 5 by binding to GU-rich motifs downstream from the exon (Ladd et al., 2005). Here, enhancement is associated with competition with the PTB and muscleblind-like (MBNL) splicing silencers. More recently it has been demonstrated that CUGBP2 interacts with protein components of U2 snRNP across the exon to facilitate branch site recognition at the upstream intron (Goo and Cooper, 2009). Furthermore, cTNT exon 5 inclusion in the developing heart is associated with CUGBP1 and 2 expression, while exon 5 skipping in the adult heart is associated with a deficiency of these factors (Ladd et al., 2005). Additional studies have shown that this postnatal switch in CUGBP1 and 2 expression is associated with over 30 splicing events during heart development (Kalsotra et al., 2008). Furthermore, CUGBP1 and 2 have been implicated in the mutually exclusive selection of a set of exons in the α-actinin transcript (Suzuki et al., 2002; Gromak et al., 2003). The SM (smooth muscle) exon is repressed in most cell types by PTB, while the NM (nonmuscle) exon is included. In muscle cells, expression of CUGBP1 and 2 drives skipping of the NM exon by an uncharacterized mechanism and inclusion

of the SM exon in part by competition with PTB. Therefore, regulation by the CUGBP1 and 2 splicing factors drives exon selection in a tissue- and developmental-specific manner.

1.6.4 Gain-of-function in DM1: from splicing defects to disease phenotypes

DM1 is an autosomal dominant neuromuscular disease associated with a $(CTG)_n$ repeat expansion in the 3´ UTR of the myotonin protein kinase (Mt-Pk) gene. These repeats have been shown to sequester CUGBP1 and 2 in the nucleus, resulting in a gain-of-function of protein activity by an unknown mechanism (Timchenko et al., 1996; Timchenko et al., 2001b; Wang et al., 2007). Defects in CUGBP1 and 2 splicing regulation contribute to the multisystemic disease phenotypes in patients with DM1 (Timchenko et al., 2004). Additional defects result from a loss-of-function of MBNL splicing factors and gain-of-function of hnRNPH (Paul et al., 2006).

CUGBP1 and 2 gain-of-function has been associated with an increase in cTNT exon 5 inclusion (Philips et al., 1998). cTNT is the heart-specific tropomyosin-binding subunit of the troponin complex, which is located on the thin filament of striated muscle and regulates muscle contraction in response to endogenous calcium levels. The increased expression of the fetal isoform, which includes exon 5, in the adult heart, may be associated with reduced myocardial function and conduction abnormalities observed in patients with DM1.

Loss of expression of the muscle-specific chloride channel, ClC-1, in patients with DM1 is associated with CUGBP1 gain-of-function and aberrant splicing of ClC-1 pre-mRNA (Charlet et al., 2002b). CUGBP1 and possibly CUGBP2 gain-of-function cause aberrant intron inclusion in the ClC-1 transcript by binding to an intronic element near the 3´ splice site of the retained intron. Intron retention results in a shift in the reading frame, which reveals a PTC that likely targets the aberrant transcript for NMD. ClC-1 is the major skeletal muscle chloride channel that is localized to the sarcoplasmic reticulum where it is responsible for chloride conductance and

stabilizing membrane potential. Defects in this channel are greatly associated with myotonia and likely contribute to muscular defects of DM1.

Functional defects of CUGBP1 and 2 have also been associated with increased skipping of the insulin receptor (IR) exon 11 (Savkur et al., 2001; Han and Cooper, 2005; Paul et al., 2006). CUGBP1 gain-of-function reduces exon 11 inclusion as a result of binding to an intronic element near the 3´ splice site (Savkur et al., 2001). Increased expression of insulin receptor isoforms that lack the corresponding peptide cassette is associated with the decreased metabolic response to insulin in patients with DM1. CUGBP2 likely silences this exon through similar intronic elements (Han and Cooper, 2005). Recently, it has been demonstrated that IR exon 11 is under coordinate regulation by the silencing effects of CUGBP1 and hnRNPH and enhancing effects by MBNL1 (Paul et al., 2006). Therefore, a combination of hnRNPH and CUGBP1 and 2 gain-of-function and MBNL1 loss-of-function contributes to the splicing defects of IR exon 11 associated with DM1.

Patients with DM1 also exhibit splicing defects in the central nervous system. CUGBP2 regulates the microtubule-associated protein Tau (MAPT) exons 2/3 splicing. As a result of CUGBP2 gain-of-function in DM1, MAPT exons 2/3 are excluded more often (Leroy et al., 2006). MAPT mutations have been associated with a number of neurological disorders including Alzheimer's disease, Pick's disease, frontotemporal dementia, cortico-basal degeneration and progressive supranuclear palsy. Therefore, collective splicing defects in the MAPT and NMDA receptor NR1 subunit transcripts as a result of CUGBP2 gain-of-function likely contribute to the cognitive problems observed in DM1 patients.

In addition to its roles as a splicing regulator, CUGBP2 has been shown to function in post-transcriptional events such as mRNA stabilization, translational silencing, and RNA editing (Good et al., 2000; Anant et al., 2001; Mukhopadhyay et al., 2003a). In the cytoplasm, CUGBP2 has been shown to regulate the cell cycle and apoptosis through translational silencing of the COX-2 transcript (Mukhopadhyay et al., 2003a; Mukhopadhyay et al., 2003b). Sequestration of

CUGBP1 and 2 in the nucleus in DM1 is associated with a decrease of these proteins in the cytoplasm and a loss-of-function in the regulation of cytoplasmic events (Timchenko et al., 2001a). This has been associated with defects in the differentiation of muscle cells because CUGBP1 and 2 are not able to enter the cytoplasm to simulate cell cycle withdrawal. Taken together, a combination of cytoplasmic and nuclear RNA processing events are regulated by CUGBP1 and 2 and misregulation of these events by a gain-of-function in the nucleus and loss-of-function in the cytoplasm contribute to the global mRNA defects associated with the disease phenotypes of DM1.

1.7 THESIS GOALS AND OBJECTIVES

Above, I described the current state-of-progress of the splicing field in great detail. From this analysis, it is clear that several fundamental questions remain unanswered. For example, how do splicing factors recognize their cognate RNA substrates given the short and degenerate nature of SREs? Does a combination of specific RNA interactions facilitated by multiple RRMs of a single protein contribute to the high specificity and affinity of splicing factor binding? Furthermore, do protein-protein interactions between alternative splicing factors and the core splicing machinery help facilitate pre-mRNA recognition? If so, how do such interactions negatively or positively affect spliceosome assembly to adjust splicing patterns? With this in mind, how can one alternative splicing factor simultaneously function as an enhancer and silencer of splicing on different substrates? In the nervous system, how is splicing factor function adjusted in response to cell signaling to allow for plasticity in splicing control? Here, the mechanisms of splicing regulation by the CUGBP2 splicing factor are studied extensively to provide insight into some of these questions.

CUGBP2 regulates a complex network of alternative splicing events in a tissue-specific and developmentally controlled manner. Furthermore, global splicing patterns are affected by its misregulation. However, codes for splicing enhancement or silencing by this factor or the mechanisms of regulation have yet to be characterized. I initially set out to study splicing silencing and enhancement by CUGBP2 using the NMDA receptor NR1 subunit NI and CI cassette exons as model systems with three main objectives in mind. The main goals of my thesis were: (1) to identify the functionally significant sequence motifs recognized by the CUGBP2 splicing factor for splicing silencing and enhancement of the NI and CI exons, respectively, (2) to determine the mechanisms of regulation at the level of spliceosome assembly, and (3) to generate a predictive code for splicing regulation that can be applied to additional target exons. In the end, I only had time to study splicing silencing by this factor in great detail and have used information from the NI exon and other published studies to predict how this factor may also function as an enhancer of splicing. Results from these analyses provide insight into the mechanisms of target exon selection and splicing silencing by CUGBP2 and can be used as a platform for the future analysis of splicing regulation by this factor and additional splicing silencers and dual functional regulators.

The functional domains of CUGBP2 involved in splicing silencing or enhancement have been analyzed in the past (Singh et al., 2004). In these studies, truncation mutants starting at either end of the protein were used to study splicing enhancement or silencing of a CUGBP2 target exon. Results from these studies remain inconclusive because of inherent flaws in the study design. It is likely that the RRMs and divergent domain act in combination to carryout the silencing and enhancing roles of the protein. Furthermore, functional redundancies between RRMs make it difficult to determine the role of an individual RRM in the context of a truncation mutant. As a side study, I use RRM domain mutants and a biochemical approach to determine the nucleotides contacted by individual RRMs. Studies with these mutants give clues to the

topology of binding for splicing silencing by CUGBP2 and also provide insight into a potential mechanism of splicing enhancement.

2.0 MATERIALS AND METHODS

2.1 PLASMID CONSTRUCTION

The pMAL-c2x-NAPOR vector for purification of recombinant CUGBP2 was generated by PCR amplification of the mouse CUGBP2 open reading frame (ORF) (coding for amino acids 1-484) from genomic DNA and insertion into the pMAL-c2x vector (NEB) between the BamHI and HindIII restriction sites with the malE ORF encoding the maltose binding protein (MBP) tag at the amino terminus. Mutant derivatives, RRM1_2, RRM2_3, and ΔDD were generated from the pMAL-c2x-NAPOR vector by PCR amplification and reinsertion between the BamHI and HindIII restriction sites of the pMAL-c2x vector backbone. RRM1_2 contained amino acids 1-204, RRM2_3 contained amino acids 104-212 and 395-484 separated by the XhoI restriction site (nucleotides, CTCGAG; amino acids, LE), and ΔDD contained amino acids 1-204 and amino acids 392-484 separated by the XbaI restriction site (nucleotides, TCTAGA; amino acids, SR).

The mouse CUGBP2 (pcDNA4/NAPOR) and rat PTB (pcDNA4/PTB) *in vivo* expression vectors were described previously (Zhang et al., 2002). The pCI/Nova-1 *in vivo* expression vector was a gift of Robert Darnell (Dredge and Darnell, 2003) and the human pcDNA3.1/CUGBP1 (FlyLQ) expression vector was a gift of Thomas Cooper (Charlet et al., 2002a). For the pcDNA3.1/CUGBP1 vector, CUGBP1 was Flag-tagged at the amino-terminus. To generate the DUPNIwt and DUP-CUGBP2 splicing reporters, the cassette exon and flanking introns were amplified from rat genomic DNA and inserted between the ApaI and BglII restriction sites of the DUP4-1 splicing reporter (Modafferi and Black, 1997). Mutant derivatives

were generated using the QuickChange Site-Directed Mutagenesis Kit according to the manufacturer's protocol (Stratagene). The DIPm93wt splicing reporter (An and Grabowski, 2007) was used to generate DIPNIwt and mutant derivatives. DIPNIwt was generated by insertion of a 39 base pair fragment containing the 3´ splice site region of the NI exon at position -13 base pairs upstream from the test exon.

pBSDUPNI and pBSDUP-CUGBP2 wild type and mutant vectors for *in vitro* transcription were generated by PCR amplification from the DUPNI or DUP-CUGBP2 wild type or mutant splicing reporters and insertion between the HindIII and EcoRI restriction sites of the pBS- phagemid vector (Stratagene). E5-8, E5-10, and E5-15 plasmids were generated in a similar manner. The pBSAd1 plasmid for *in vitro* transcription of Ad1 pre-mRNA was described previously (Konarska et al., 1985). Plasmid sequences were confirmed by restriction digestion and DNA sequencing. For detailed plasmid maps see Appendix.

2.2 CUGBP2 PROTEIN PURIFICATION

Two-250 ml starter cultures of *DH10B Escherichia coli* transformed with the pMALc2x-NAPOR vector were grown in LB broth containing 10 μg/ml ampicillin. After 16 hours, bacteria were diluted 1:10 (500 ml to 5 L) and OD_{595} was monitored until it reached 0.5-0.6. At this time, IPTG (Isopropyl β-D-1-thiogalactopyranoside) was added at a final concentration of 0.3 mM to induce MBP-tagged CUGBP2 expression. After two hours, cells were pelleted at 3,500 rpm for 10 minutes at 4^{o}C in the IEC tabletop centrifuge. The bacteria pellet was stored at -80^{o}C overnight. The next day, the pellet was thawed and resuspended in chilled 1X Column Buffer (20mN Tris-HCl, 200 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol, pH 7.4). For every 1 mg pellet, 10 ml Column Buffer was added. Bacteria were sonicated with 8, 15 second pulses on setting 3 while on ice. Sonicated bacteria were pelleted at 9,000 xg for 30 minutes at 4^{o}C in the Sorvall

centrifuge to obtain a crude extract (supernatant). Using the BioRad Protein Assay and a bovine serum albumin (BSA) standard curve, the protein concentration of the crude extract was estimated at OD_{280}. At this time 100 μl crude extract (input) was set aside for SDS-PAGE analysis. The crude extract was diluted to 3 mg/ml total protein using 1X Column Buffer.

The amylose column was prepared by pouring amylose resin into a 2.5 x 20 cm column until the packed column volume was 80 ml. The column was washed with 10X Column Buffer, followed by 1X Column Buffer until the pH of the flow through was 7.4. Crude extract was loaded onto the column and flow through was collected for SDS-PAGE analysis. The column was washed with 12 column volumes (960 ml) of 1X Column Buffer overnight. Note that the column flow rate was ~2 ml/minute. MBP-tagged CUGBP2 was eluted from the column by the addition of 200 ml 1X Column Buffer containing 10 mM maltose. Fractions were collected and assayed with the BioRad Protein Assay as before. Input, flow through, and several fractions were separated on a 10% SDS-polyacrylamide gel and stained with Coomassie Blue for protein analysis (Figure 5A).

The most concentrated fractions were pooled together and treated with Factor Xa protease (NEB) for 48 hours at 4^{o}C to separate the MBP tag from the CUGBP2 protein. Factor Xa was added at a ratio of 1:150 (for example: 44 mg protein/150 = 293 μl Factor Xa). SDS-PAGE, Coomassie Blue gel analysis was carried out to verify protein cleavage (Figure 5B). The MBP/CUGBP2 protein mixture was then dialyzed against 100X volume DLS Buffer (20 mM Tris-HCl, 25 mM NaCl, 2 mM dithiothreitol (DTT), pH, 8.0) at 4^{o}C using dialysis tubing with a molecular weight cut off (MWCO) of 3,500. Tubing was boiled in 1 L water containing ~2.5 g $NaHCO_3$ for 20 minutes and soaked in water overnight before use. After 3 hours, the DLS Buffer was replaced and the mixture was dialyzed for an additional 3 hours.

The diethylaminoethyl (DEAE)-sepharose column was prepared by loading the DEAE-sepharose resin into a 2.5 x 20 cm column until the packed column volume was 45 ml. The

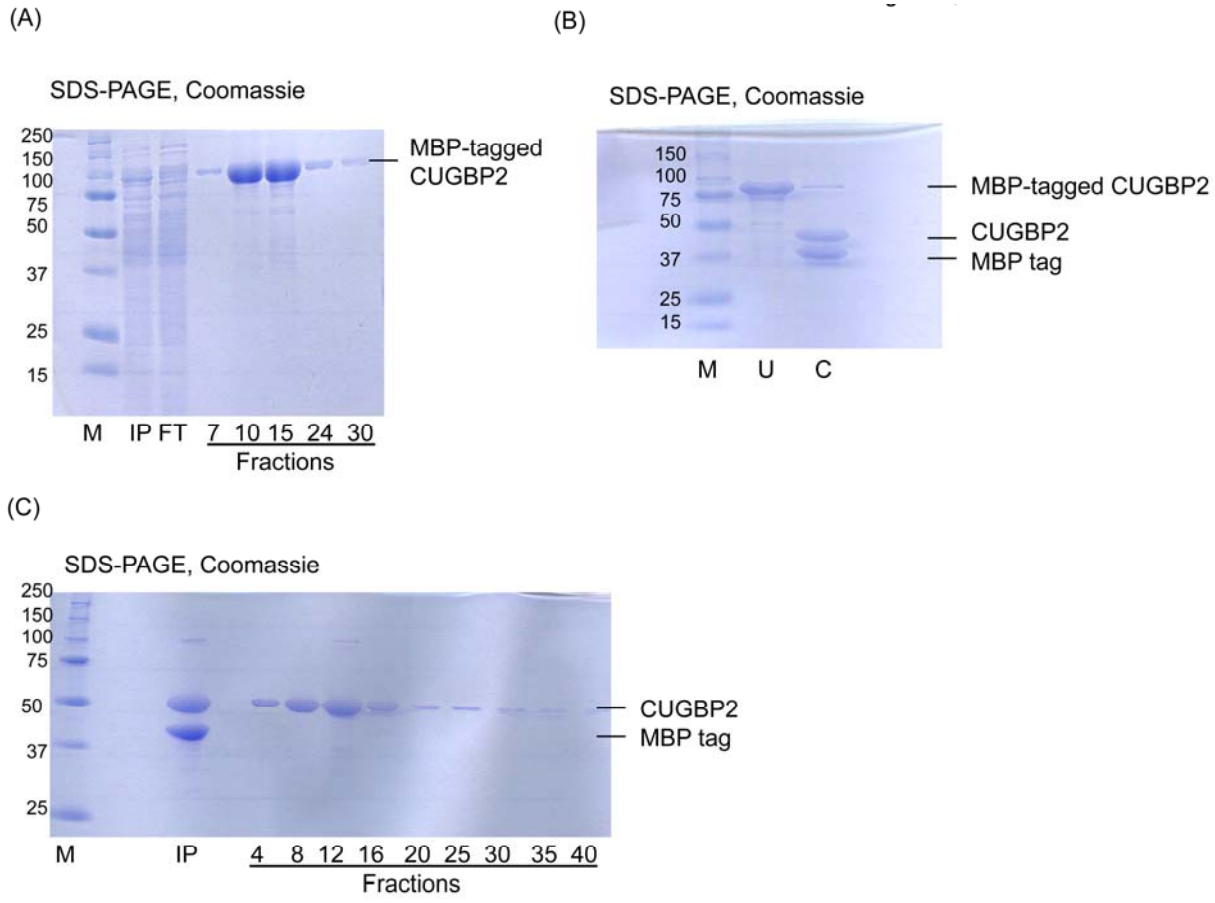


Figure 5. Recombinant CUGBP2 protein purification. (A) Amylose affinity purification of MBP-tagged CUGBP2. SDS-PAGE, Coomassie Blue gel analysis of total protein from input (IP), flow through (FT), and eluted fractions 7, 10, 14, 24, and 30. Precision Plus Protein standard, M. Molecular weight in kDa at left. (B) Factor Xa cleavage of MBP-tagged CUGBP2. SDS-PAGE, Coomassie Blue gel analysis of MBP-tagged CUGBP2 uncut (U) or cut (C) with Factor Xa. The relative positions of MBP-tagged CUGBP2, CUGBP2, and the MBP tag are shown at right. (C) DEAE sepharose ion-exchange chromatography to separate purified CUGBP2 from the MBP tag. SDS-PAGE, Coomassie Blue gel analysis of IP and flow through fractions 4, 8, 12, 16, 20, 25, 30, 35, and 40. Molecular weights are as follows: MBP-tagged CUGBP2, 91 kDa; CUGBP2, 51 kDa; MBP-tag, 40 kDa.

column was washed with 10X DLS Buffer, followed by 1X DLS Buffer until the pH of the flow through was 8.0. The dialyzed protein mixture was passed through the DEAE-sepharose column and 6 ml fractions were collected at a rate of ~2.5 ml/minute while washing. The protein concentration in each fraction was estimated using the BioRad Protein Assay and select fractions were analyzed by SDS-PAGE (Figure 5C). The most concentrated fractions were combined and concentrated using an Amicon ultrafiltration unit with a MWCO of 30,000 (Rosenberg lab) to obtain a final concentration of at least 0.54 mg/ml or 10 μM. The protein solution was separated into 100 μl aliquots, flash frozen in liquid nitrogen, and stored at -80°C until use. CUGBP2 mutant proteins RRM1_2, RRM2_3, and ΔDD, were purified in a similar manner.

2.3 *IN VITRO* TRANSCRIPTION

To generate the DNA template for *in vitro* transcription, plasmid DNA was purified from *DH5α E. coli* using the QIAprep Spin Miniprep Kit (Qiagen) and digested in a 200 μl reaction containing 10 μg plasmid DNA, 20 μl 10X NEBuffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM DTT, pH 7.9) and 2 μl EcoRI (40 U) at 37°C overnight. Plasmids used for *in vitro* transcription include: E5-8, E5-10, E5-15, pBSDUPNI, pBSDUPNIm1, pBSDUPNIm3, pBSDUPNIm1,2,3, pBSAd1, and pBSDUP-CUGBP2. Complete digestion was verified on a 1% agarose gel next to uncut plasmid. Uncut plasmids mostly contained supercoiled DNA with less than 10% nicked circle DNA. Cut plasmid DNA was purified by phenol/chloroform extraction followed by ethanol precipitation. All DNA pellets were washed with 70% ethanol and resuspended in 10 μl sterile ddH_2O to obtain a final DNA concentration of ~1 μg/μl. DNA templates were stored at -20°C until use.

In vitro transcription was carried out in a 25 μl reaction containing 1X transcription buffer (40 mM Tris-HCl pH 8.0, 8 mM $MgCl_2$, 50 mM NaCl, 2 mM spermidine), 400 μM ATP, 400 μM CTP, 400 μM UTP, 400 μM GTP, 30 mM DTT, 1 μg DNA template, 10 U T3 RNA polymerase (Stratagene), and 10 U RNasin (Promega). For "hot" transcription, 25 μCi α-^{32}P-UTP (Perkin Elmer) was added to the reaction and unlabeled UTP was at a final concentration of 300 μM. Importantly, α-^{32}P-UTP was less than two weeks old. Transcription reactions were incubated at 37°C for 30 minutes before addition of 1.25 U RQ1 RNase-Free DNase (Promega) and incubation at room temperature for an additional 10 minutes. Reactions were then loaded onto a Sephadex G-25 column generated in 1X New Buffer (50 mM Hepes pH 7.0, 50 mM KCl, 2 mM magnesium acetate, 0.1 mM EDTA) with a packed column volume of 3-5 ml. Columns were centrifuged at 2,000 rpm at 4°C for 4 minutes in the tabletop IEC centrifuge. Flow through containing newly transcribed RNA was collected in a 15 ml conical tube and transferred to a 1.5 ml eppendorf tube. Free nucleotides were retained in the column. RNA was ethanol precipitated in 1/10 volume 3 M sodium acetate, 1/50 volume 10 mg/ml glycogen, and 2.5 volumes 100% ethanol on crushed dry ice for 10 minutes followed by centrifugation at 13,000 xg for 10 minutes at 4°C. RNA pellets were washed with 70% ethanol and resuspended in 25 μl sterile ddH_2O. One-thousand cpm radiolabeled RNA was separated on a 10% polyacrylamide 7 M urea gel, dried under vacuum, and visualized on a BAS-2500 phosphoimaging system (Fujifilm) to verify RNA integrity and size. RNA was stored at -20°C for less than one week before use.

2.4 NITROCELLULOSE FILTER BINDING ASSAY

For nitrocellulose filter binding assays, a modified version of the protocol described by Draper et al. (1988) was used. RNA substrates were ^{32}P-UTP-labeled by *in vitro* transcription and used at

a final concentration of 1000 cpm/μl (20-100 nM). RNA was heated at 85°C for 5 minutes and then cooled to 37°C for 5 minutes to remove long-range secondary structures. Serial dilutions of protein (0.1-1.0 μM final concentration) were prepared on ice in binding buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1 mg/ml tRNA, 2 mM DTT, 20 units RNasin (Promega)) in a final volume of 199 μl. Protein solutions were warmed to 37°C for 5 minutes before adding 1 μl of diluted RNA (final RNA concentration 100-500 pM). RNA-protein complexes were assembled in triplicate at 37°C for 30 minutes before filtration through 25 mm BA85 filters backed by DE81 filters in a Millipore 1225 vacuum manifold. Filters were separated and dried at room temperature overnight. The cpm retained on the BA85 filter corresponded to RNA bound to protein and the cpm retained on the DE81 filter corresponded to free RNA. Bound/total RNA was plotted as a function of increasing protein concentration using KaleidaGraph Synergy Software and data were fit to a hyperbola to estimate the dissociation constant (K_d) according to the equation bound/total RNA = [CUGBP2]/([CUGBP2] + K_d).

2.5 ELECTROPHORETIC MOBILITY SHIFT ASSAY

Binding reactions were prepared as described above for the nitrocellulose filter binding assay except 50 μl reactions were used and protein-RNA complexes were formed on ice for 1 hour followed by the addition 3 μl native loading dye (50% glycerol, 0.1% xylene cyanol, 0.1% bromphenol blue). For electrophoretic mobility shift assays (EMSAs), 0.1-4.0 μM final concentration protein was used. Free RNA and RNA-protein complexes were separated on a 5% native polyacrylamide gel (80:1 acrylamide to bis ratio) cast in 50 mM tris-glycine buffer at 250 volts for 2 hours in the cold room. Gels were dried under vacuum and visualized with a BAS-2500 phosphoimaging system (Fujifilm) and ImageGauge software.

2.6 CMCT MODIFICATION FOOTPRINTING

RNA footprinting assays were modified from (Brunel and Romby, 2000). RNA-protein complexes were assembled as for the nitrocellulose filter binding assay for 30 minutes at 37°C with 0.18 μM unlabeled RNA substrate and purified protein in a final volume of 50 μl. Each sample was then combined with an equal volume of 42 mg/ml 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimidemetho-*p*-toluene-sulfonate (CMCT) in BMK buffer (80 mM potassium borate pH 8.0, 100 mM NH_4Cl) without mixing and chemical modification was carried out at 37°C for 7 minutes. Reactions were terminated by ethanol precipitation by the addition of 2.5 volumes 100% ethanol, 1/10 volume 3 M sodium acetate, and 1/50 volume 10 mg/ml glycogen followed by mixing, incubation on crushed dry ice for 10 minutes, and centrifugation at 13,000 xg for 10 minutes at 4°C. Recovered RNA pellets were resuspended in 200 μl proteinase K mix (10 mg/ml proteinase K, 50 mM Tris-HCl pH 7.5, 10 mM EDTA, 10 mM NaCl, 0.2% SDS) and incubated at room temperature for 10 minutes. Proteinase K treatment was followed by phenol/chloroform extraction and ethanol precipitation as before. Modified RNA pellets were washed with 70% ethanol and resuspended in 10 μl sterile ddH_2O.

Primer extension was carried out in a reaction containing 5 μl modified RNA, 1 μl of 1 μM 5′^{32}P-labeled primer (see below), 1X First Strand Buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$), 10 mM DTT, 0.5 mM dNTPs, 10 U RNasin (Promega), and 100 U Superscript II Reverse Transcriptase (Invitrogen) in a final reaction volume of 10 μl at 42°C for 50 minutes. RNA was recovered by ethanol precipitation by the addition of 45 μl 0.3 M sodium acetate, 150 μl 100% ethanol, and 1 μl 10 mg/ml glycogen and incubation on crushed dry ice for 15 minutes followed by centrifugation at 13,000 xg for 15 minutes at 4°C. RNA pellets were washed with 70% ethanol and resuspended in 6 μl formamide loading buffer (97% deionized formamide, 0.1% xylene cyanol, 0.1% bromphenol blue). Sequencing ladders were generated

from the parent plasmid using Thermo Sequenase Cycle Sequencing Kit (USB) according to manufacturer's protocol. cDNA was heated to 85^{o}C for 2 minutes and 2 μl were separated on a 10% polyacrylamide 7 M urea gel at 45 watts for 2 hours and dried under vacuum. Gel images were recorded on a BAS-2500 phosphoimaging system (Fujifilm) and analyzed with ImageGauge software.

The JBE5-2 (5′-ATAGGACAGTTGGTCGAGGT-3′) and CUGBP2-PE (5′-TCTGTGCCATTGCCCTGGTAG-3′) primers were 5′^{32}P-labeled with T4 polynucleotide kinase (PNK) in a 25 μl reaction containing 20 μM PAGE-purified primer, 10 U T4 PNK (NEB), 1X PNK buffer (70 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, pH 7.6), and 150 μCi γ-^{32}P-ATP (Perkin Elmer). Reactions were incubated at 37^{o}C for 30 minutes then loaded onto a Sephadex G-25 column with a total packed column volume of 3-5 ml. Columns were centrifuged at 2,000 rpm at 4^{o}C for 4 minutes in the tabletop IEC centrifuge. Flow through containing labeled primer was collected in a 15 ml conical tube and transferred to a 1.5 ml eppendorf tube. Free nucleotides were retained in the column. Primers were ethanol precipitated in 1/10 volume 3 M sodium acetate, 1/50 volume 10 mg/ml glycogen, and 2.5 volumes 100% ethanol on crushed dry ice for 10 minutes followed by centrifugation at 13,000 xg for 10 minutes. Primer pellets were washed with 70% ethanol and resuspended in 20 μl sterile ddH_2O to obtain a final concentration of ~1 μM. Note that JBE5-2 was used for primer extension of the E5-10 RNA substrate and CUGBP2-PE for the DUP-CUGBP2 RNA substrate. The corresponding primers were also used for the generation of sequencing ladders from the parent plasmids.

2.7 *IN VIVO* SPLICING REPORTER ASSAY

C2C12 and N18TG2 cells were grown in DMEM (Cellgro), 10% (v/v) fetal bovine serum (FBS). Twenty-four hours prior to transfection 1.5 x 10^5 C2C12 cells or 2 x 10^5 N18TG2 cells were seeded on 35 mm plates to achieve 60-80% confluency. For transient co-transfection, 1 μg pcDNA4 His/Max vector backbone or 1 μg pcDNA4/NAPOR, pC1/Nova-1, pcDNA4/PTB, or pcDNA3.1/CUGBP1 expression vector and 0.25 μg splicing reporter (DUPNIwt, DIPNIwt, DUP-CUGBP2, or mutant derivatives) were mixed with 250 μl Opti-MEM (Gibco) followed by addition of an equal volume of Opti-MEM mixed with 2.5 μl Lipofectamine 2000 (Invitrogen) and incubated at room temperature for 20 min. Media on cells was replaced with 1.5 ml DMEM, 10% (v/v) FBS prior to transfection. Typically 20-30% of cells were transfected and both plasmids entered the same cell. Total RNA was isolated 36 hours after transfection using TRIZOL reagent (Invitrogen) according to manufacturer's protocol.

Total RNA was reverse transcribed in a reaction containing 2 μg total RNA, 0.5 μg random primers (Promega), 0.5 mM dNTPs, 10 mM DTT, 1X First Strand Buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$), and 200 U MMLV Reverse Transcriptase (Invitrogen) in a final volume of 20 μl. Reverse transcription (RT) was carried out at 37°C for 1 hour and terminated by heat inactivation at 65°C for 10 minutes. Note that the random primers were annealed to the total RNA in water by incubation at 65°C for 10 minutes followed by slow cooling to room temperature for 10 minutes prior to reverse transcription.

Polymerase chain reaction (PCR) was carried out in 10 μl reactions containing 1 μl of the reverse transcription reaction, 0.1 μM of each primer, 2 mM $MgCl_2$, 0.2 mM dNTPs, and 2.5 U Taq DNA polymerase (Promega). Forward primers were 5′-^{32}P-labeled as described above for the CMCT modification footprinting assay. The BGJB1 forward primer (5′-TGG TGCATCTGACTCCTGAGG-3′) and BGJB2 reverse primer (5′-CTGGGTCCAAGGGTAGA

CCAC-3´) were used for amplification of RNA from the DUPNI and DUP-CUGBP2 wild type and mutant splicing reporters and the GFP-f forward primer (5´-AGTGCTTCAGCCGCTACCC-3´) and GFP-r reverse primer (5´-GTGTCGCCCTCGAACTTCACC-3´) were used for amplification of RNA from the DIPNI wild type and mutant splicing reporters. Cycling parameters were adjusted to give amplification in the linear range. Conditions were as follows: initial denaturation at 94^{o}C for 2 minutes followed by denaturation at 94^{o}C, 1 minute; annealing at 60^{o}C, 1 minute; and elongation at 72^{o}C, 1 minute for 22 cycles with a final elongation step at 72^{o}C for 10 minutes. PCR samples were resolved on 6% polyacrylamide 5 M urea gels at 30 watts for 1 hour and 15 minutes and dried under vacuum. Gel images were captured using a BAS-2500 phosphoimaging system (Fujifilm) and data were quantified using Image Gauge software. The profile mode was used to subtract background and calculate percent exon inclusion of individual lanes.

2.8 SDS-PAGE AND WESTERN BLOTTING

Unless otherwise stated, whole cell lysates were used for Western blot analysis. Whole cell lysates were prepared by first growing N18TG2, HEK293T, or C2C12 cells in a 10 cm dish with DMEM growth media (Cellgro) supplemented with 10% FBS. Once cells reached 90% confluency, media was removed and cells were scraped off the plate in 200 μl SDS-RIPA buffer (3 parts RIPA buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton x100), 1 part 2X laemmli (100 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS)). Cells were transferred to a 1.5 ml eppendorf tube and sonicated on setting 3 for 10 seconds on ice. Cell lysates were flash frozen in liquid nitrogen and stored at -80^{o}C until use. Note that I did not check the concentration of cell lysates because they contain SDS. Therefore a loading control was used to compare protein levels.

For SDS-PAGE analysis, SDS-polyacrylamide gels were prepared with a 4% stacking gel mixture cast on top of 10-12% laemmli separating gel (37.5:1 acrylamide to bis ratio). Protein samples were prepared by first spinning out cell debris at 10,000 rpm for 2 minutes at 4^{o}C and then adding 6X SDS-loading buffer (1.0 M Tris-HCl pH 6.8, 30% glycerol, 10% SDS, 600 mM DTT, 0.012% bromphenol blue) to a final concentration of 1X followed by boiling at 100^{o}C for 5 minutes. Gels were run in 1X Running Buffer (400 mM glycine, 50 mM Tris-HCl, 0.1% SDS) until the bromphenol blue ran out of the gel into the bottom reservoir. Gels were either stained with Coomassie Blue to assay total proteins or transferred to a nitrocellulose membrane for Western blotting.

For Western blot analysis, proteins were transferred from the polyacrylamide gel to a Immobilin-P nitrocellulose membrane (Millipore) at 15 volts overnight. Transfer was confirmed by Ponceau S staining. Membranes were blocked in PBST (1X phosphate buffered saline (PBS), 0.1% Tween 20) containing 5% milk for 1 hour at room temperature while shaking, followed by a quick wash in PBST. Membranes were then incubated in primary antibody for 1 hour while rotating at 4^{o}C. Primary antibodies were diluted in PBST + 1% BSA as follows: αCUGBP2 (1H2), 1:1000; αCUGBP1/2 (3B1), 1:1000; αhnRNPA1 (9H10), 1:2000; αXpress, 1:3000. Importantly, the 1H2 antibody is highly specific for CUGBP2 (Ladd et al., 2005). Membranes were then subjected to four 10-minute washes in PBST at room temperature. After washing, membranes were incubated in secondary antibody (1:20,000 dilution of αmouse-HRP (Roche) in PBST + 1% BSA) for 1 hour while shaking in the cold room. Membranes were then washed as before and treated with Western Lightning Plus-Enhanced Chemiluminescence Substrate (Perkin Elmer) for 1 minute before development with the LAS-3000 Intelligent Dark Box (Fujifilm) for 1-5 minutes and analysis with ImageGauge software.

2.9 CULTURING, TRANSFECTION, AND PROTEIN ANALYSIS OF PRIMARY CORTICAL NEURONS

Two-dozen fresh E18 rat cerebral cortexes (hemispheres) were ordered from Zivic Miller in sterile 1X PBS. Immediately after arrival, tissue was gently rinsed three times with 20 ml sterile 1X PBS (Ca^{2+}, Mg^{2+} free). Tissue was digested in 20 ml 1X trypsin diluted in trypsin media (MEM (Gibco), 0.1% glucose (sterile filtered), 2 mM L-glutamine (sterile filtered, light sensitive), 20 mM HEPES (Sigma)) at 33-36^{o}C for 20 minutes, tapping the tube gently every 5 minutes. After digestion, trypsin was removed and tissue was rinsed three times with 20 ml fresh trypsin media. For rinsing, media was added, tissue was allowed to settle to the bottom of the tube for 3 minutes, and then media was removed. Digested tissue was resuspended in 8 ml trypsin media, transferred to a sterile 10 ml syringe, and passed through an 18 gauge needle 8-10 times into a 10 cm dish to homogenize the tissue. Importantly, tissue was never aspirated into the syringe it was only passed out of the needle. After pipetting up and down five times, digested tissue was transferred to a 50 ml conical tube. The 10 cm dish was rinsed two times with 10 ml trypsin media, which was also added to the conical tube. Homogenized tissue was allowed to sit at room temperature for 5 minutes to allow connective tissue and debris to settle to the bottom of the tube. The supernatant was transferred to two 15 ml falcon tubes. The pellet containing cell debris was rinsed with 30 ml trypsin media and the supernatant of this wash was transferred to two more 15 ml falcon tubes. The four tubes were centrifuged at 800-900 rpm for 3-4 minutes (setting 4, tissue culture room) to pellet the cells. Cell pellets were washed two times with 5 ml growth media (Neurobasal media (Gibco) supplemented with 0.5 mM L-glutamine (sterile filtered, light sensitive) and 2% B-27 supplement (Invitrogen)). While washing, pellets were combined into one tube. The final cell suspension was turbid without clumps and was resuspended in 25 ml growth media containing 200 ng/ml gentamicin. Cells were counted on a hemacytometer and 6-10 x 10^5 cells were plated per well of a 6 well dish and 3 x10^6 cells were

plated per 10 cm dish. Dishes were pre-coated with poly-D-lysine (BD Biosciences). If coverslips were used, three poly-D-lysine-coated coverslips (BD Biosciences) were placed into each well of an uncoated 6 well dish and 6-10 x 10^5 cells were plated per well. Half of the growth media was replaced after 4 days and primary cortical cultures were incubated in a 37^{o}C tissue culture incubator adjusted to 6% CO_2. Primary cortical cultures contained mostly neurons and few supporting glial cells.

For splicing reporter analysis in primary cortical cultures, 1 μg pcDNA4 His/Max vector backbone and 0.25 μg splicing reporter (DUPNIwt or mutant derivatives) were mixed with 250 μl Opti-MEM (Gibco) followed by addition of an equal volume of Opti-MEM mixed with 2.5 μl Lipofectamine 2000 (Invitrogen) and incubated at room temperature for 20 min. Media on cells was replaced with 1.5 ml Opti-MEM prior to transfection. Typically 10-15% of cells were transfected. For this experiment, the pcDNA4 His/Max vector backbone was used as carrier DNA. After 5 hours, the Opti-MEM transfection mix was replaced with growth media containing 200 ng/ml gentamicin. Total RNA was isolated 24 hours after transfection using TRIZOL reagent (Invitrogen) according to manufacturer's protocol. RT-PCR and gel analysis were carried out as described in section 2.7.

Nuclear lysates of primary cortical cultures were prepared from six 5-day-old 10 cm dishes. Cells were harvested by scraping them off the plate in 1 ml 1X PBS per plate, followed by centrifugation at 800-900 rpm for 3-4 minutes (setting 4, tissue culture room). Cells from all plates were pooled into a 15 ml conical tube, resuspended in 4 ml chilled Buffer I (0.32 M sucrose, 3.0 mM $CaCl_2$, 2.0 mM magnesium acetate, 0.1 mM EDTA, 10 mM Tris-HCl pH 8.0, 0.5% NP-40, 1.0 mM DTT), and incubated on ice for 30 minutes while douncing 9-12 times followed by vortexing every 10 minutes. Glass homogenizers were baked and chilled prior to use. Cell lysis was verified by trypan blue staining. The lysed cell mixture was overlaid onto 3.5 ml chilled Buffer II (2M sucrose, 5 mM magnesium acetate, 0.1 mM EDTA, 10 mM Tris-HCl pH

3.0, 1 mM DTT) in an SW41 ultracentrifuge tube (Beckman 331372). Note that protease and phosphatase inhibitors can be added to Buffer I and Buffer II to decrease protein degradation. Cell lysates were ultracentrifuged at 30,000 xg (~15,500 rpm) for 45 minutes at 4^{o}C. Supernatants containing cell debris and cytoplasmic lysate were carefully removed. Nuclei were lysed in 150 μl SDS-RIPA buffer (3 parts RIPA buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton x100), 1 part 2X laemmli (100 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS)) for 30 minutes on ice. Nuclear lysates were transferred to a 1.5 ml eppendorf tube on ice, sonicated on setting 3 for 10 seconds, flash frozen in liquid nitrogen, and stored at -80^{o}C until use. SDS-PAGE and Western blotting were carried out as described in section 2.8.

2.10 IMMUNOFLUORESCENCE

Immunofluorescence analysis was carried out with primary cortical cultures grown on poly-D-lysine-coated coverslips. For all steps, it is important that coverslips are always completely covered with liquid and do not dry out. Coverslips were washed two times with 1X PBS and cells were fixed with 3.7% formaldehyde in 1X PBS for 7 minutes at room temperature. Note that all washes were carried out for 3 minutes each. After fixation, coverslips were washed three times with 1X PBS and fixed cells were permeabilized with 0.1% Triton X100 in 1X PBS for 1 minute at -20^{o}C followed by 5 minutes at 4^{o}C. Coverslips were then washed three times with Incubation Buffer (1X PBS, 1% BSA, 0.1% gelatin, 0.01% sodium azide). Note that the Incubation Buffer was prepared by first dissolving gelatin in PBS at 65^{o}C before the addition of BSA and sodium azide. Cells were then blocked in Incubation Buffer containing 5% goat serum for 3 hours at room temperature, followed by three washes in Incubation Buffer. Coverslips were then transferred from the 6 well dish to a humidified chamber (tissue culture dish with parafilm on the

bottom of the dish and a piece of damp filter paper on the inside of the lid) for the rest of the incubations.

The appropriate dilution of primary antibody was prepared in Incubation Buffer and pipetted onto the coverslips so that the surface of the coverslips were completely covered with liquid. Cells were incubated in primary antibody at 4°C overnight. It is important to be extra careful that coverslips do not dry out during this step. For co-immunofluorescence, both primary antibodies were added at the same time. Primary antibody dilutions were as follows: mouse αCUGBP1/2 (3B1), 1:500; mouse αCUGBP2 (1H2), 1:500; rabbit αGFAP (Chemicon), 1:1000; rabbit αNeurofilament L (Chemicon), 1:1000. GFAP (Glial Fibrillary Acidic Protein) is a glial cell-specific marker that is only expressed in the cytoplasm and Neurofilament L (Light) is a neuron-specific marker that is expressed throughout the cell and nucleus. After incubation in primary antibody, coverslips were washed three times with Incubation Buffer as before.

Next, the appropriate dilution of secondary antibody in Incubation Buffer was pipetted onto the coverslips and incubated for 2 hours at 4°C in the dark. Secondary antibody dilutions were as follows: Alexa Fluor 488 goat αmouse-IgG (Molecular Probes), 1:500; Alexa Fluor 568 goat αrabbit-IgG (Molecular Probes), 1:500. All subsequent steps were carried out in the dark. Coverslips were washed three times with Incubation Buffer and then incubated in DAPI Nucleic Acid Stain (Invitrogen) diluted to 300 nM in 1X PBS for 3 minutes at room temperature. Coverslips were subjected to three final washes in 1X PBS and mounted cell-side down onto glass slides with 2-5 μl mounting media (see below). Cells were visualized with the confocal microscope and 60X oil immersion lens and slides were stored in a dark box at -20°C for future reference.

To prepare mounting media, 2.4 g Mowiol 4-88 (Calbiochem) and 6.0 g glycerol (liquid) were diluted in 6.0 ml sterile ddH_2O and incubated at room temperature overnight. The next day, 12.0 ml 0.2 M Tris-HCl was added and the mixture was heated to 50°C for 10 minutes. The

mixture was then centrifuged at 5000 xg for 15 minutes. The supernatant was transferred to a new tube and 0.6 g DABCO (Sigma) was added. The mixture was then aliquoted and stored at -20^{o}C until use.

2.11 PREPARATION OF HELA NUCLEAR EXTRACTS

HeLa nuclear extract was prepared as described by Dignam et al. (1983). Six liters of HeLa spinner cells were grown to a cell density of 3 x 10^5 cells/ml (~1.8 x 10^9 cells total). JMEM growth media (Sigma) was supplemented with 20 mM Hepes and 24 mM sodium bicarbinate, adjusted to pH 7.0, sterile filtered, and then supplemented with 5% horse serum before use. Cells were grown in sterile spinner flasks at 37^{o}C. The day before the prep, all buffers were prepared without DTT or glycerol and chilled in the cold room, centrifuge rotors and buckets were chilled, and glass homogenizers were baked, cooled to room temperature, and chilled in the cold room overnight.

Using 1 L centrifuge bottles, cells were pelleted at 1000 rpm for 10 minutes at 4^{o}C in the IEC tabletop centrifuge. The supernatant was carefully discarded and this step was repeated until all cells were condensed into one bottle. Note that cells were kept on ice at all times and the cell pellet was very loose. Next, the cell pellet was condensed further into one 50 ml conical tube by rinsing with 1X PBS. At this time, the packed cell volume (PCV) was 7.5 ml. The pellet was resuspended in 5 PCV (37.5 ml) Buffer A (10 mM Hepes pH 7.6, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT) and incubated on ice for 10 minutes. Cells were then centrifuged at 1500 rpm for 10 minutes at 4^{o}C in the IEC tabletop centrifuge and the cell pellet was resuspended in 2 PCV (15 ml) Buffer A. Next, the cells were dounced with a tight pestle A on ice until ~90% of cells were lysed as verified by trypan blue staining.

Lysed cells were then quickly transferred to a 50 ml round bottom tube and centrifuged at 2000 rpm for 10 minutes at 4^{o}C in the Sorvall centrifuge. The supernatant was removed, set aside on ice, and used to generate the S100 extract (see below). Note that the pellet is not very distinct from the supernatant. The pellet was centrifuged at 30,500 xg for 20 minutes at 4^{o}C in the Sorvall centrifuge. The supernatant from this step was added to the supernatant from the last step.

The nuclear pellet was gently resuspended in 5.4 ml Buffer C (20 mM Hepes pH 7.5, 25% glycerol, 0.42 M KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT) by swirling with a pipette tip and then rotated at 4^{o}C for 30 minutes for nuclear extraction. Note that Buffer C was added at 3 ml/10^9 cells. The nuclei were then centrifuged at 30,500 xg for 30 minutes at 4^{o}C in the Sorvall centrifuge to obtain the nuclear extract (supernatant). Nuclear extracts were dialyzed against 1 L Buffer D (20 mM Hepes pH 7.6, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT). Dialysis tubing was 18 mm wide and had a MWCO of 3500. After 2 hours, the Buffer D was replaced and the extract was dialyzed for an additional 2 hours. The final protein concentration was ~6 mg/ml. Nuclear extracts were divided into 200 μl aliquots, flash frozen in liquid nitrogen, and stored at -80^{o}C until use.

For preparation of the S100 extract, 0.11 volume Buffer B (0.3 M Hepes pH 7.5, 1.4 M KCl, 30 mM $MgCl_2$) was mixed with the S100 fraction and centrifuged at 100,000 xg in the ultracentrifuge using the SW41 rotor for 1 hour. The supernatant was flash frozen in liquid nitrogen and stored at -80^{o}C for future use.

2.12 *IN VITRO* SPLICING ASSAY

pBSDUPNI, mutant derivatives, or pBSAd1 were digested with EcoRI for *in vitro* transcription in the presence of α-^{32}P-UTP. *In vitro* splicing reactions contained 20 mM Hepes pH 7.5, 44%

HeLa nuclear extract (prespun at 10,000 rpm for 2 minutes at 4°C), 60 mM KCl, 2.2 mM $MgCl_2$, 5 mM creatine phosphate, 1.5 mM ATP, 40,000 cpm pre-mRNA, and 0, 0.8, or 1.6 μM CUGBP2, RRM1_2, RRM2_3, or ΔDD protein in a final volume of 25 μl. Reaction mixtures were pre-incubated at 30°C for 10 minutes prior to the addition of pre-mRNA. Parallel splicing reactions were prepared in the absence of ATP, creatine phosphate, or recombinant protein as negative splicing controls. Note that these conditions were adjusted from the standard splicing reaction, which contains 2 mM $MgCl_2$ and 44 mM KCl. It is important that the concentration of KCl in the HeLa nuclear extract is considered when calculating the final concentration of KCl in the splicing reaction. Splicing was carried out for 20-60 minutes at 30°C. Splicing was stopped by the addition of 25 μl proteinase K mix (10 mg/ml proteinase K, 50 mM Tris-HCl pH 7.5, 10 mM EDTA, 10 mM NaCl, 0.2% SDS) and incubation at 30°C for an additional 10 minutes. Reactions were then placed on ice until all samples were ready. Samples were adjusted to 1X HHS (10X: 0.2 M Hepes pH 7.4, 2.5 M sodium acetate, 10 mM EDTA) and subjected to phenol/chloroform extraction. RNA was ethanol precipitated by the addition of 1/10 volume 3 M sodium acetate, 1/50 volume 10 mg/ml glycogen, and 2.5 volume 100% ethanol and incubation on crushed dry ice for 10 minutes followed by centrifugation at 13,000 xg for 10 minutes. RNA pellets were washed with 70% ethanol and resuspended in 6 μl formamide loading buffer (97% deionized formamide, 0.1% xylene cyanol, 0.1% bromphenol blue). Samples were heated to 65°C for 5 minutes before half of the sample was separated on an 8% polyacrylamide 7 M urea gel at 400 volts for 1.5 hours. Gels were soaked in water for 1 hour (3 changes), dried completely under vacuum, and visualized with the BAS-2500 phosphoimaging system (Fujifilm) and ImageGauge software.

2.13 BRANCHPOINT MAPPING

pBSDUPNI or mutant derivatives were digested with EcoRI and subjected to *in vitro* transcription without α-^{32}P-UTP. *In vitro* splicing assays and recovery of splicing products and intermediates were carried out as described above. RNA pellets were resuspended in 5 μl sterile ddH_2O. Branchpoints were detected by primer extension in a reaction containing 5 μl recovered RNA, 200 U MMLV Reverse Transcriptase (Invitrogen), 10 mM DTT, 1 mM dNTPs, 1X First Strand Buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$), and 50 nM 5´^{32}P-labeled JBE5-2 primer in a total reaction volume of 20 μl. Primer and 5´-end labeling were as for the CMCT modification footprinting assays (section 2.6). Reactions were incubated at 37^{o}C for 30 minutes and were terminated by ethanol precipitation by the addition of 2.5 volumes 100% ethanol, 1/10 volume 3 M sodium acetate, and 1/50 volume 10 mg/ml glycogen followed by incubation on crushed dry ice for 10 minutes and centrifugation at 13,000 xg for 10 minutes at 4^{o}C. Pellets were washed with 70% ethanol and resuspended in 6 μl formamide loading buffer (97% deionized formamide, 0.1% xylene cyanol, 0.1% bromphenol blue). Samples were heated to 85^{o}C for 2 minutes and 2 μl of the sample was separated on an 8% polyacrylamide 7M urea gel next to a sequencing ladder. Sequencing ladders were generated from the pBSDUPNI plasmid using the JBE5-2 primer and the Thermo Sequenase Cycle Sequencing Kit (USB) according to manufacturer's protocol. Gels were run at 45 watts for 2 hours, dried under vacuum, and visualized with the BAS-2500 phosphoimaging system (Fujifilm) and ImageGauge software. Primer extension of Ad1 splicing intermediates and products was carried out with the Ad1-PE primer (5´-ACTGGAAAGACCGCGAAGAG-3´).

For the debranching reaction, HeLa S100 extract was pretreated with micrococcal nuclease by incubating 200 μl S100 extract with 1 mM $CaCl_2$ and 6 U micrococcal nuclease for 30 minutes at 30^{o}C. Reactions were stopped by the addition of EGTA to a final concentration of

4 mM. The debranching mix contained 20 mM Hepes pH 7.6, 20% glycerol, 100 mM KCl, 8 mM EDTA, 0.5 mM DTT, and 150 μl micrococcal nuclease-treated S100 HeLa extract in a final volume of 750 μl. For the debranching reaction, 5 μl RNA from the *in vitro* splicing reaction was incubated with 25 μl debranching mix at 30°C for 30 minutes. Reactions were terminated by proteinase K treatment, followed by phenol/chloroform extraction, ethanol precipitation, and primer extension as described above.

2.14 ANALYSIS OF SPLICING COMPLEX ASSEMBLY

Spliceosome complexes were assembled on the E5-10 wild type or mutant RNA or the Ad1 pre-mRNA in a 10 μl reaction containing 20 mM Hepes pH 7.4, 44% HeLa nuclear extract, 2.2 mM $MgCl_2$, 60 mM KCl, 1.5 mM ATP, 5 mM creatine phosphate, and 0, 1.6, or 3.2 μM CUGBP2 at 30°C for 15 minutes. Spliceosome assembly was stopped by the addition of heparin at a final concentration of 2 mg/ml and incubation at 30°C for an additional 3 minutes. Half of the reaction was separated on a 3.75% native polyacrylamide gel (80:1 acrylamide to bis ratio) cast in 50 mM tris-glycine buffer and run at 4 watts at 4°C for 4 hours. Gels were dried under vacuum and visualized with the BAS-2500 phosphoimaging system (Fujifilm) and ImageGauge software. For assembly of the ATP-independent E complex, the nuclear extracts were preincubated at 30°C for 10 minutes to deplete ATP and complexes were assembled in the absence of ATP or creatine phosphate for only 8 minutes.

Oligonucleotide-directed cleavage of U1 and U2 snRNAs was carried out as described previously (Hoffman and Grabowski, 1992). Spliceosome assembly reactions were prepared as above except 1.5 U RNase H (Promega) and 400 ng antisense oligo were added to the reaction and the final volume was 50 μl. Samples were preincubated at 30°C for 45 minutes prior to the

addition of RNA for spicing complex assembly. Mock reactions lacked RNase H and antisense oligos. I found that mock reactions lacking exogenous RNase H but containing an antisense oligo still carried out snRNA cleavage because RNase H was present in the nuclear extract. For U1 snRNA cleavage, the P007 antisense oligo (5´-GCCAGGTAAGTAT-3´) was used and for U2 snRNA cleavage, the P009 antisense oligo (5´-CAGATACTACACTTG-3´) was used. For verification of snRNA cleavage, half of the reaction was ethanol precipitated, resuspended in 10 μl water and 6.7 μl formamide loading buffer (97% deionized formamide, 0.1% xylene cyanol, 0.1% bromphenol blue) and separated on a 10% polyacrylamide 7 M urea gel run at 400 volts for 1.5 hours. Gels were soaked in ethidium bromide and visualized using the LAS-3000 Intelligent Dark Box (Fujifilm) and Image Gauge software.

2.15 CALCIUM PHOSPHATE TRANSFECTION AND ANALYSIS OF ENDOGENOUS TARGET EXONS

HEK293T cells were grown in DMEM (Cellgro), 10% (v/v) FBS. Twenty-four hours prior to transfection, 2 x 10^5 HEK293T cells were seeded in 35 mm dishes precoated with poly-L-lysine (Sigma). Cells were approximately 50% confluent at the time of transfection. Before transfection, all reagents were brought to room temperature. For one well, 1.6 μl of 1 μg/μl pcDNA4/NAPOR vector or pcDNA4 His/Max vector backbone were mixed with 16.1 μl 2.5 M $CaCl_2$ by vortexing briefly. Next, 65.8 μl water was added and the $CaCl_2$-DNA mixture was pipetted over 83.5 μl 2X BBS, pH 7.15 (50 mM N,N-bis(2-hydroxyethyl)-2-aminoethane sulfonic acid, 280 mM NaCl, 1.5 mM Na_2HPO_4) and vortexed for 3 seconds. The pH of the BBS solution should be titrated and tested every time it is prepared to optimize transfection efficiency. This solution can be aliquoted and stored at -20^{o}C for future use. Mixtures were incubated at room

temperature for 10 minutes before 164 μl was added drop-wise to each dish. After transfection, cells were incubated at 3% CO_2 for 36 hours prior to RNA isolation using TRIZOL reagent (Invitrogen). A white precipitate was visible under the microscope 24 hours after transfection. RT-PCR was as described for the *in vivo* splicing reporter assay (section 2.7) except the following modifications were used for PCR amplification: final primer concentration was increased to 2.5 μM, unlabeled primers were used, the annealing temperature was adjusted to 55^{o}C, and 26 cycles were carried out. A complete list of primers used for amplification of endogenous target exons is shown in Table 1. For amplification using the CUGBP2_E1 and E5/7 junction primers, 32 cycles were used. PCR samples were resolved on 2% agarose gels next to a 100 bp DNA ladder (Invitrogen) at 80 volts for 1-3 hours or until exon included and skipped bands were separated. Gel images were captured using the LAS-3000 Intelligent Dark Box (Fujifilm) and quantified with Image Gauge software. The profile mode was used to subtract background and to calculate percent exon inclusion of individual lanes.

Table 1. Primers used for amplification of endogenous exons.

Transcript	Exon #	Upstream Primer (5´-3´)	Downstream Primer (5´-3´)
GRIN1*	NI	AGATGATGCGCGTCTACAACTGGAAC	TCAGCAGAGCCGTCACATTCTTG
GRIN1*	CI	ATGCCCGTAGGAAGCAGATGC	CGTCGCGGCAGCACTGTGTC
GAPDH**		GACCCCTTCATTGACCTCAACTACATGG	TTGCCCACAGCCTTGGCAG
MAPT	2	CTTTGAACCAGGATGGCTGA	GGTGTCTCCAATGCCTGCTTC
MAP4	15	CTTCTGCTCCTGATCTGAAGAATG	ACCCACACTTGGAGGAGAC
SORBS1	5	GGGCCTCGTCTTCCTACAG	CACTTCTATAGCCCTTGGCAG
PPF1BP1	19	GCAGAAAGCAGGCCATTTGG	CAGCCAGATCTAGGTGCTCTG
SMARCE1	4	AAAGACCATCTTATGCCCCAC	CGGCTTATCTGGTGGCTTTG
FOX2	11	GTCTCCCTTTAGTTCCTGGCTTC	GGCTGTGCATATCTGTAGGCTG
NFAT5	2	GGACCTAGACCTGGAATCGC	ACCGCTTGTCTGACTCATTGATG
CTBP1	2	CAGCTCGCACTTGCTCAAC	CATCCAGCAATGCCACCAG
PTER	3	ACCTGTAGTCAACCACAGTCTG	GTCAAAGGCCAGGAGCAAC
MLLT10	13	ATGCAGTATCGGCATGATGGAG	GAGGTAAGTGAGAGCTGGAGAC
CUGBP2**	6	CTGATGGGCTGAGTCGAG	AGGTGGCAGTGTTGAGCTG
CUGBP1	6	GACCTGATGGCCTGAGC	TGCTGCATCTGCTGCTG
CUGBP2	1-5/7 JNX	GATTGGCGGAGCGCGAG	GTGAAGAGCAGCCCTCGACTC
NDRG3	3	CAGCTCACAGAGATCAAACCAC	CTCTTATGGTGACGTGGACC
SCAMP3	6	GTCCAGTTCAGCCCTGC	CTGGAGGACAAAGAGCACATC
EHMT1	24	CACCAGGCACCTTTGTCTGC	CTTTGATGTCCCAGAAGCGCTCT

Primers are complementary to exons that flank the target exon. Unless otherwise noted, primers are specific for human exon sequences. Asterisk, primers are specific for rat exon sequences; double asterisk, primers are universal.

3.0 FUNCTIONAL CUGBP2 BINDING SITES ARE INTIMATELY INVOLVED IN THE MECHANISM OF NI EXON SILENCING

3.1 INTRODUCTION

The CUGBP2 splicing factor is versatile in its ability to carryout dual functions with opposite effects on pre-mRNA splicing. The silencing face of CUGBP2's dual character was examined to determine how specific CUGBP2-nucleotide contacts alter spliceosome assembly to change the splicing pattern from exon inclusion to exon skipping. A general understanding of how this protein operates on the NI exon of the NMDA NR1 receptor will provide a framework to study splicing silencing of other CUGBP2 target exons and insight into the enhancing function of this splicing factor. This study serves as a template to study the functions of other splicing factors that either act solely as splicing silencers or as dual functional regulators.

3.2 CUGBP2 BINDS TO GU-RICH MOTIFS AT THE PERIMETERS OF THE PREDICTED NI BRANCH SITE REGION

A silencing role for CUGBP2 was shown for the NI exon of the GRIN1 transcript in a previous study but the mechanism of silencing was not characterized (Zhang et al., 2002). To gain insight into the mechanism, I sought to extend this analysis to identify the sequence and spatial arrangement of motifs associated with direct binding of CUGBP2 and silencing of the NI exon. A

nitrocellulose filter binding assay was used initially to locate the RNA region involved in stable binding by CUGBP2. Bound/total RNA was plotted as a function of increasing protein concentration and data were fit to a hyperbola to estimate the dissociation constants (K_d) for binding to individual RNA substrates. These RNA substrates included the NI exon, the NI exon and flanking introns, and the upstream intron, substrates E5-8, E5-10, and E5-15, respectively (Figure 6A). CUGBP2 was found to bind to the E5-10 and E5-15 substrates containing the upstream intron region with apparent dissociation constants in the nanomolar range (96 and 92 nM K_d values), but not to the E5-8 substrate containing the exon alone (Figure 6B). Because the downstream intron region is not a common feature of the high affinity binding substrates, this region, as well as the exon, must be dispensable for high affinity binding. Note that because binding curves are hyperbolic and not sigmoidal, it is likely that CUGBP2 binds as a monomer or preformed dimer and that binding is not cooperative at the protein concentrations tested.

To identify the specific nucleotides contacted by CUGBP2, I next carried out chemical modification footprinting with the E5-10 substrate, which contains the high affinity region identified by filter binding. CMCT modification at the N3 position of uracil and the N1 position of guanine causes termination of reverse transcription initiated at a downstream primer (Brunel and Romby, 2000). This chemical was chosen for footprinting because of the binding preference of human ETR-3 for UG-rich motifs as indicated by an iterative selection procedure (Faustino and Cooper, 2005). A representative footprint of the region upstream from the NI exon is shown (Figure 6C). These results reveal that a core UGUGU and upstream flanking GU motif are protected by CUGBP2 in a dose-dependent manner. At high CUGBP2 concentrations (>14.4 μM), an additional low affinity UGUG binding site was detected downstream from the core motif (data not shown). Regions within the NI exon and in the downstream intron were also subjected to RNA footprinting with CUGBP2, but no additional binding sites were detected in agreement with the filter binding experiments (data not shown). To verify the specificity of the assay, footprinting reactions were carried out with purified splicing factors U2AF and PTB. Protected

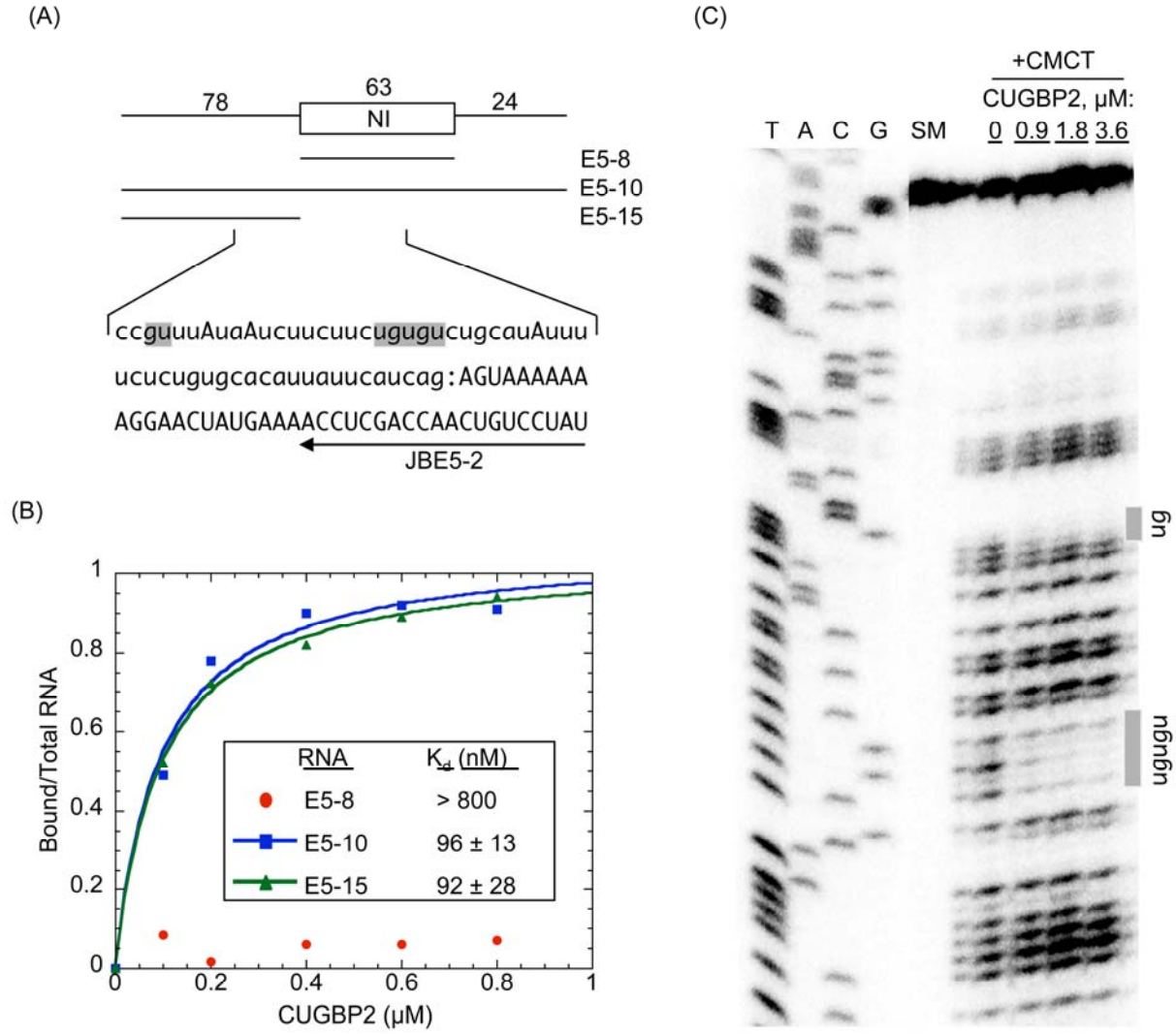


Figure 6. CUGBP2 contacts GU-rich motifs in the intron upstream from the NI exon. (A) Schematic of the NI exon (rectangle) and flanking introns (lines) is shown with corresponding nucleotide lengths (numbers, top). RNA substrates (E5-8, E5-10, and E5-15) used for filter binding assays are indicated below. Sequence analyzed by footprinting using primer JBE5-2 (arrow) includes the intron (lower case) and exon (uppercase) region of the 3′ splice site (colon). Shaded sequences represent regions protected from chemical modification by CUGBP2. Individual adenosines in uppercase represent predicted branchpoints. (B) Nitrocellulose filter binding analysis. ^{32}P-labeled RNA substrates from (A) were assembled with purified recombinant CUGBP2 protein and separated into protein bound and unbound fractions.

Representative graphs for each RNA substrate are shown. Inset: K_d values were calculated as the average of three experiments; ±, standard deviation. (C) RNA footprint analysis. E5-10 RNA was chemically modified with CMCT in the presence or absence of purified CUGBP2 protein. Modified positions were detected by primer extension (lanes +CMCT), in reference to a sequencing ladder generated from plasmid, E5-10 (lanes T,A,C,G). Shaded rectangles at right represent protected regions highlighted in the sequence shown in (A). Primer extension of starting material without modification is shown (lane SM).

regions were distinct from those observed for CUGBP2 (Figure 7, lanes 1-8) and consistent with the known RNA binding specificities of these factors (Singh et al., 1995).

I next asked if CUGBP2 was binding as a monomer or dimer. To address the issue of stoichiometry of binding by CUGBP2 to the core and upstream motifs, I carried out a competition footprinting assay with CUGBP2 and PTB. CUGBP2 and PTB bind to the E5-10 RNA substrate with similar affinities (96 and 63 nM, respectively) and the PTB binding sites overlap with the core but not upstream motif (Figure 7, lanes 1-4). Therefore, if CUGBP2 bound as a monomer to these two sites, PTB would be able to compete with CUGBP2 for binding to not only the core motif but also the upstream motif. This was indeed the case (Figure 7, lanes 9-12), suggesting that one protein simultaneously contacts both of these sites. Furthermore, CUGBP2 binds to the rNAPORbs RNA substrate containing the core but not upstream GU motif with a much higher dissociation constant (456 nM) than a similar RNA containing both motifs (92 nM), which indicates much lower affinity binding. Therefore, flanking interaction sites are necessary for high affinity binding by CUGBP2.

Taken together, these results support a model in which CUGBP2 contacts a pair of high affinity GU-rich motifs in the intron upstream from the NI exon. Notably, the core and flanking motifs protected by CUGBP2 are located at the boundaries of the predicted branch site region with the nucleotides protected by U2AF also within these boundaries. Thus, these results together with the high sequence conservation of the motifs (100%) across human, rat, mouse, fruit fly, and chicken genomes, support their involvement in the mechanism of silencing.

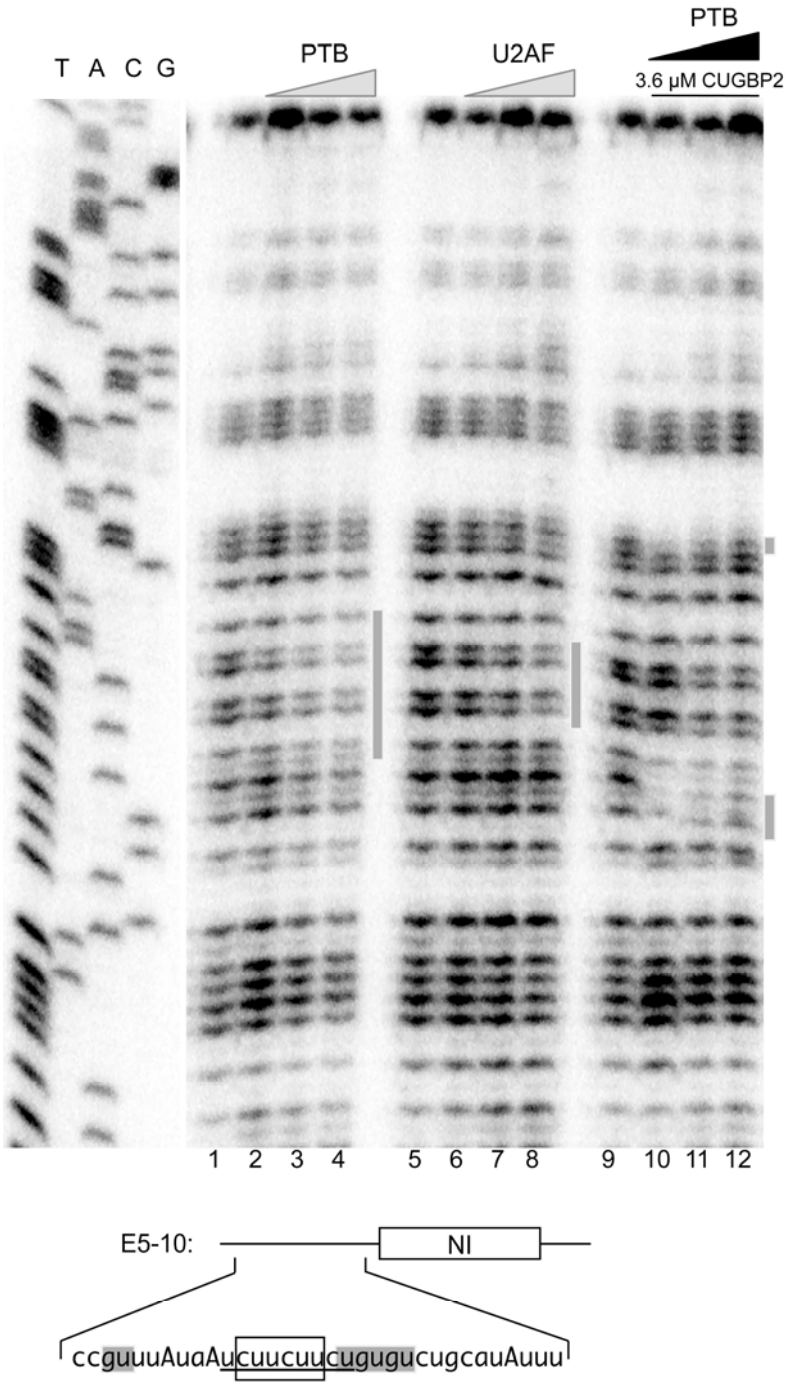


Figure 7. PTB and U2AF[65] bind to the polypyrimidine tract between GU-rich motifs and PTB can effectively compete with CUGBP2 for binding to the core and upstream motifs. RNA footprint analysis and sequencing ladder were carried out as in Figure 6 with the E5-10 substrate and JBE5-2 primer. Gel, lanes 1-8: CMCT modification footprint with recombinant PTB

and purified HeLa U2AF. Left lane of each set, no protein; gray wedge, 1.8, 3.6, and 7.2 μM protein added. Shaded rectangles represent protected regions. Lanes 9-12: Competition footprint with recombinant PTB and CUGBP2. Left lane, no protein; right lanes, 3.6 μM CUGBP2; black wedge, 0, 3.6, and 7.2 μM PTB added. Shaded rectangles at right represent regions where CUGBP2 protection was lost by competitive binding by PTB. PTB footprint (UCUUCUUCU) is underlined, U2AF footprint (CUUCUU) is boxed, and CUGBP2 footprint (GU, UGUGU) is shaded in schematic.

3.3 CUGBP2 MUTANTS PROVIDE INSIGHT INTO THE TOPOLOGY OF RNA BINDING BY INDIVIDUAL RRMS

To learn more about the topology of binding by individual RRMs of CUGBP2, I carried out additional filter binding and footprinting experiments with purified recombinant CUGBP2 mutant proteins. These mutants included RRM1_2, RRM2_3, and ΔDD. RRM1_2 contained RRMs 1 and 2, RRM2_3 contained RRMs 2 and 3, and ΔDD contained RRMs 1, 2, and 3. None of these proteins contained the divergent domain. Note that there were amino acid and charge differences in the linker region between RRMs 2 and 3 in the RRM2_3 and ΔDD mutants (RRM2_3, LE; ΔDD, SR). See Figure 4 for the predicted structure of the wild type protein and relative positions of RRM domains.

To determine the role of the divergent domain in RNA binding, I carried out RNA footprinting with the ΔDD mutant. This mutant protected the same nucleotides as the wild type protein, indicating that the divergent domain is dispensable for intimate protein-nucleotide interactions (Figure 8, lanes ΔDD). Perhaps one role of the divergent domain is to position RRMs relative to each other to provide flexibility in binding to flanking interaction sites separated by various distances. In support of this, ΔDD bound to the shorter rNAPORbs RNA (containing only the core motif) better than the wild type protein but not as well to the longer E5-10 and E5-15 RNA substrates (containing both motifs) (Table 2). RRM2_3 protected the same nucleotides as the wild type protein (Figure 8, lanes RRM2_3) and bound to the rNAPORbs RNA substrate with highly similar affinity (Table 2). Therefore, it seems that RRM1 and the divergent domain are dispensable for high affinity protein-nucleotide interactions.

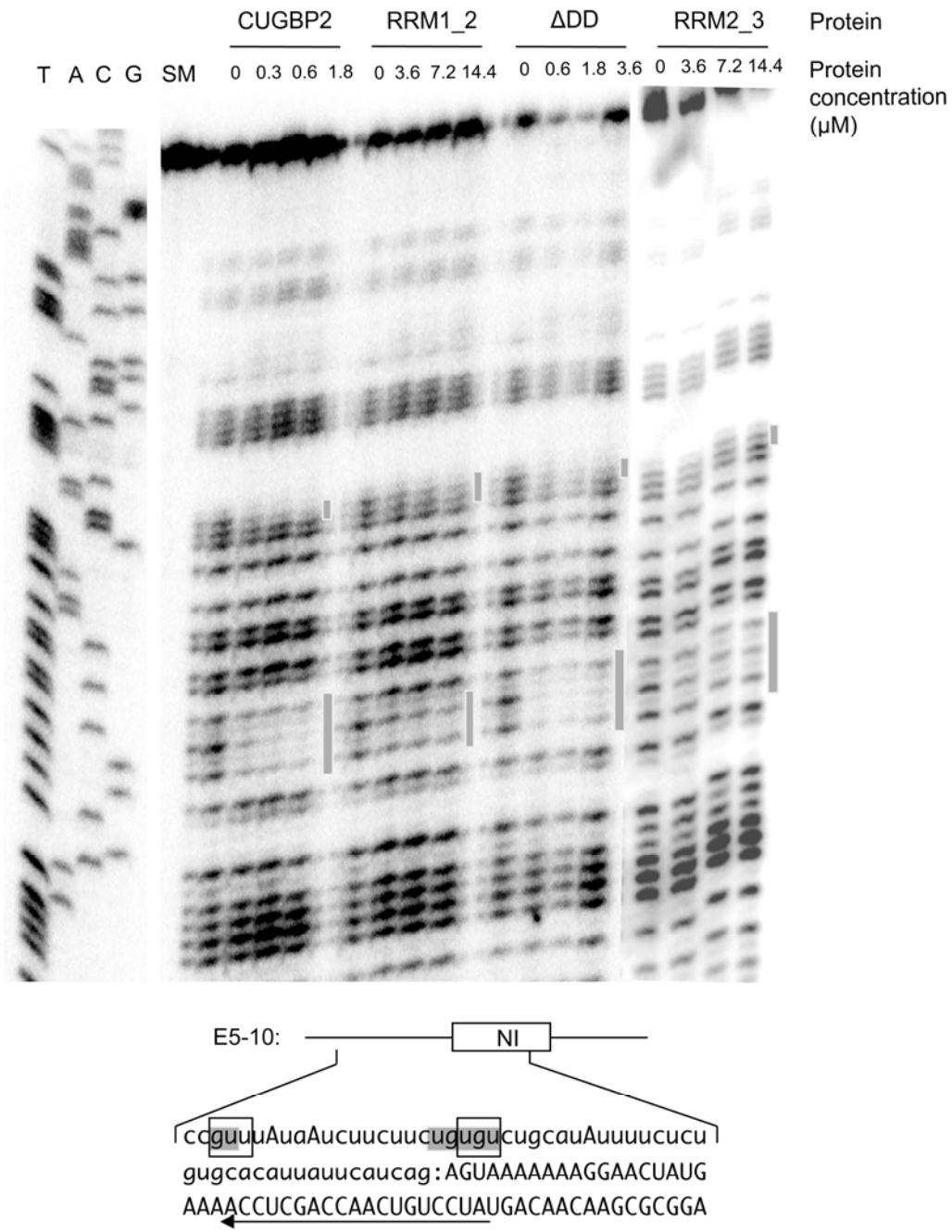


Figure 8. CUGBP2 RRMs 2 and 3 are sufficient for protection of the core and upstream GU-rich motifs. CMCT modification footprint analysis was carried out as in Figure 6 with protein concentrations ranging from 0 to 14.4 μM as indicated above the gel. Shaded rectangles represent protected regions. Primer extension of starting material without modification is shown (SM). Protected nucleotides are indicated on the sequence below the gel. Regions protected by

CUGBP2, RRM2_3, and ΔDD are shaded and regions protected by RRM1_2 are boxed. Arrow indicates the location of the JBE5-2 primer used for primer extension of modified RNA. Sequencing ladder was generated from plasmid, E5-10 (lanes T,A,C,G), using the JBE5-2 primer.

RRM1_2 bound to the E5-10 and E5-15 RNA substrates with approximately five times less affinity than the wild type protein (Table 2). In the footprinting assay, this mutant only contacted a UGU from the core and GU from the upstream motif, but not to a complete set of motifs (Figure 8, lanes RRM1_2). Therefore, RRMs 1 and 2 are insufficient for high affinity binding and proper nucleotide protection. RRM2 may be involved in stable binding to the shorter upstream motif and may contact part of the core motif when binding in an alternate register.

Table 2. Summary of nitrocellulose filter binding assays.

	CUGBP2	RRM1_2	ΔDD	RRM2_3	PTB
E5-10	96	473	179	NT	63
E5-15	92	546	153	NT	100
rNAPORbs	456	NT	352	435	NT

Wild type and mutant CUGBP2 and PTB bind to E5-10, E5-15, and rNAPORbs RNA substrates with different affinities. Note that these proteins do not bind to the E5-8 RNA substrate. Mutant proteins contained RRMs 1 and 2 (RRM1_2), RRMs 1, 2, and 3 (ΔDD), or RRMs 2 and 3 (RRM2_3) but lacked the divergent domain (DD). The E5-10 and E5-15 RNA substrates are illustrated in Figure 6. rNAPORbs, UCUUCUGUGUCUGCA, contains the core motif of the NI intron plus five flanking nucleotides on either side. All units are nM. NT, not tested.

Taken together, these data support a model in which RRM2 contacts the upstream motif and RRM3 contacts the core motif with the divergent domain providing flexibility in the positioning of RRMs 2 and 3 relative to each other (Figure 9). The divergent domain and RRM1 are dispensable for high affinity binding and may be involved in protein-protein interactions or nuclear localization (Ladd and Cooper, 2004; Goo and Cooper, 2009). Additionally, CUGBP2 may bind in an alternate register of less affinity by contacting a GU at the core motif and UGUG at the downstream motif. Still, these results do not rule out binding by CUGBP2 as a monomer to all three motifs nor can I confirm that individual RRMs function properly when taken out of the wild type context. Structural studies are necessary to fully understand the topology of binding.

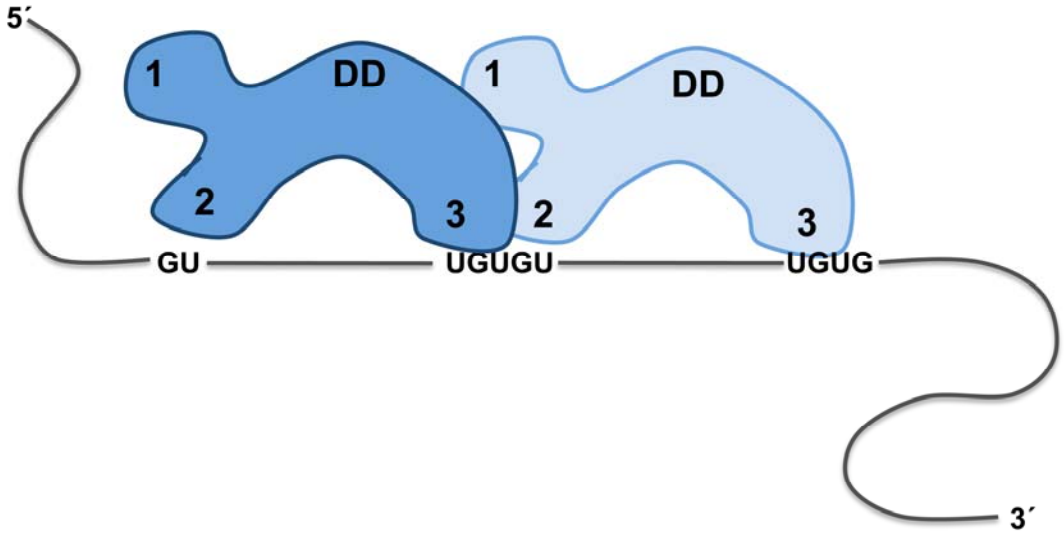


Figure 9. Model of the proposed topology of RNA binding by CUGBP2. Relative positions of GU-rich motifs are shown on the intronic RNA sequence (curved line) in the 5´ to 3´ direction. CUGBP2 (dark blue) binds with highest affinity to the upstream GU and core UGUGU motif with RRM2 (2) contacting the upstream GU and RRM3 (3) contacting the core UGUGU motif. CUGBP2 (light blue) also binds with less affinity to a GU at the core motif and a UGUG at the downstream motif. The divergent domain (DD) and RRM1 (1) are not necessary for protection in the footprinting assays and may mediate CUGBP2 protein-protein interactions or nuclear localization. The divergent domain functions to correctly position RRMs 2 and 3 relative to each other.

3.4 CUGBP2 BINDING SITES ARE NECESSARY, SUFFICIENT, AND SPECIFIC FOR EXON SILENCING

In order to determine whether nucleotides in contact with CUGBP2 upstream from the NI exon are necessary for its silencing role, I generated an *in vivo* splicing reporter with the NI exon and its immediate adjacent flanking introns inserted between β globin exons 1 and 2 (DUPNIwt) (Figure 10A). Mutant derivatives of this reporter contained site-directed mutations in the GU dinucleotide and UGUGU core motifs (m1 and m2 motifs, respectively) as identified by footprinting. A nearby UGUG motif (m3) was also mutated, since it showed weak protection in the footprinting assays.

Splicing reporters were co-expressed in the presence and absence of CUGBP2 in C2C12 mouse myoblast cells or N18TG2 mouse neuroblastoma cells, which have little or no endogenous protein as shown by Western blotting (Figure 11). The change in exon inclusion value, ΔEI, was then calculated as the difference between the % exon inclusion ± CUGBP2. The ΔEI value is used here as a convenient measure of the effectiveness of CUGBP2 to induce exon skipping (or silencing). While CUGBP2 expression in C2C12 cells induced exon skipping of the wild type substrate with a ΔEI value of –33% (Figure 10B, lanes wt) mutations in all three motifs eliminated silencing entirely as indicated by a ΔEI value of 0% (lanes m1,2,3). The double mutations also showed a significant reduction in the silencing effect of CUGBP2 (lanes m1,2, m2,3, m1,3). Of this group, combined mutations at positions m1 and m2 showed the smallest degree of silencing (ΔEI -5%), suggesting that these sites are intimately involved in the mechanism of action. Single mutations also showed a reduction in silencing indicative of additive effects (lanes m1, m2, m3). Similar results were observed in N18TG2 cells with

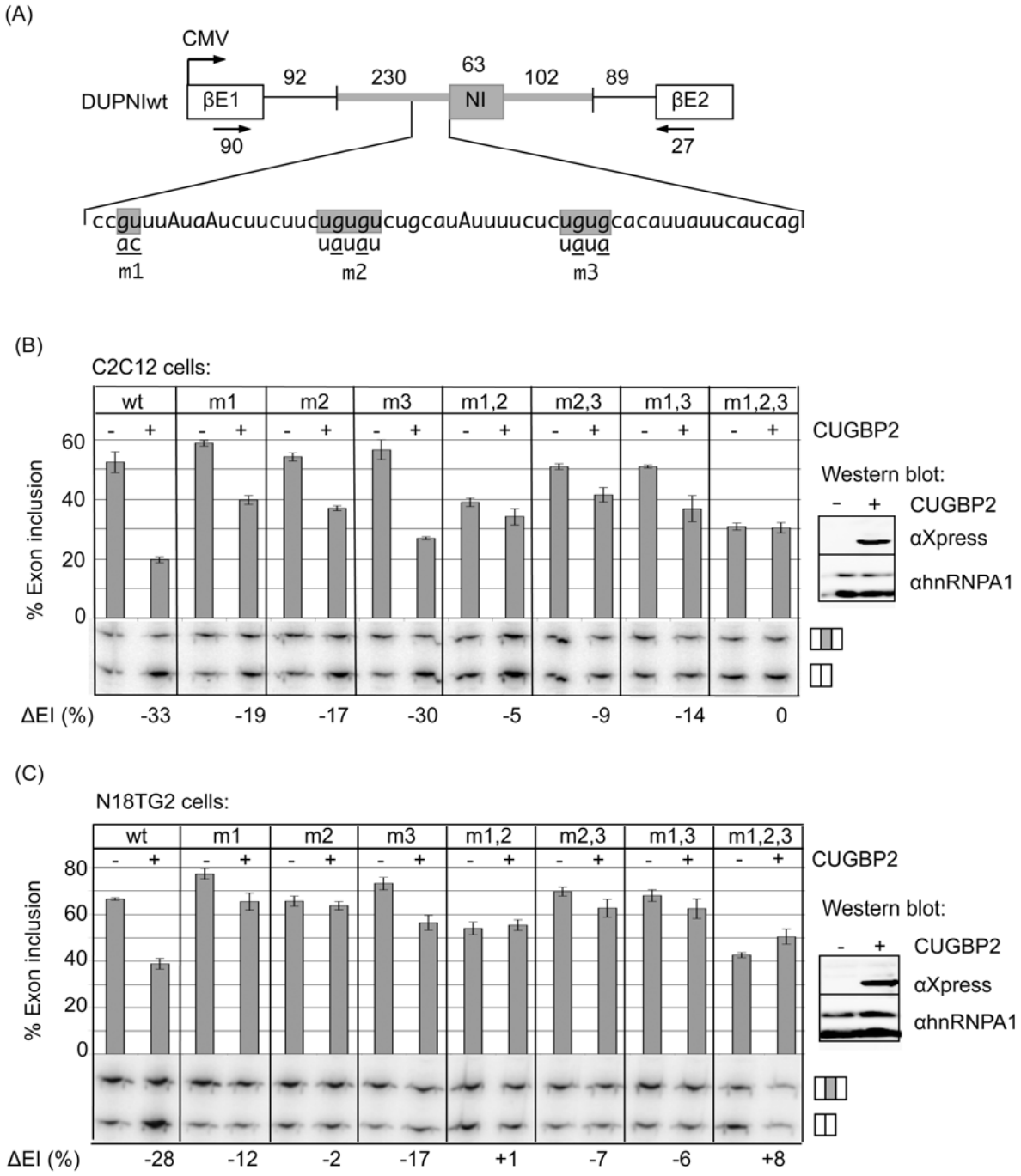


Figure 10. The UGUGU core motif and flanking GU dinucleotides are functionally important for CUGBP2 silencing. (A) Schematic of wild type (wt) and mutant versions of the DUPNI splicing reporter. The NI cassette exon (shaded rectangle) and regions of the native flanking introns (shaded lines) were inserted between β-globin exons βE1 and βE2 downstream

of a CMV promoter. Nucleotide lengths in base pairs are indicated above and below schematic. Sequence shows an expanded view of the NI 3′ splice site region with mutations m1, m2, and m3 indicated by underscores below the shaded regions. Arrows indicate the location of primers used for RT-PCR amplification of exon included and exon skipped mRNAs. (B) Splicing reporter analysis in C2C12 cells. Splicing reporters with no mutation (wt), with single (m1, m2, m3), or combined (m1,2, m2,3, m1,3, m1,2,3) mutations were expressed with vector backbone control (-) or with pcDNA4/CUGBP2 expression vector (+). The gel panel is a representative polyacrylamide gel image with the top band corresponding to the exon included and the bottom band corresponding the exon skipped mRNA. The bar graph shows the percent exon inclusion as an average of three separate experiments. The change in percent exon inclusion (ΔEI) as a function of CUGBP2 expression is shown below the gel panel. Inset: Western blot analysis was used to verify Xpress-tagged CUGBP2 expression; endogenous hnRNPA1 was a loading control. (C) Experiments were as in (B) except N18TG2 cells were used.

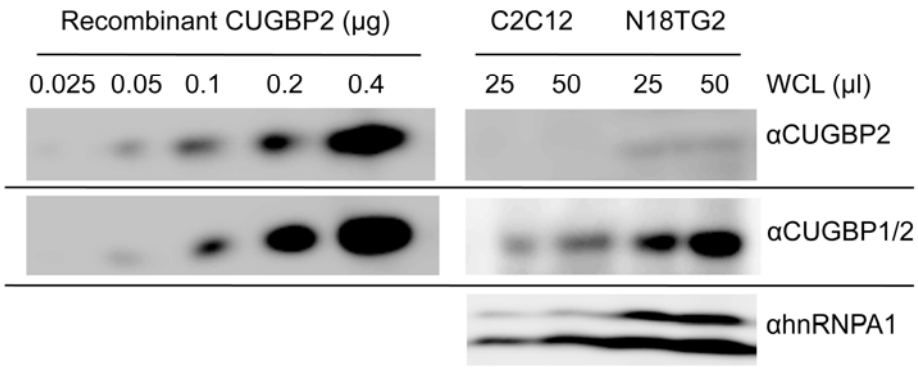


Figure 11. Quantitative Western blot analysis of whole cell lysates demonstrates that endogenous CUGBP2 levels are low in C2C12 and N18TG2 cells. Whole cell lysates (WCL) were obtained from 90% confluent 10 cm dishes of each cell type as indicated (top right). Twenty-five or 50 µl of cell lysate was loaded and Western blotting was carried out with αCUGBP2 (1H2), αCUGBP1/2 (3B1), or αhnRNPA1 (9H10) antibodies as indicated at right. Left panel is Western blot analysis of recombinant CUGBP2 as a control for antibody sensitivity and relative protein levels. αhnRNPA1, loading control.

consistent loss of silencing by CUGBP2 with each mutation (Figure 10C). New to this cell type is higher basal levels of inclusion in the absence of CUGBP2 and reduced silencing by CUGBP2 on all substrates tested. This could be the consequence of the differential expression of splicing factors in these two cell lines. That is, an enhancer may act on the NI exon in N18TG2 cells and may be better able to compete with CUGBP2 when its binding sites are compromised. A good candidate enhancer is Fox because a perfect match to its enhancer element, (U)GCAUG (Auweter et al., 2006), is located near the NI 5´ splice site in the downstream intron.

I also demonstrate that the m1, m2, and m3 motifs are involved in tissue-specific silencing of the NI exon in the context of the DUPNI splicing reporter in transfected primary cortical cultures (Figure 12A). It is important to note that I did not overexpress CUGBP2 in the primary cultures, therefore reporter splicing patterns are a result of endogenous splicing factors acting on pre-mRNA from transfected plasmids. The NI exon is only included 50% of the time when all three motifs are intact (lanes DUPNI) as compared to 95% of the time when the m1 and m2 motifs are mutated (lanes m1,2 and m1,2,3). Furthermore, individual and combinations of mutations confirm additive effects in primary cultures (lanes m1, m2, m3, m2,3, and m1,3). I verify nuclear expression of CUGBP2 in primary cortical cultures by Western blotting (Figure 12B) and show that CUGBP2 is expressed in both neurons and glial cells by immunofluorescence (Figure 12C). DAPI nucleic acid stain was used to verify nuclear localization in both neurons and glial cells (data not shown). Taken together, CUGBP2 is expressed in the nucleus of cortical neurons where it likely regulates NI exon splicing through the m1, m2, and m3 motifs.

To further investigate the roles of the m1, m2, and m3 motifs for exon silencing by CUGBP2, I introduced a 39 nucleotide region containing the three motifs upstream of a constitutive exon in a different context (Figure 13A). Constitutive exon 3 of the DIP13β transcript was tested, since its splicing pattern is insensitive to CUGBP2 regulation, and unlike the NI

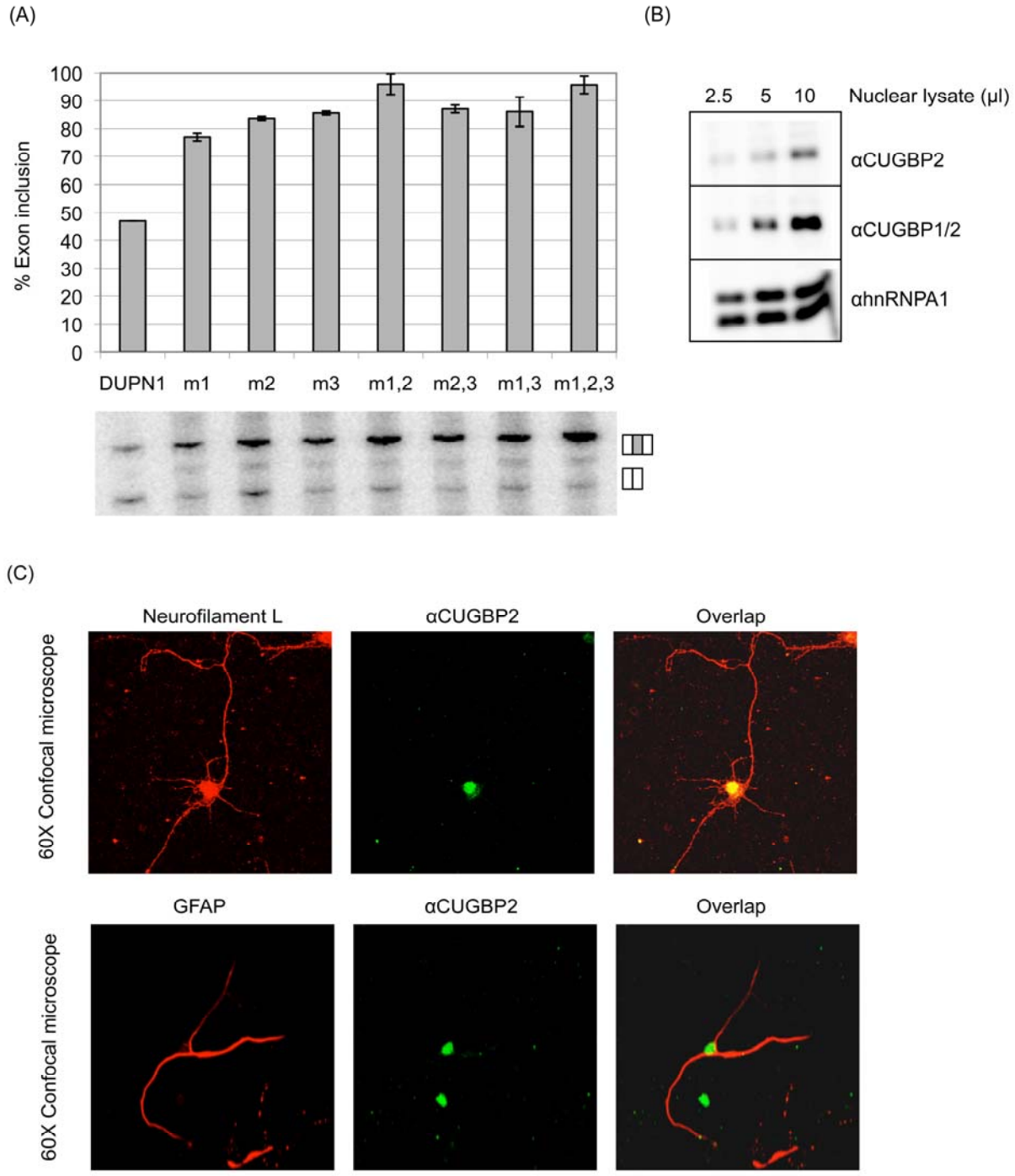


Figure 12. NI exon skipping in primary cortical cultures depends on the m1, m2, and m3 motifs and correlates with CUGBP2 expression. (A) *In vivo* splicing reporter analysis. DUPNIwt (DUPNI) and mutant (m1, m2, m3, m1,2, m2,3, m1,3, m1,2,3) splicing reporters were transfected into rat primary cortical neurons for 24 hours prior to RNA isolation and RT-PCR

analysis. The gel panel is a representative polyacrylamide gel image with the top band corresponding to the exon included and the bottom band corresponding the exon skipped mRNA. The bar graph shows the percent exon inclusion as an average of two separate experiments. (B) Western blot analysis of nuclear lysates with αCUGBP2 (1H2) and αCUGBP1/2 (3B1) antibodies shows endogenous CUGBP2 expression in rat primary cortical cultures. Two and a half, 5, or 10 μl nuclear lysate was loaded as indicated at top. αhnRNPA1 was used as a loading control. (C) Immunofluorescence analysis demonstrates that CUGBP2 is enriched in the nucleus and soma of rat primary cortical neurons and glial cells. Mixed cultures of neurons and glial cells were grown on coverslips and assayed for cell type-specific CUGBP2 expression (αCUGBP2) by costaining with a neural (Neurofilament L) or glial (GFAP) cell-specific antibody. An overlap of confocal images is shown at right.

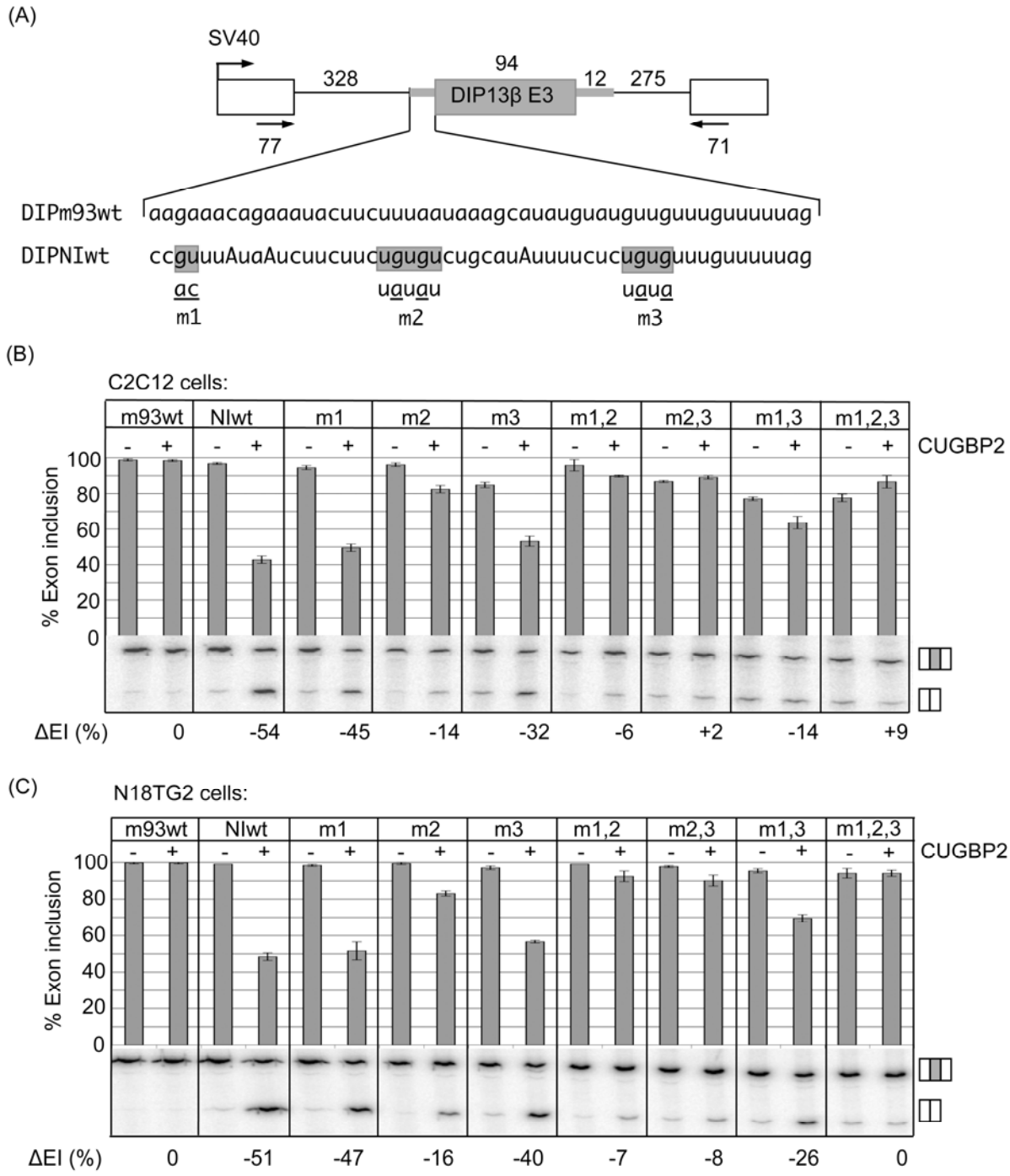


Figure 13. CUGBP2 silencing motifs are functionally transferable. (A) A 39 nucleotide region containing the m1, m2, and m3 silencing motifs from the NI intron was inserted upstream from DIP13β exon 3 (middle exon) in the SIRT1 splicing reporter context (DIPNIwt). Expression was driven by the SV40 promoter. Nucleotide lengths in base pairs are indicated. Arrows

indicate primers used for RT-PCR amplification; numbers below give nucleotide lengths contributing to PCR products. Individual or combinations of mutations were introduced at CUGBP2 regulatory sites as in Figure 10. Sequence of control region (DIPm93wt) corresponds to 3′ splice site region of constitutive exon, DIP13β exon 3. (B) Splicing reporter plasmids (m93wt, NIwt, m1, m2, m3, m1,2, m2,3, m1,3, m1,2,3) were cotransfected in C2C12 cells with vector backbone control (-) or with pcDNA4/CUGBP2 expression vector (+). The graph, gel image, and ΔEI calculation are as described for Figure 10. (C) Experiments were as in (B) except N18TG2 cells were used.

exon is not under alternative splicing control. The introduction of the 39 nucleotide region conferred strong silencing by CUGBP2 (Figure 13B, lanes NIwt; ΔEI, -54%), in contrast to the parent plasmid, which was unregulated by CUGBP2 (lanes m93wt; ΔEI, ~0%). Single and combined mutations in the m1, m2, and m3 motifs were also tested in this context (lanes m1, m2, m3, m1,2, m2,3, m1,3, and m1,2,3). Exon silencing by CUGBP2 was nearly eliminated when site-directed mutations were introduced into both the m1 and m2 motifs (lanes m1,2; ΔEI, -6%). Mutations in both the m2 and m3 motifs also led to a significant reduction in silencing (lanes m2,3). Thus, the general requirement for a pair of proximal CUGBP2 motifs, and the additive effects of the single mutations were verified in this context. This heterologous reporter was also tested in N18TG2 cells (Figure 13C). In this context, results were consistent between cell lines. Furthermore, compared to the NI exon, DIP13β E3 has strong 5′ and 3′ splice sites and individual mutations have less of an effect on the basal level of inclusion. Therefore, this reporter is a good system to study isolated effects of CUGBP2 on the m1, m2, and m3 motifs.

I also tested the silencing role of a closely related family member, CUGBP1, on wild type and mutant substrates since it is expressed in both cell lines tested (Figure 11). CUGBP1 silences the NI exon (Figure 14A, lanes DUPNI wt) with a dependence on the same motifs (lanes DUPNI m1,2,3). However, CUGBP1 silencing is much weaker in the DIPNI context (lanes DIPNI wt, ΔEI -16 compared to ΔEI -54 for CUGBP2) indicating that additional sites outside of the transferred region are necessary for strong silencing by CUGBP1. In support of this, exon 3 of the DIPNIwt reporter is included >99% of the time in N18TG2 cells (Figure 13C, lanes NIwt) despite high levels of endogenous CUGBP1 (Figure 11). Therefore, CUGBP1 shows a weaker silencing role compared to CUGBP2 and splicing effects from endogenous CUGBP1 should not interfere with CUGBP2 overexpression studies. As additional controls, I carried out similar experiments with overexpression of PTB or Nova, since both of these factors are known to silence the NI exon through distinct motifs (UCUU and YCAY, respectively; Y, pyrimidine)

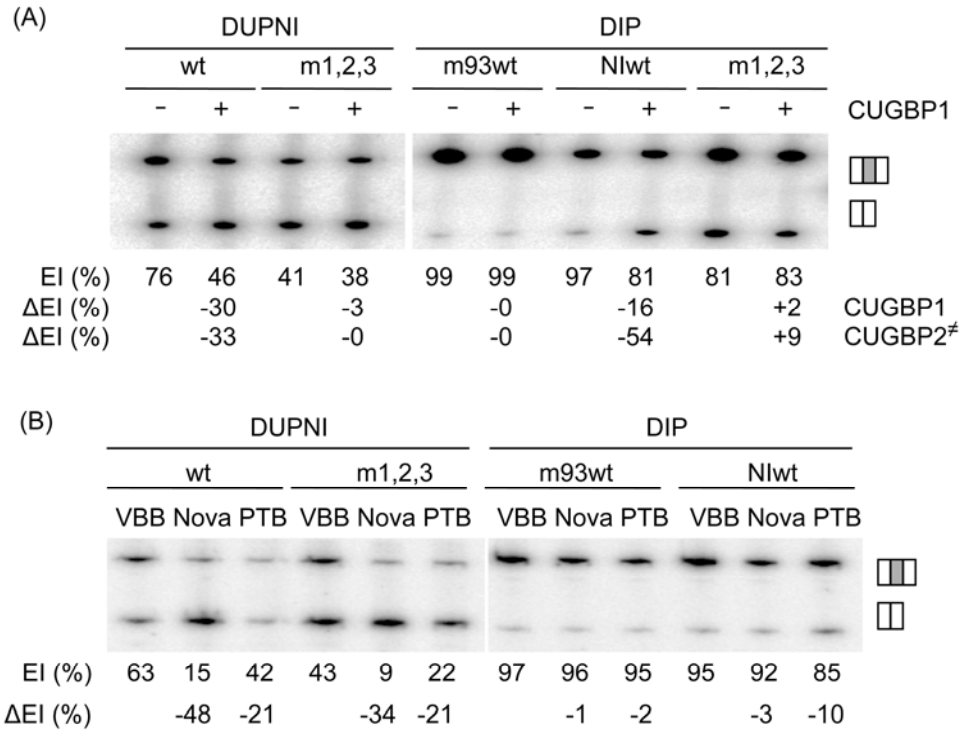


Figure 14. CUGBP2 regulatory motifs are specific. (A) Splicing reporters (top) were transfected into C2C12 cells in the presence (+) or absence (-) of CUGBP1 as indicated and included and skipped forms of spliced reporter RNA were assayed by RT-PCR and separated on a polyacrylamide gel. Gel image is representative of three separate experiments. Percent exon inclusion (EI (%)) and the change in percent exon inclusion (ΔEI (%)) with CUGBP1 overexpression is indicated below the gel panel. ≠The effect of CUGBP2 overexpression is shown for comparison. (B) Splicing reporter assays were carried out as in (A) except Nova and PTB protein expression vectors were used. VBB, vector backbone control.

(Zhang et al., 1999). As expected, PTB and Nova were active in silencing the NI exon in the context of the DUPNIwt splicing reporter, and these effects were maintained in the presence of the m1,2,3 triple mutation (Figure 14B). Furthermore, the region transferred to the DIPNIwt splicing reporter was insufficient for strong splicing silencing by Nova or PTB. Thus, the m1, m2, and m3 motifs are specific for silencing by CUGBP2.

3.5 CUGBP2 BLOCKS BRANCHPOINT FORMATION BETWEEN GU-RICH MOTIFS

I hypothesized that CUGBP2 may function to silence the NI exon by blocking branchpoint formation in the upstream intron. Based on complementarity to U2 snRNA, two candidate branch sites, A1 and A2, are located between the m1 and m2 motifs, and a third, weaker candidate, A3, resides just downstream between the m2 and m3 motifs (Figure 15A). As a test of this hypothesis, I measured branchpoint formation for the DUPNI substrate under *in vitro* splicing conditions in the presence and absence of recombinant CUGBP2. Note that endogenous CUGBP2 levels are not detectable in HeLa nuclear extracts by immunoblotting with the 1H2 antibody, which is highly specific for CUGBP2 (Ladd et al., 2005). According to the perimeter-binding model, the addition of recombinant CUGBP2 to the extract should bind and preferentially occupy motifs m1 and m2 on the wild type substrate with the resulting inhibition of one or more of the branch sites in this neighborhood. The protein can also bind in an alternate register of lower affinity by contacting a GU at the m2 site and UGUG at the m3 site. Alternately, one protein may contact all three sites simultaneously.

Branchpoints were detected by primer extension as for the experiments in Figure 6C. The results for the wild type substrate verified the use of the predicted branchpoints with a preference for A1 and A2, compared to A3 (Figure 15B, lanes 2,3). The A1 and A2 branchpoints satisfied the criteria for ATP dependence (lane 1). Primer extension of reactions following

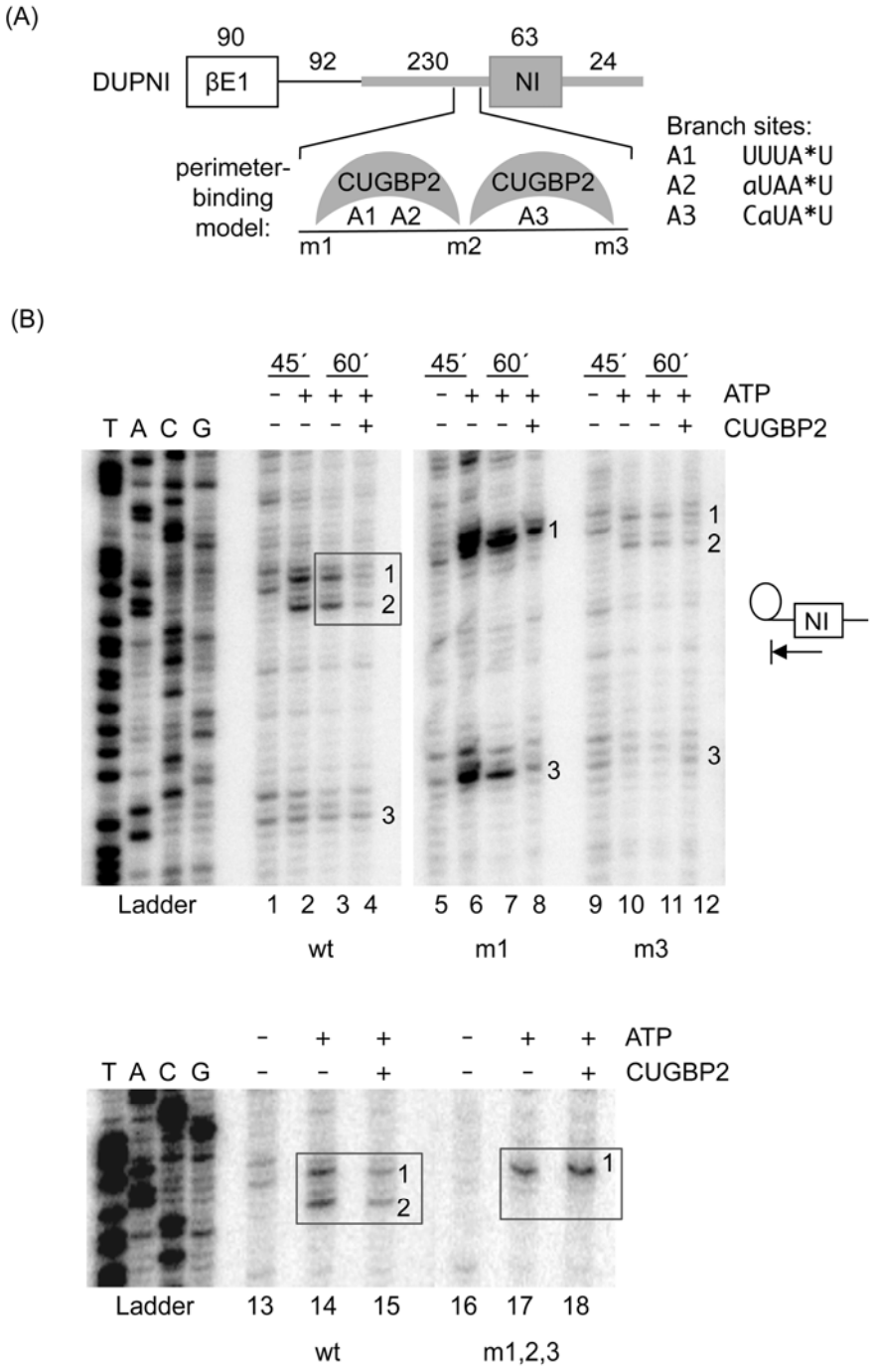


Figure 15. CUGBP2 blocks branchpoint formation between RNA-protein contact sites. (A) Schematic of the DUPNI pre-mRNA used for *in vitro* splicing assays. The NI exon and adjacent intron regions are shaded; numbers above schematic indicate nucleotide sizes. Expanded region shows perimeter-binding model with relative positions of CUGBP2 motifs (m1, m2, m3)

and branchpoint adenosines (A1, A2, A3) mapped in these experiments. Branch site sequences are shown at right; branchpoint adenosine, asterisk; nucleotides matching the consensus, YUNAY (Y, pyrimidine; N, any nucleotide), uppercase; mismatches, lowercase. (B) Gel panels show primer extension analysis of splicing intermediates of the wild type (wt), m1, m3, and m1,2,3 DUPNI RNA substrates (bottom) with primer, JBE5-2. Schematic at right illustrates the termination of reverse transcriptase at branchpoint positions in the assay. Branchpoint numbers on gel correspond to positions indicated on sequence in (A). A sequencing ladder is shown for the pBSDUPNI plasmid (ladder). For lanes 1-12, splicing reactions containing ATP (+ATP lanes) were incubated for 45 and 60 min with (+) or without (-) 1.6 μM recombinant CUGBP2. Control reactions lacked ATP. For lanes 13-18, 60 min splicing reactions were used. Boxes emphasize the inhibition of branchpoints A1 and A2 on the wild type substrate but not on the m1,2,3 mutant RNA substrate.

debranching showed a loss of stops at A1, A2, and A3 providing confirmation that all three of these adenosines are used as branchpoints (Figure 16). Notably, branchpoint formation was inhibited when the *in vitro* splicing reactions were supplemented with recombinant CUGBP2 (Figure 15B, lane 4). An analysis of the corresponding *in vitro* splicing reactions confirmed that CUGBP2 inhibited the formation of splicing intermediates of these reactions (Figure 17, lanes wt). Note that I also carried out the *in vitro* splicing reactions in the presence of a PTB competitor because PTB levels are artificially high in HeLa nuclear extract and PTB has been shown to inhibit splicing of the NI exon (Zhang et al., 1999; Zhang et al., 2002). However, I did not observe a significant difference in splicing or CUGBP2 regulation in the presence of the PTB competitor under all conditions tested (data not shown).

I next examined the effect of the single mutation in motif m1 as a test of whether branchpoint inhibition occurs between core and flanking motifs. That is, a single mutation in m1 should permit the binding of CUGBP2 to the remaining intact sites (m2 and m3) leading to preferential inhibition of branchpoint A3. Indeed, under conditions in which the m1 site was mutated, branchpoint A3 was preferentially inhibited as expected for a model involving flanking interaction motifs (Figure 15B, lanes 7,8). Consistent with this observation, the corresponding *in vitro* splicing gel showed that CUGBP2 inhibited the formation of one lariat intermediate but not the other (Figure 17, lanes m1). For the single mutant, m3, which should display the reciprocal pattern of inhibition by CUGBP2, branchpoints A1 and A2 were preferentially inhibited relative to A3 (Figure 15B, lanes 11,12). Finally, the triple mutant, m1,2,3, was tested. Here, the elimination of all three binding motifs neutralized the inhibitory effects of CUGBP2 on branchpoint formation (lanes 17,18). Again, these results were consistent with the splicing intermediates of these reactions (Figure 17, lanes m3 and m1,2,3). Adenovirus major late (Ad1) pre-mRNA was tested as a control, because Ad1 pre-mRNA lacks CUGBP2 motifs in the upstream intron. Both the *in vitro* splicing and branchpoint formation of Ad1 were unaffected by the addition of recombinant CUGBP2 (Figure 18A/B). Under *in vitro* splicing conditions tested, it

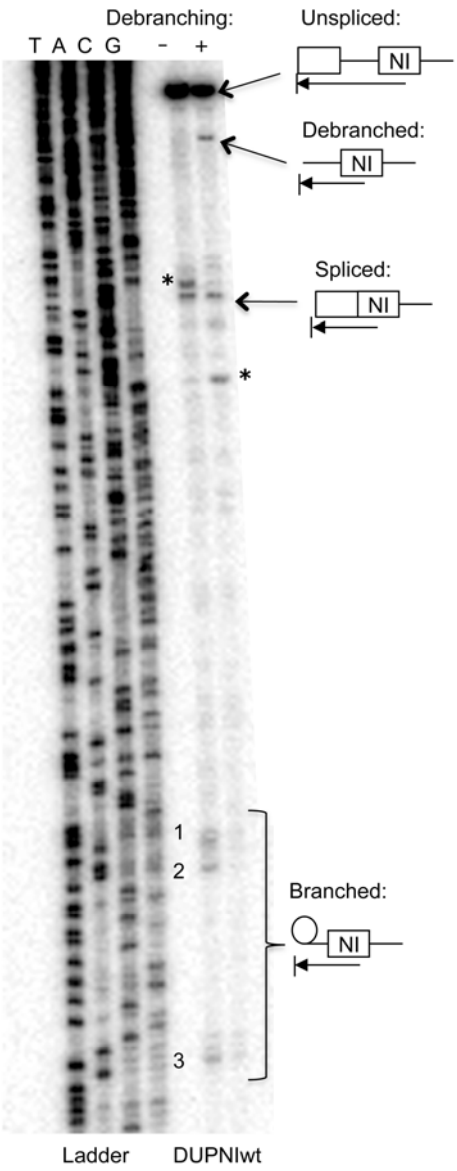


Figure 16. Three branchpoints in the intron upstream from the NI exon were confirmed with a debranching control. Branchpoint analysis of the DUPNI pre-mRNA was carried out as in Figure 15 with (+) or without (-) debranching prior to primer extension. Branchpoint numbers correspond to numbers in Figure 15. Schematics at right illustrate the termination of reverse transcription on various RNA intermediates and products. Asterisks mark unidentified bands that may represent different transcription start sites or the use of cryptic splice sites. Sequencing ladder generated from the pBSDUPNI plasmid is shown, Ladder.

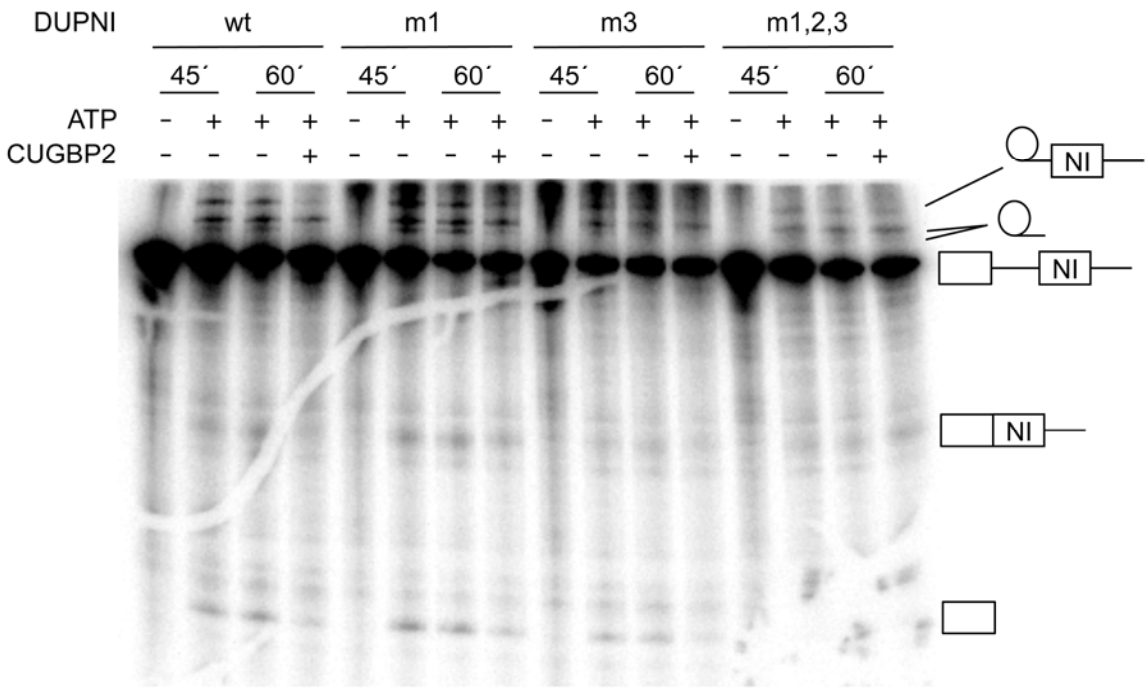


Figure 17. Recombinant CUGBP2 inhibits splicing of the intron upstream of the NI exon *in vitro*. *In vitro* splicing reactions were carried out using the two exon pre-mRNAs, DUPNI wild type (wt) and mutant derivatives m1, m3, and m1,2,3 (top). The presence (+) or absence (-) of ATP or 1.6 μM recombinant CUGBP2 and the time of incubation are indicated at the top of the gel. The structures of RNA intermediates and products are indicated at right. Note that the time dependence of the accumulation of branchpoints mapped in Figure 15 coincides with the appearance of the intron lariat-3´ exon intermediate. The doublet band in the vicinity of the intron lariat is consistent with branchpoints at varying distances from the 3´ splice site (see for example, lane 6 from left).

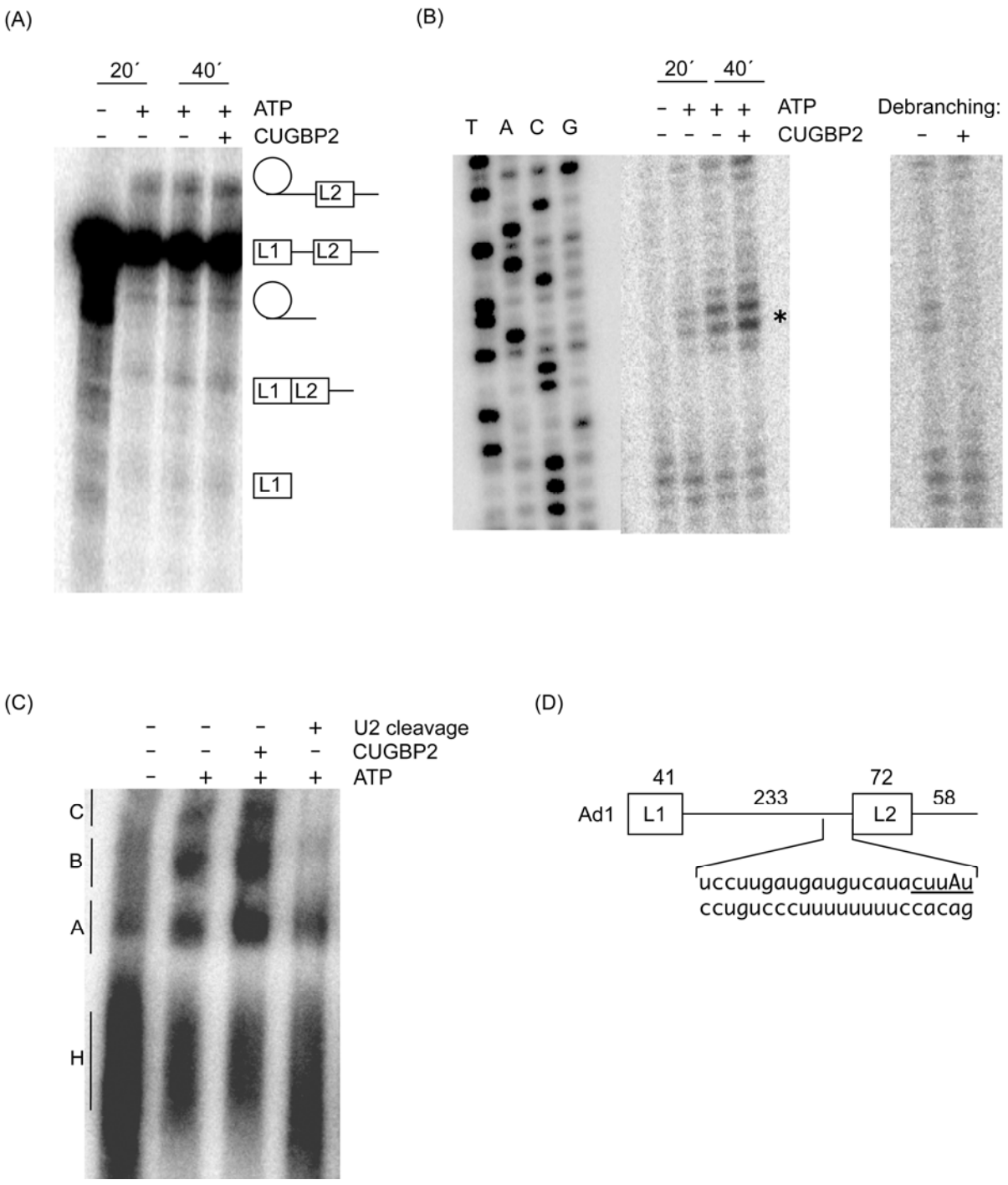


Figure 18. Ad1 pre-mRNA is resistant to splicing silencing by CUGBP2. (A) *In vitro* splicing reactions were carried out with the Ad1 two exon pre-mRNA in the presence (+) or absence (-) of ATP or 1.6 μM recombinant CUGBP2 protein for 20 or 40 minutes as indicated at the top of the gel. RNA intermediates are shown at right. (B) Branchpoint analysis of Ad1 pre-mRNA was

carried out in the presence (+) or absence (-) of ATP or 1.6 μM recombinant CUGBP2. Splicing reactions were carried out for 20 or 40 minutes as indicated above the gel. Asterisk, position of branchpoint mapped in these experiments. Right, debranching reaction of the branchpoint detected after 20 minutes in the presence of ATP without CUGBP2. (C) Analysis of splicing complex assembly on the Ad1 substrate in the presence (+) or absence (-) of ATP, 1.6 μM recombinant CUGBP2, or U2 snRNA cleavage. Location of complexes E, A, B, and C are shown at left. (D) Schematic of Ad1 pre-mRNA with the L1 and L2 exons (rectangles) and intron regions (lines) shown. Sizes in nucleotides are indicated above the schematic and the sequence of the 3′ splice site region is shown below with the branch site region underlined and the branchpoint adenosine indicated with a capital A.

seems that CUGBP2 has a stabilizing effect on the nuclear extract and actually appears to enhance Ad1 splicing.

To provide support for the RNA binding model in Figure 9, I asked if the RRM2_3 mutant was capable of blocking branchpoint formation and *in vitro* splicing of the DUPNI pre-mRNA. Both branchpoint formation (Figure 19A, lanes RRM2_3) and *in vitro* splicing reactions (Figure 19B, lanes RRM2_3) were inhibited by the addition of recombinant RRM2_3 to the nuclear extract. Furthermore, the RRM1_2 mutant was insufficient for branchpoint (Figure 19A, lanes RRM1_2) or *in vitro* splicing (Figure 19B, lanes RRM1_2) inhibition, consistent with a model involving branchpoint inhibition by physical occlusion of the branch site region by binding to flanking interaction sites.

Surprisingly, the ΔDD mutant activated branchpoint formation (Figure 19A, lanes ΔDD) and *in vitro* splicing (Figure 19B, lanes ΔDD) of the DUPNI pre-mRNA. A possible explanation for this is that in the absence of the divergent domain, RRM1 may interact with enhancing factors to activate splicing. Additionally, differences in the charge between RRMs 2 and 3 in the RRM2_3 (negative) and ΔDD (positive) mutants may contribute to the different functions of these proteins (RRM2_3 silences and ΔDD enhances). In the future, it would be interesting to investigate the role of RRM1 in splicing enhancement or the effects of charge differences in the divergent domain on protein function.

Note that the DUPNI two-exon substrate used in these reactions contained the NI exon and upstream exon. The RNA substrate containing the NI exon and downstream exon did not splice *in vitro,* probably because of a weak 5´ splice site at the NI exon. Therefore, it is likely that splicing of the upstream intron activates splicing of the downstream intron. If NI splicing naturally occurs in this order, it makes sense that CUGBP2 inhibits the first step.

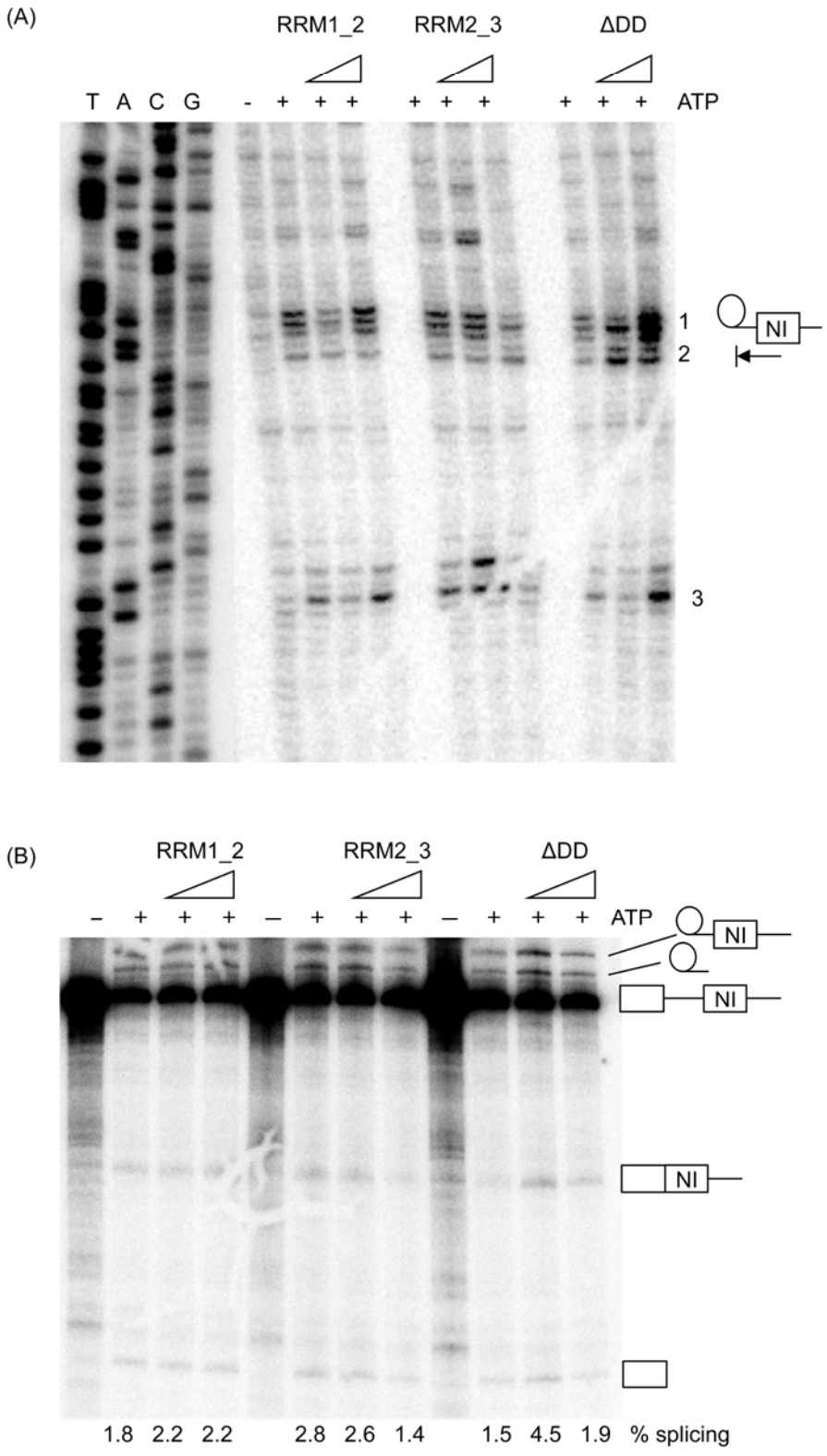


Figure 19. RRMs 2 and 3 are sufficient for branchpoint inhibition. (A) CUGBP2 mutants have differential effects on branchpoint formation on the DUPNI pre-mRNA substrate. *In vitro*

splicing was carried out for 60 minutes in the presence (+) or absence (-) of ATP and 0.8 or 1.6 μM recombinant CUGBP2 mutant proteins (wedge) as indicated above the gel. Branchpoints were detected by primer extension with the JBE5-2 primer as indicated by the schematic at right. Branchpoints in the gel are numbered as in Figure 15. A sequencing ladder generated from the pBSDUPNI plasmid with the JBE5-2 primer is shown for reference (lanes T,A,G,C). (B) CUGBP2 mutants have differential effects on *in vitro* splicing of the DUPNI pre-mRNA substrate. *In vitro* splicing was carried out for 60 minutes in the presence (+) or absence (-) of ATP or recombinant CUGBP2 mutant proteins, RRM1_2, RRM2_3, or ΔDD, as indicated above the gel. Wedge indicates 0.8 and 1.6 μM protein. Splicing intermediates and products are labeled at right. Below the gel, % splicing is indicated as the intensity of the lariat 3′-exon intermediate band normalized to unspliced pre-mRNA for each lane.

3.6 CUGBP2 INHIBITS COMPLEX A FORMATION AND U2 snRNP BINDING

To determine what step before branchpoint formation is specifically affected by the addition of CUGBP2 to the splicing reaction, I analyzed complex assembly on the E5-10 RNA substrate in the presence or absence of recombinant CUGBP2. I confirmed that CUGBP2 blocks U2 snRNP association because CUGBP2 inhibited complex A formation on the wild type substrate but not on the triple mutant substrate (Figure 20A). I verified that the complex was complex A and involved U2 snRNP because it did not form upon U2 snRNA cleavage (Figure 20A/B). In contrast, parallel samples assembled in the absence of ATP showed no effect of CUGBP2 on complex E (Figure 20C). I also carried out splicing complex formation on the Ad1 pre-mRNA in the presence and absence of recombinant CUGBP2 as a negative control (Figure 18C). Here, Ad1 splicing complex formation was resistant to CUGBP2.

The results shown above are consistent with a model in which site-specific binding of CUGBP2 surrounding the branch site region mediates exon skipping. Because exon definition could potentially affect branchpoint formation in the upstream intron by interactions involving U1 snRNP and U2AF across the exon, I asked whether strengthening the 5´ splice site of the NI exon would antagonize the silencing effect of CUGBP2. For this purpose, I increased the complementarity of the 5´ splice site to U1 snRNP and tested the ability of CUGBP2 to induce silencing *in vivo*. This mutation had no detectable effect on silencing by CUGBP2 (Figure 21). I also show that U2AF and CUGBP2 can contact the same RNA at the same time indicating that branchpoint inhibition occurs after U2AF but before U2 snRNP binding (Figure 22). This, together with the lack of effect of CUGBP2 on complex E, which contains U1 snRNP, U2AF, and SF1, is consistent with a mechanism involving inhibition at a step subsequent to exon

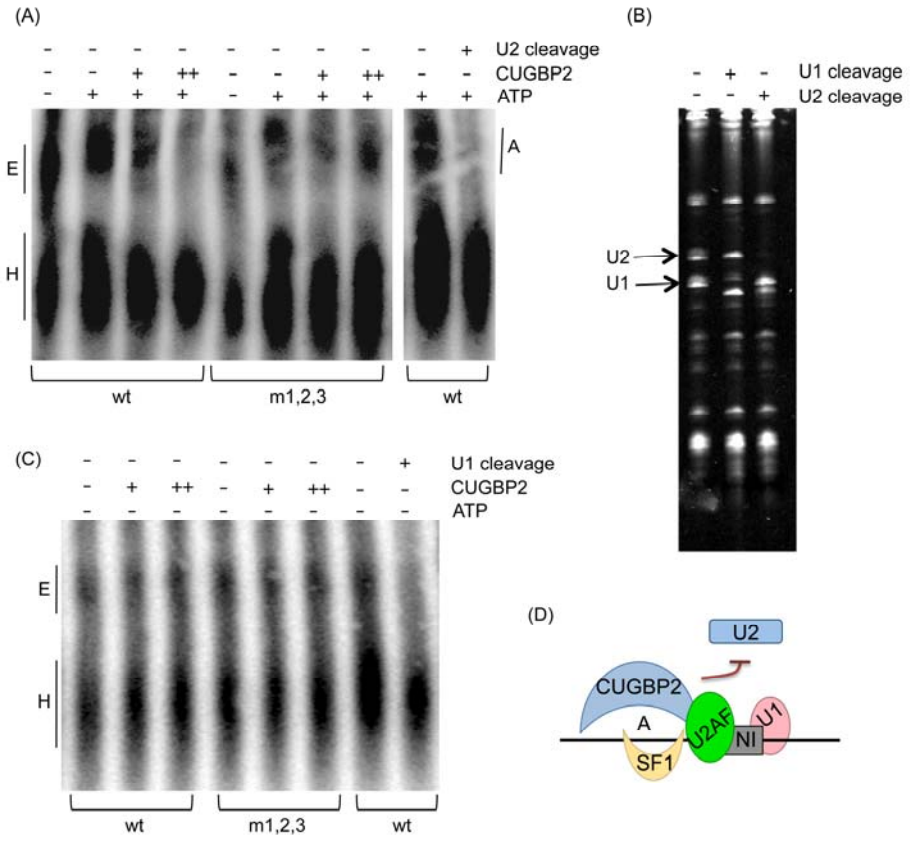


Figure 20. Recombinant CUGBP2 inhibits U2 snRNP binding and complex A but not complex E formation. (A) CUGBP2 inhibits complex A formation and U2 snRNP binding. Splicing complex formation was carried out in the presence (+) or absence (-) of ATP, 1.6 (+) or 3.2 (++) μM CUGBP2, or U2 snRNA cleavage as indicated at top on either wild type (wt) or triple mutant (m1,2,3) E5-10 RNA substrates as indicated below for 15 minutes at 30°C. The position of the ATP-independent complexes E and H are shown at left and the ATP-dependent complex A is shown at right. (B) Conformation of oligonucleotide-directed cleavage of U1 and U2 snRNA (indicated at top). Positions of uncleaved U1 and U2 snRNA are shown at left. (C) CUGBP2 does not inhibit complex E formation. Splicing complex formation was carried out as in (A) except ATP was omitted and reactions were incubated for only 8 minutes. U1 snRNA cleavage is shown to confirm the identity of the U1 snRNP containing complex E. (D) Model for splicing complex assembly on the E5-10 wild type RNA substrate in the presence of CUGBP2.

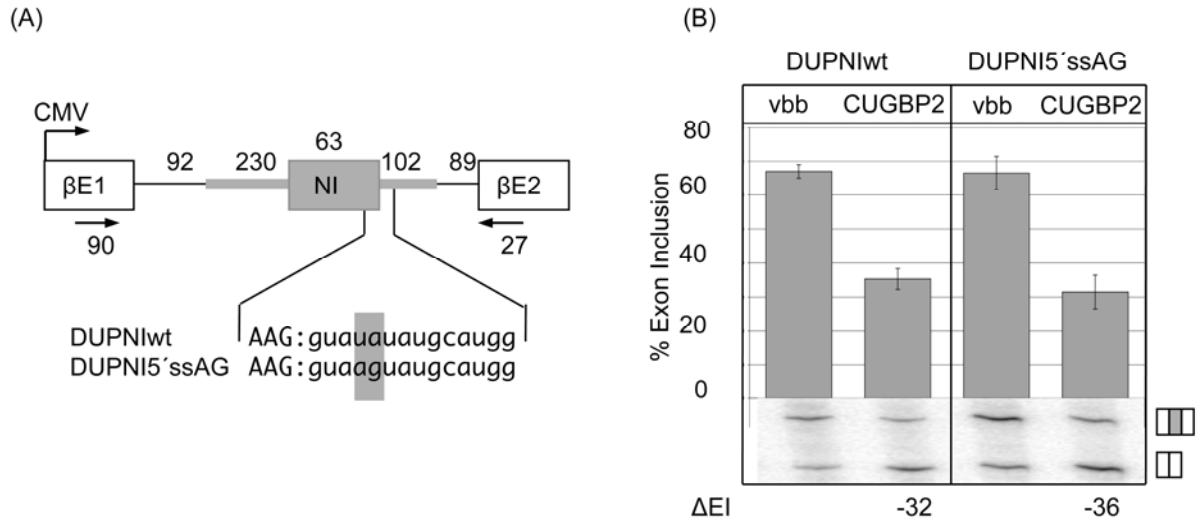


Figure 21. Strengthening the NI exon 5´ splice site complementarity to U1 snRNP does not affect splicing silencing by CUGBP2. (A) Schematic of mutations made to the 5´ splice site of the DUPNIwt splicing reporter to strengthen U1 snRNP binding (DUPNI5´ssAG, shaded nucleotides). Schematic is as described in Figure 10. (B) *In vivo* splicing reporter assay with co-expression of a vector backbone control (vbb) or CUGBP2 protein and the DUPNIwt or DUPNI5´ssAG splicing reporter as indicated at top. Gel, bar graph, and ΔEI calculation are as in Figure 10.

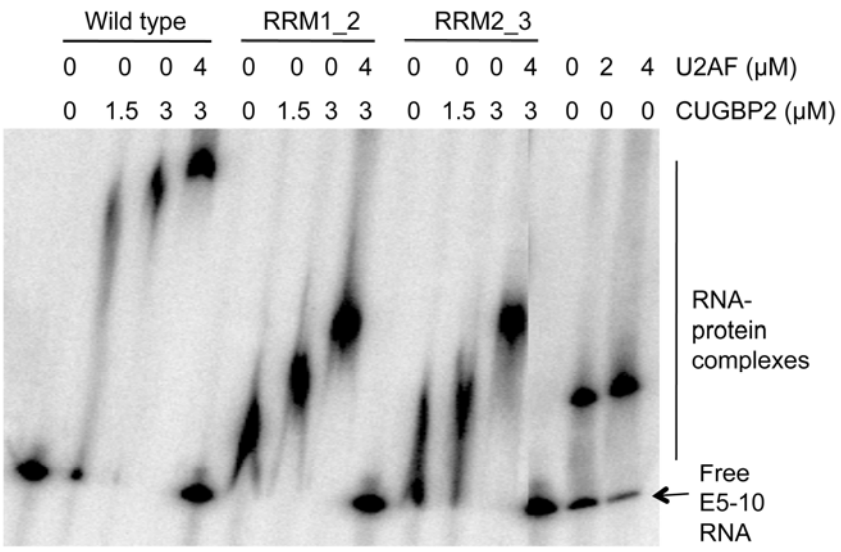


Figure 22. CUGBP2 and U2AF can bind to the same RNA at the same time. Increasing amounts of recombinant wild type or mutant CUGBP2 were bound to E5-10 RNA in the presence or absence of HeLa purified U2AF. Protein concentrations are labeled on the top of the gel and free RNA and RNA-protein complexes are labeled at right. U2AF binding alone to the E5-10 RNA substrate is shown for comparison (last three lanes at right).

definition. Thus, the inhibitory role of CUGBP2 is likely to involve direct antagonism of U2 snRNP binding at the NI branch site region.

3.7 DISCUSSION

In this study, I reveal a novel mode of exon silencing by the CUGBP2 splicing factor. I clearly demonstrate that CUGBP2 binds to an arrangement of functional GU-rich motifs at the perimeters of the NI branch site region. I show that by binding to flanking interaction sites, CUGBP2 inhibits branchpoint formation and splicing at the upstream intron. I provide additional support for a model involving direct branchpoint inhibition by demonstrating that CUGBP2 inhibits U2 snRNP association but has no effect on exon definition events. Taken together, I conclude that the silencing face of the CUGBP2 splicing factor silences an ensemble of branchpoints from flanking interaction sites to cause skipping of the NI exon. This study underscores the importance of the branch site region in alternative splicing regulation and provides a framework to study splicing regulation of the CI cassette and other CUGBP2 target exons throughout the transcriptome.

4.0 A PREDICTIVE CODE FOR CUGBP2 REGULATION WITH IMPLICATIONS FOR AUTOREGULATION

4.1 INTRODUCTION

The CUGBP2 splicing factor is highly enriched in the brain where it regulates alternative splicing and has a general impact on cognitive brain function (Philips et al., 1998). However, only a few CUGBP2 target exons have been identified. Therefore, I wanted to identify a code for splicing silencing that could be applied to additional CUGBP2 target exons. Regulation of predicted target exons would provide general support for the branch site-perimeter-binding model and will be useful in understanding the complete network of exons that are regulated by CUGBP2.

4.2 PREDICTION, IDENTIFICATION, AND CHARACTERIZATION OF CUGBP2 TARGET EXONS

A recent publication by Yeo, et al. (2007) used computational approaches to identify intronic splicing regulatory elements (ISREs) in the introns upstream and downstream from skipped exons (Yeo et al., 2007). One of the ISREs identified was a UGUGUU motif with the propensity to be found within 400 nucleotides of conserved skipped exons. The Yeo study identified 168 skipped exons with a UGUGUU motif in their upstream intron. I obtained this list for further analysis.

In order to extend this analysis to identify additional exons that are potentially silenced by CUGBP2, I searched the list of 168 exons for the following sequence features: (1) the presence of conserved pairs of UGUGU and GU motifs within 100 nucleotides of the 3´ splice site of the skipped exon with a spacing of 10-30 nucleotides between the motifs, and (2) the presence of potential branch site(s) between the motifs. Because the m1 and m2 motifs are sufficient to inhibit branchpoints A1 and A2 on the DUPNI substrate, I rationalized that two motifs flanking the branch site would be sufficient to predict CUGBP2 regulation. Potential branch sites were required to match the human consensus sequence, YUNAY (Y, pyrimidine; N, any nucleotide) with one mismatch allowed (Gao et al., 2008). From this analysis I determined that 48 of the 168 exons (29%) fit these criteria. I chose 27 exons to test for regulation by CUGBP2. Sequence features of regulated exons and controls are shown in Figure 23.

To analyze the response of these endogenous exons to CUGBP2 overexpression, I optimized a calcium phosphate transfection method to obtain >90% transfection efficiency in HEK293T cells. HEK293T cells were chosen for these experiments, since there is no detectable expression of CUGBP2 (Figure 24A, Western blot). RNA was harvested from the cells and the test exon region was amplified by RT-PCR with primers specific for the flanking exons. Ten predicted exons showed an increase in exon skipping when CUGBP2 was overexpressed (Figure 24A, panels MAP4_E15, SORBS1_E5, PPF1BP1_E19, SMARCE1_E4, FOX2_E11, CUGBP2_E6, NFAT_E2, CTBP1_E2, PTER_E3, and MLLT10_E13). Of the 17 exons that were not affected by CUGBP2, one was constitutively included and resistant to CUGBP2 (NDRG3_E3), 8 were not expressed in HEK293T cells (HNT_E7a, MED15_E5, SEC14L4_E8, RGS7_17a, KCNMA_E19a, SCUBE1_E6a, CKLF_E3, DYNC1I1_E5), and 8 were always skipped, therefore, CUGBP2 could not induce additional skipping (MYO9A_E26a, PLCH1_E21, SRPK2_E14a, CTNND1_E16, BNIP2_E4a, AFAP1_E13, BBX_E3a, ATP6VOA1_E19). Taken together, 10 out of 11 testable exons were regulated by CUGBP2 indicating that I have

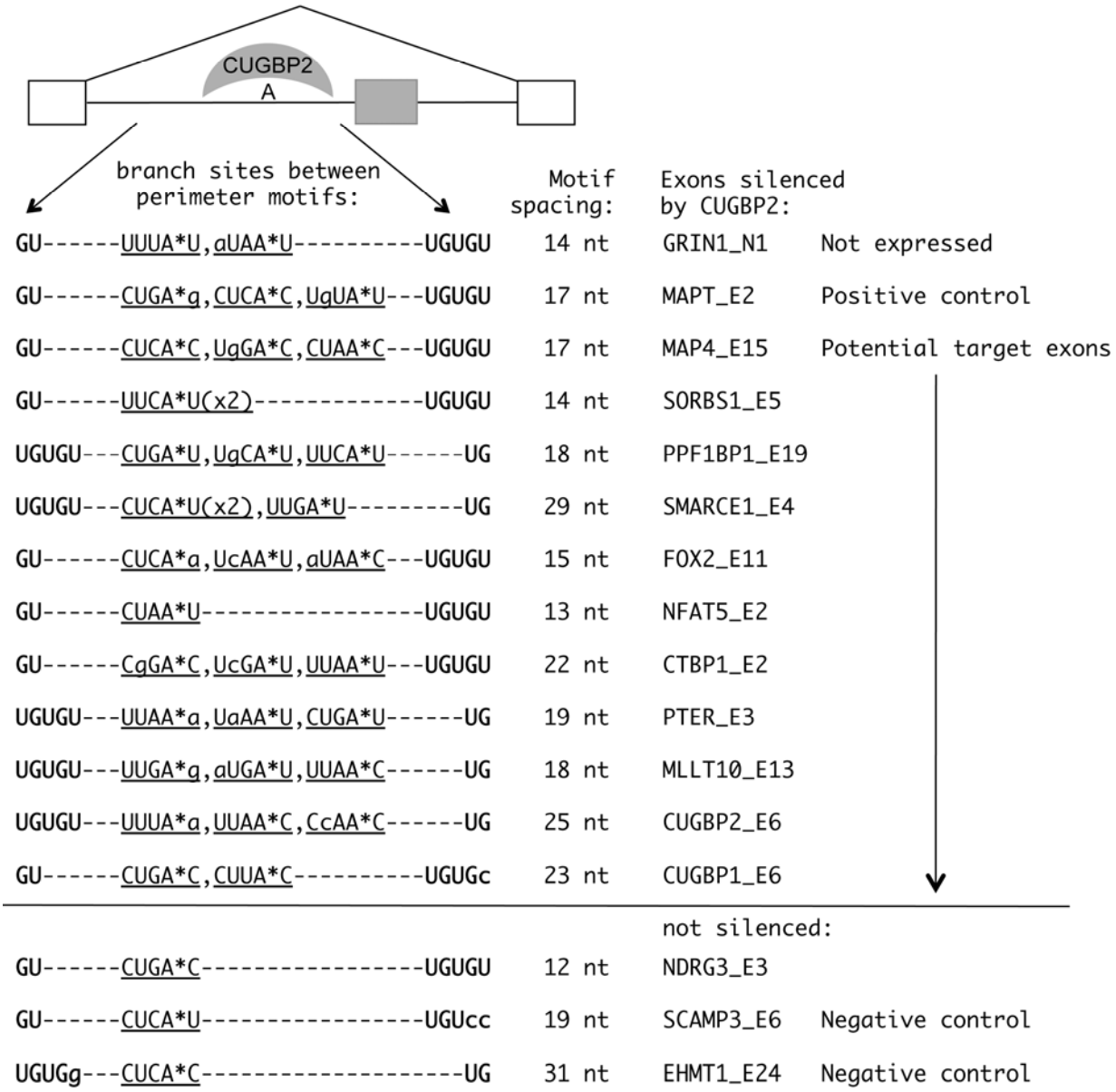


Figure 23. Prediction of additional CUGBP2 target exons. Model depicting the inhibition of branchpoint formation by the binding of CUGBP2 to flanking interaction sites. Below schematic: For all predicted target exons, the gene name and exon number are labeled and the predicted branch sites between GU-rich motifs are shown. Branchpoint adenosine (A*); lowercase letters indicate mismatches to the branch site consensus. Spacing in nucleotides (nt) between the perimeter motifs is shown for all tested exons.

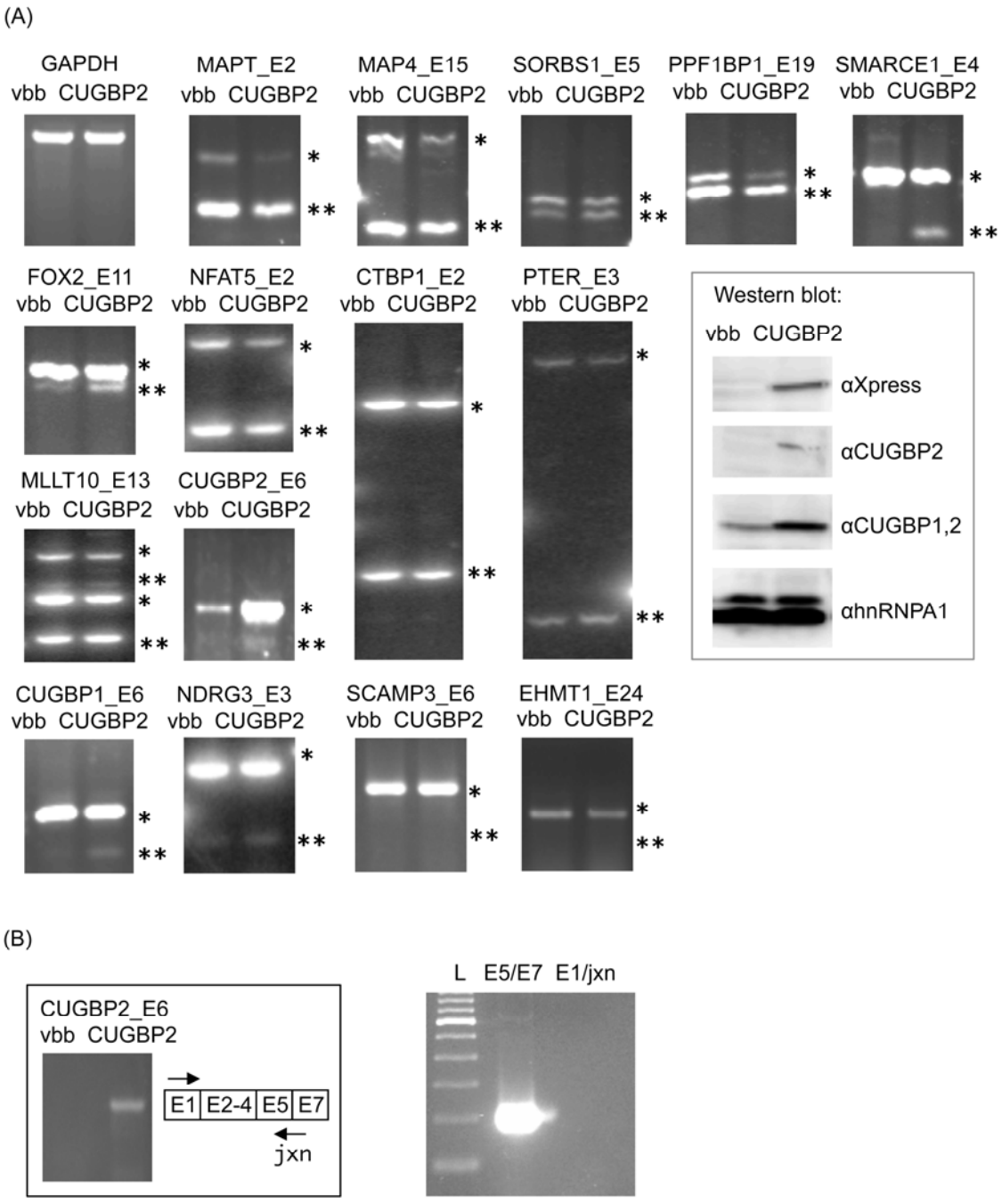


Figure 24. A motif code for splicing silencing reveals novel endogenous exons that are silenced by CUGBP2. (A) HEK293T cells were transfected for 36 hours with CUGBP2 protein expression vector (CUGBP2) or vector backbone control (vbb) prior to RNA isolation and RT-PCR analysis of endogenous target exons. Exon included (*) and skipped (**) RT-PCR products are shown after separation on 2% agarose gels. The gene name and exon number are labeled

above each panel. Note that relative band locations are not shown to scale. For MLLT10, primers are specific for exons 12 and 15 with alternatively spliced exons 13 and 14 in between. Splicing of exon 14 is resistant to CUGBP2. Note that included and skipped bands are designated as exon 13 included and skipped. Inset: Western blot analysis confirming CUGBP2 expression (αXpress); total CUGBP2 as detected by 1H2 antibody (αCUGBP2), CUGBP1 and 2 using 3B1 antibody (αCUGBP1,2), and loading control (αhnRNPA1). (B) Black box: RT-PCR detection of the E5/E7 junction (jxn) of the endogenous CUGBP2 transcript with overexpression of CUGBP2. Right, the E1/jxn primer set cannot amplify the CUGBP2 protein expression vector but the E5/E7 primer set can. Lane L, 100 bp ladder.

identified a specific code that can be used to accurately predict CUGBP2 regulation. For all of the confirmed target exons, the core and flanking motifs were separated by 14-29 nucleotides, and multiple branch site candidates were located between the CUGBP2 motifs (Figure 23). Note that exon 6 of the SCAMP3 transcript, which contains two mismatches to the GU-rich motif pattern, was insensitive to silencing by CUGBP2. MAPT_E2 is shown as a positive control because GRIN1 is not expressed in HEK293T cells (data not shown). Note that across all CUGBP2 target exons, it does not seem that the strength of the 5′ and 3′ splice sites as determined by ESEfinder 3.0 (Cartegni et al., 2003; Smith et al., 2006) influences CUGBP2 regulation, providing support for CUGBP2 acting independently of exon definition events (Table 3). Furthermore, the size of the exon does not appear to influence the ability of CUGBP2 to function as a splicing silencer. Taken together, these data suggest that branch site inhibition from flanking interaction motifs is a general mechanism for CUGBP2 splicing silencing and that CUGBP2 does not act on a subset of alternatively spliced exons.

An interesting observation was the appearance of an exon skipped product of the CUGBP2 transcript itself, which was specific for conditions in which CUGBP2 was overexpressed (Figure 24A, CUGBP2_E6). However, the primers in this case also amplified mRNA expressed from transfected CUGBP2, thereby complicating interpretation. For this reason, I designed a downstream primer specific for the junction between exons 5 and 7, since such a junction primer should amplify only the skipped product from the endogenous mRNA. The junction primer was used together with a forward primer specific for the first exon. The results with this primer set clearly showed the accumulation of the exon 6 skipped version of the endogenous CUGBP2 transcript upon overexpression of CUGBP2 (Figure 24B, black box). Note that this primer set (E1/jxn) did not amplify the CUGBP2 protein expression vector or mRNA from transfected CUGBP2 (right). It is also important to note that although CUGBP2 protein is not detected by Western blotting, there are low levels of CUGBP2 RNA in HEK293T cells. This may indicate that trace amounts of the protein are present in these cells or that the

mRNA is translationally repressed. Furthermore, because there is an enrichment of CUGBP2 protein in the rat cerebral cortex and a deficiency in the cerebellum (Zhang et al., 2002), I predicted and confirmed that there would be more skipping of exon 6 in the cortex (Figure 25). To establish the identity of the exon 6 skipped product, I cut the band out of the gel, cloned and sequenced it. The cloned product exactly matched the exon 5-7 junction sequence demonstrating its identity as the skipped product.

Table 3. Summary of target exon size and strength.

Transcript_exon	ΔEI ± SD	Exon Size (nts)	3´ss score	5´ ss score
GRIN1_NI	NA	63	4.32	8.33
MAPT_E2*	-5.3 ± 0.6	87	6.84	10.26
MAP4_E15*	-11.6 ± 3.5	93	9.54	8.70
SORBS1_E5	-8.0 ± 1.7	27	5.01	9.71
PPF1BP1_E19*	-12.3 ± 2.1	33	5.30	5.91
SMARCE1_E4	-21.6 ± 5.5	105	7.58	7.81
FOX2_E11	-10.0 ± 2.0	40	4.06	7.83
NFAT5_E2*	-9.0 ± 1.7	54	10.00	6.29
CTBP1_E2	-5.0 ± 1.0	195	4.23	5.42
PTER_E3	-32.0 ± 5.0	480	8.43	8.66
MLLT10_E13*	-13.3 ± 2.3	33	3.77	9.70
CUGBP2_E6	NA	80	11.23	4.44
CUGBP1_E6	-4.7 ± 0.6	80	11.68	10.15
NDRG3_E3	-1.3 ± 0.6	36	10.23	2.90
SCAMP3_E6	0	160	10.58	9.98
EHMT1_E24	0	79	12.47	6.96

Change in exon inclusion (ΔEI) upon CUGBP2 expression (±, standard deviation), exon size in nucleotides (nts), 3´ splice site (ss) score, and 5´ ss score of CUGBP2 target exons and controls are listed. Pink, ΔEI >-10%; orange, ΔEI -4 - -9%; asterisk, transcripts that may be transcriptionally silenced upon CUGBP2 overexpression. NA, not applicable.

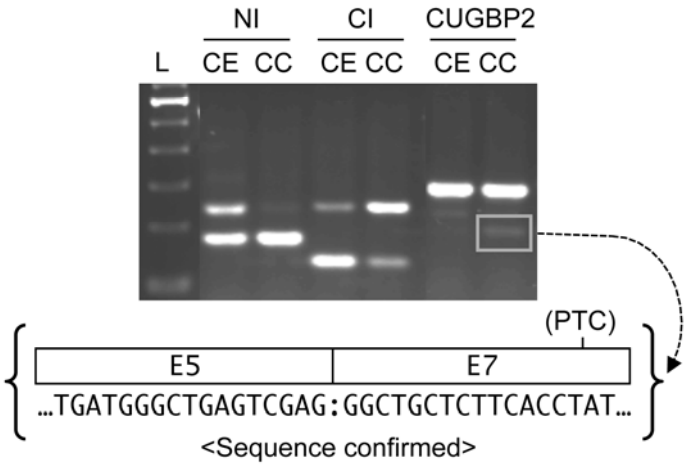


Figure 25. CUGBP2 exon 6 is skipped in the cerebral cortex but not the cerebellum. RT-PCR of RNA from cerebral cortex (CC) and cerebellum (CE) of postnatal day 28 rat shows the E6 skipped mRNA in CC (lanes CUGBP2); white box, E6 skipped product cloned for sequence confirmation. Tissue-specific NMDA R1 receptor splicing patterns are shown as controls (lanes NI and CI). Lane L, 100 bp ladder. Brackets indicate sequence confirmed region of E5/E7 junction. PTC, premature termination codon introduced into the reading frame as a result of E6 skipping.

I also tested an exon in the CUGBP1 transcript that is homologous to CUGBP2 exon 6 (Figure 23, CUGBP1_E6). Here, there is one mismatch to the core motif and although CUGBP2 can silence this exon, the effect is less than other target exons with perfect matches to the consensus motifs (Figure 24, CUGBP1_E6 and Table 3). This suggests that CUGBP2 regulation can be titrated depending on the sequence content and binding affinity to target motifs.

4.3 AUTOREGULATION BY CUGBP2

To determine whether CUGBP2 silences its own exon by a mechanism similar to that shown for the NI exon, exon 6 and the adjacent introns of CUGBP2 pre-mRNA were cloned into the DUP splicing reporter (Figure 26A). In this context overexpression of CUGBP2 had a robust silencing effect changing the exon 6 splicing pattern from 100% to 18% inclusion in transfected HEK293T cells (Figure 26B, lanes Wild type). To address the functional significance of the CUGBP2 binding sites at the boundaries of the predicted branch sites, I tested site-directed mutations in the core (CORE) and downstream (DSM) motifs (Figure 26A). One perfect match to the branch site consensus (A2) and two additional candidates with a single mismatch (A1,A3) are the only plausible branch sites located within 100 nucleotides of the 3´ splice site of exon 6. Mutations in the CORE and DSM motifs resulted in a reduction of splicing silencing by CUGBP2 (Figure 26B, lanes CORE mt, DSM mt), and these effects were additive in the double mutant (lanes CORE/DSM mt). These are similar to the results shown for the NI exon providing additional support for the perimeter-binding model. Note that, although mutations did not completely eliminate CUGBP2 regulation of exon 6, footprinting experiments documented additional contact points extending from the CORE and DSM motifs suggesting that alternate binding registers

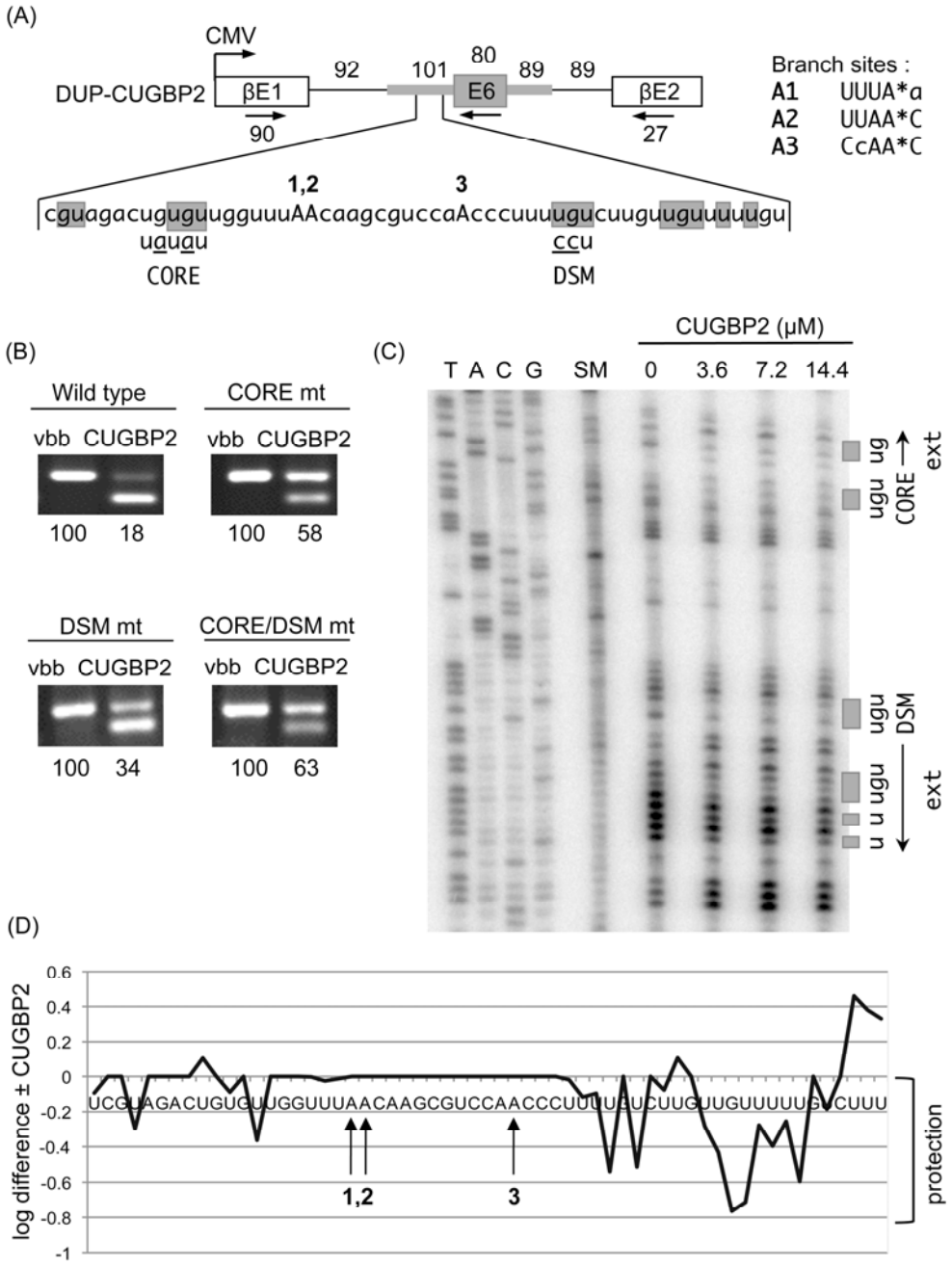


Figure 26. CUGBP2 is autoregulated by silencing from exon 6 branch site perimeters. (A) DUP-CUGBP2 splicing reporter with sequence showing predicted branch sites (A1, A2, A3) and CUGBP2 binding motifs as detected by RNA footprinting in (D) (shaded regions). Primers (arrows) used for RT-PCR (βE1 and βE2) or footprinting analysis (E6) are shown. Mutations in core (CORE) and downstream (DSM) motifs are indicated (underscores). (B) Autoregulation of

E6 depends upon CUGBP2 binding motifs surrounding the branch site region. Gel panels represent RT-PCR analysis of wild type or mutant derivatives of the splicing reporter co-expressed with vector backbone control (lanes vbb) or CUGBP2 protein expression vector (lanes CUGBP2) in HEK293T cells and are representative of three separate experiments. Percent exon inclusion values are shown below gel panels. (C) Footprinting analysis of the predicted branch site region upstream of E6 of CUGBP2 pre-mRNA. Starting material (lane SM), or RNA treated with CMCT in the absence (lane 0) or presence (lanes 3.6, 7.2, 14.4 μM) of CUGBP2 protein is indicated. Protected regions are indicated as shaded boxes at right with extended regions of protection (ext) from the CORE or DSM motifs indicated by arrows. Protected regions correspond to shaded nucleotides in (A). (D) Log difference of the band intensities of modified nucleotides in the absence or presence of 14.4 μM CUGBP2. Negative values represent regions protected by CUGBP2. Note that the strength of protection decreases with increasing distance from the primer used for primer extension.

might allow for some residual silencing (see below). Also note that I tested the possible role of an intronic UGUGU motif located 70 nucleotides downstream from exon 6. Mutation of this motif to UAUAU had a negligible effect on splicing silencing by CUGBP2, ruling out effects across the exon and further supporting our model (data not shown). Furthermore, I show that CUGBP1 can also regulate this exon but has a weaker silencing role compared to CUGBP2 in this context, and does not act through the CORE and DSM motifs like CUGBP2 (Figure 27).

Finally, I used RNA footprinting to identify CUGBP2 contact sites in the neighborhood of the predicted branch sites upstream of exon 6 (Figure 26C/D). GU-rich motifs flanking A1, A2, and A3 were protected by the addition of purified CUGBP2 similar to that observed above for the NI 3′ splice site (Figure 26C, last three lanes at right). That is, two protected regions at the borders of the predicted branch sites overlap with the UGUGU core and UG flanking motifs in agreement with the perimeter-binding model. A difference in the pattern of protection, however, was the finding that two sets of motifs on either side of the branch site region extend outward indicating variations in the mode of binding compared to the NI exon.

4.4 DISCUSSION

Here, I have identified a combinatorial code that can be used to accurately predict CUGBP2 target exons. This study underscores the importance of understanding direct nucleotide contact sites and the significance of a functional reference point in the prediction of splicing factor target exons. This code will be useful in the characterization of additional alternatively spliced exons. Furthermore, insights from this study can be applied to systematically examine the role of intronic mutations in the neighborhood of the branch site underlying mechanisms of human disease. The identification of an autoregulated exon in the CUGBP2 transcript and the finding

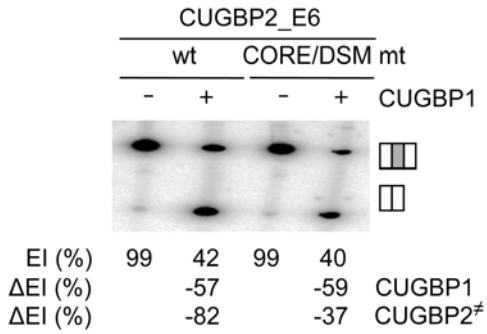


Figure 27. CUGBP2 autoregulatory motifs are specific. Splicing reporters (top) were transfected into C2C12 cells in the presence (+) or absence (-) of CUGBP1 as indicated and included and skipped forms of spliced reporter RNA were assayed by RT-PCR and separated on a polyacrylamide gel. Percent exon inclusion (EI (%)) and the change in percent exon inclusion (ΔEI (%)) with CUGBP1 overexpression is indicated below the gel panel. ≠The effect of CUGBP2 overexpression is shown for comparison.

that the CUGBP1 and Fox2 splicing factors are cross-regulated by CUGBP2, implicates CUGBP2 in a network of splicing factor regulation. This network could have a general impact on the fine-tuning of splicing patterns in a brain region- and developmental- specific manner.

5.0 DISCUSSION

An in depth analysis of the silencing face of the CUGBP2 splicing factor has revealed the functional nucleotide contacts involved in high affinity binding by CUGBP2, a novel mechanism for splicing silencing, and a combinatorial code for CUGBP2 regulation. This study has provided insight into global regulation by this splicing factor and has revealed a role for CUGBP2 in its own autoregulation and cross-regulation of family members. Furthermore, analysis of CUGBP2 mutant proteins has demonstrated that multiple RRMs facilitate multiple recognition events at flanking GU-rich contact sites for high affinity binding. Results from these studies provide a starting point to study the role of CUGBP2 in the regulation of additional target exons and to examine the enhancing face in more detail.

5.1 A COMBINATORIAL CODE FOR SPLICING SILENCING BY CUGBP2

In this study, I identified a combinatorial code that can be used to predict exon silencing by the CUGBP2 alternative splicing factor. I initially used the NI cassette exon of the NMDA receptor NR1 subunit as a model exon to study RNA recognition by this factor. The first hint of how this exon target is recognized was revealed by chemical modification RNA footprinting of a high affinity binding region, which showed two contact sites – a core UGUGU and flanking GU – closely positioned in the neighborhood of the branch sites upstream from the NI exon (Figure 6). These contact sites and a third weaker footprinting site were shown to modulate alternative

splicing of the NI exon *in vivo* (Figure 10). I further demonstrate that transfer of these motifs is sufficient to confer CUGBP2 splicing silencing to an unregulated exon (Figure 13). Furthermore, I use this motif code to predict and identify additional CUGBP2 target exons with an accuracy of 91% (10 out of 11 testable exons were regulated), satisfying the criteria for a splicing code as indicated in Chapter 1. I provide additional support for this code by identifying a similar arrangement of functional motifs in the intron upstream from the autoregulated exon 6 of the CUGBP2 transcript (Figure 26). That a set of motifs flanking a functional reference point can be used in combination to predict regulation by a splicing factor with high accuracy is a novel finding of this study and has general implications for refining computer programs with predictive value.

Our results support and extend those of a previous study, which reported the identification of a UGUGUU motif as an intronic splicing regulatory element (U17) enriched within 400 nucleotides upstream of conserved skipped exons (Yeo et al., 2007). This previous study reported the association of the U17 element with exon inclusion in brain tissue as indicated by microarray analysis. In contrast, this study shows that the UGUGU core of the U17 element is generally associated with exon silencing when the motif is paired with a flanking GU surrounding the functional branch sites. This is not necessarily a discrepancy, but more likely a reflection of a mechanism operating on a subset of exons containing a U17-related element. This observation provides support for the value of the branch site as a functional reference point that can be used together with the precise binding interaction motifs of a splicing factor to computationally predict new splicing regulatory targets.

Our results are consistent with the types of binding motifs identified for CUGBP2 using a SELEX approach, although the relationship of the binding motifs to the branch site region and the autoregulatory role of CUGBP2 were not determined (Faustino and Cooper, 2005). Furthermore, the types of motifs identified for CELF proteins on the α-actinin transcript are in agreement with our results (Suzuki et al., 2002). In this study, the authors demonstrated that

CUGBP2 can bind to a short RNA containing a set of GU-rich motifs but not to a similar RNA containing one GU-rich stretch. Additionally, the GU-rich silencing motifs studied were located upstream of the regulated exon positioned on either side of the branch site. Note that the authors did not comment on either of these observations. In further support of our splicing code, insulin receptor exon 11 and MAPT exon 2 contain a similar arrangement of motifs flanking the branch site region in their upstream introns. GU- and U-rich stretches are also located in the vicinity of the branch site of the aberrantly retained intron of ClC-1 pre-mRNA. Taken together, these observations support the generality of the splicing code for CUGBP2 regulation although the functional significance of these motifs has not been determined.

The splicing code described here provides an explanation for how splicing factors can achieve high specificity given the short and degenerate nature of their cognate RNA binding sites. The use of multiple binding motifs separated by variable space increases the specificity of recognition. Furthermore, the use of the branch site as a functional reference point further increases the specificity of the code. However, it is not known if recognition is mediated by the branch site or by factors interacting at the branch site. The footprinting assay used in this study is limited in that interactions with unmodified nucleotides and loose interactions that allow the chemical to enter the RNA-protein complex would be missed. In the future, we could further refine the code for CUGBP2 regulation by determining the limits in spacing between the GU-rich motifs. At present, it seems that a 12-30 nucleotide space is tolerated, although lower and higher limits have not been tested. Furthermore, we could determine if there is flexibility in the sequence content of motifs. For example, can CUGBP2 recognize a UCUGU core motif or flanking CU? These details will collectively help to refine the code for CUGBP2 splicing silencing.

Microarray analysis comparing wild type primary cortical cultures to those without CUGBP2 would be an alternate approach to identify a complete list of target exons that are either silenced or enhanced by CUGBP2. Analysis of CUGBP2 recognition motifs in target

exons could be used to further refine the code for CUGBP2 regulation. Results from these studies would also be helpful in generating a map of the positional significance of target motifs for either enhancement or silencing by CUGBP2 as was done for Nova splicing factors (Ule et al., 2006). However, such studies are limited because we have not been able to generate a CUGBP2 knock out mouse and lentiviral approaches for shRNA-mediated knock down of CUGBP2 in primary cortical cultures have been unsuccessful so far. Tissue-specific expression of CUGBP2 makes it difficult to carryout siRNA in cell lines and conventional RNAi by transfection of siRNAs into primary cortical cultures is not possible because of very low transfection efficiencies (approximately 10% of neurons are transfected by lipofectamine or calcium phosphate transfection methods). Therefore, the best method currently available to study the effects of CUGBP2 on alternative exons is by overexpression in cell lines lacking endogenous CUGBP2. Still, analysis can be hindered by minimized silencing effects because of endogenous CUGBP1 acting on a subset of target exons or by the lack of expression of many neural-specific transcripts in cell lines tested. Furthermore, cross-regulation of Fox2, CUGBP1, and possibly other splicing factors by CUGBP2 complicates interpretation of overexpression and even RNAi approaches. This cross-regulatory effect probably explains why overexpression of CUGBP2 in N18TG2 cells causes an increase in NI exon inclusion when all three GU-rich motifs are mutated (Figure 10C, lanes m1,2,3). Therefore, it is best to use a combination of experimental approaches to study splicing regulation as was done in this study.

5.2 SPLICING SILENCING BY BRANCHPOINT INHIBITION

In this study, I focused on the silencing face of the dual functional splicing factor, CUGBP2, to understand how it silences the NI exon. Here, the positions of the branch sites were mapped between the core and flanking CUGBP2 target motifs (Figure 15). These branch sites were

collectively inhibited by CUGBP2 with a dependence on the presence of flanking GU-rich binding sites. Thus, guilt-by-association places CUGBP2 at the boundaries of the branch sites it regulates in support of the three-motif occupancy model illustrated in Figure 15A. Furthermore, I clearly demonstrate that CUGBP2 inhibits branchpoint formation by blocking U2 snRNP association (Figure 20) but has no effect on exon definition events (Figures 20, 21, and 22), in support of the silencing model illustrated in Figure 28. The regulation of an ensemble of branchpoints by a perimeter-type binding model, and the discovery that an exon in the CUGBP2 transcript is itself silenced by a similar arrangement of binding motifs, are novel findings of this study.

Branchpoint formation reflects a critical step in the catalysis of the splicing reaction, but its role in the regulation of alternative splicing across the transcriptome represents largely uncharted territory. Only a small number of branchpoints have been experimentally mapped, however, and there are often multiple candidate branch sites in the 3´ splice site region that match the consensus sequence (Gao et al., 2008). Examples of alternative splicing regulation through the use of a suboptimal branchpoint include the calcitonin/calcitonin gene-related peptide exon 4 and fibroblast growth factor receptor 2 exon IIIc (Emeson et al., 1989; Adema and Baas, 1991; Zandberg et al., 1995; Hovhannisyan and Carstens, 2005). Branch site selection has also been implicated in the regulation of mutually exclusive exons of beta tropomyosin and in the processing of human growth hormone pre-mRNA (Hartmuth and Barta, 1988; Helfman and Ricci, 1989).

What advantages would a perimeter-binding model provide for the control of access to the branch site region? The pre-mRNA branch site and flanking sequences are sequentially contacted by several factors during spliceosome assembly (Staknis and Reed, 1994b; Chiara et al., 1996). The branch site interacts with the RS domain of $U2AF^{65}$ bound to the polypyrimidine-tract of the 3´ splice site, followed by interactions with the RS domain of an SR splicing factor

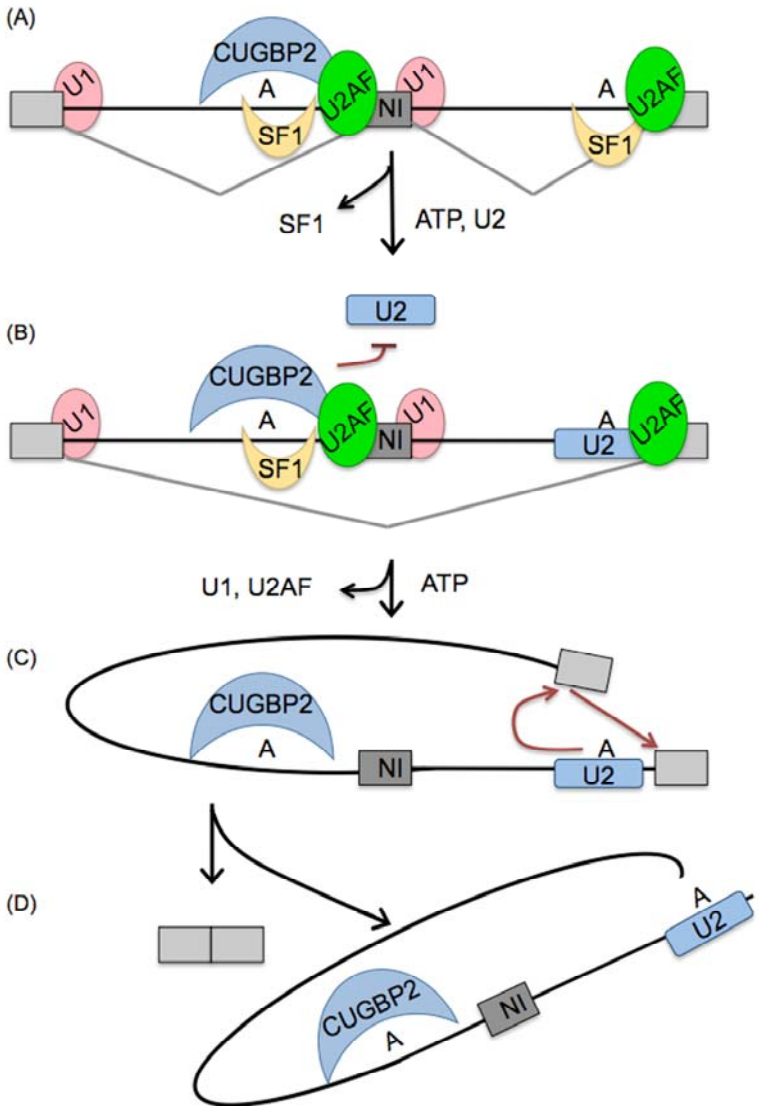


Figure 28. Model for splicing silencing by CUGBP2. (A) U1 snRNP, U2AF, and SF1 bind to the 5′ and 3′ splice sites and branch site sequence for effective complex E formation on the NI substrate containing the exon (rectangle) and flanking introns (lines) to define the exons to be spliced. (B) CUGBP2 at the perimeters of the branch site inhibits ATP-dependent U2 snRNP association and complex A formation. U2 snRNP can effectively replace SF1 at the downstream branch site to position the branchpoint adenosine for nucleophilic attack. (C) Two ATP-dependent transesterification reactions and several spliceosomal rearrangements (not shown) occur. (D) The upstream and downstream exons are joined and the lariat intron containing the upstream intron, NI exon, and downstream intron is removed.

bound to an enhancer element in the adjacent exon (Valcarcel et al., 1996; Shen and Green, 2004; Shen et al., 2004). SF1 makes direct contacts with the branch site during complex E assembly (Abovich and Rosbash, 1997; Liu et al., 2001). In complex A, SAP155 binds to sites flanking the branch site and replaces SF1 to recruit U2 snRNP (Berglund et al., 1998; Gozani et al., 1998). Here multiple contact sites furnish CUGBP2 with the stability to inhibit the association of U2 snRNP with the branch site region to block conformational transitions of the spliceosome. By binding to the perimeters of the branch site, CUGBP2 may allow for SF1 association during complex E but block access of the U2 snRNP-associated SAP155. Moreover, a recent study has demonstrated that CUGBP2 can physically interact with SF3b of the U2 snRNP and RNA helicases DDX1 and DHX9 among other proteins (Goo and Cooper, 2009). Therefore, CUGBP2 may in part function to recruit a helicase to the U2 snRNP-pre-mRNA duplex to disrupt base pairing if complex A has already formed. However, the spliceosome assembly assays used here are limited in that they cannot differentiate between these possibilities. The perimeter-type binding model described here is significant in allowing for the coordinate regulation of multiple branchpoints to control alternative splicing of a cassette exon. Furthermore, by binding in alternate registers around the branch site region, CUGBP2 may be able to inhibit association of other factors on different RNA substrates and therefore function by alternate mechanisms.

Why would CUGBP2 inhibit complex A and not complex E, allowing the exon to be defined but not spliced? In this way, the splicing machinery could more quickly adjust to environmental cues, such as cell depolarization. Post-translational modification of CUGBP2 in response to environmental signals, for example, may adjust its binding and silencing ability. Quick removal of CUGBP2 from the pre-mRNA would allow access of U2 snRNP and assembly of complexes A, B, and C for catalysis of splicing. Inhibiting a step later than exon definition, therefore, may allow for faster adjustments in exon selection, which would be particularly useful in the nervous system where CUGBP2 is expressed.

5.3 MODULAR BINDING BY CUGBP2 FACILITATES HIGH AFFINITY AND SEQUENCE SPECIFICITY

Here I demonstrate that multiple RRMs collaborate for high affinity binding and sequence recognition by CUGBP2. I show that RRMs 2 and 3 are sufficient for protection of flanking GU-rich sequence elements in the footprinting assay (Figure 8). In support of this, RRM3 is necessary but RRM1 and the divergent domain are dispensable for high affinity binding by CUGBP2 (Table 2). It is likely that RRMs 2 and 3 bind to the upstream and core motifs, respectively, with the divergent domain allowing for flexibility in the spacing between motifs (Figure 9). In support of this, the ΔDD mutant binds to the same motifs as the wild type protein (Figure 8) but with half the affinity (Table 2). It is likely in this case that two proteins bind to the upstream and core motifs with less affinity than one protein binding to both of these motifs as for the wild type protein. That is, without the divergent domain to position RRMs 2 and 3 relative to each other, one protein cannot simultaneously contact distance sites. The finding that flanking interaction sites are necessary for high affinity binding and the suggestion that multiple RRMs cooperate to facilitate this binding are of interest to this study.

CUGBP2 is a modular protein containing three RRMs in which a divergent domain of unknown function separates RRMs 2 and 3. The domain structure of the protein may be geared to facilitate binding of a monomer to a pair of core and flanking motifs forming a bridge between them as our model indicates (Figure 9). Alternatively, a single monomer might bind to all three GU-rich motifs. Both models would limit access to the branch sites by factors sliding along the RNA from upstream and downstream directions. Our footprinting results with CUGBP2 are in agreement with previous structural studies showing that a single RRM can contact ~2-7 nucleotides of its bound RNA ligand (Figure 8) (Ding et al., 1999; Oberstrass et al., 2005; Sickmier et al., 2006). Given the inherent flexibility of RNA binding proteins, it would not be surprising that breathing motions of CUGBP2 could adjust the relative conformations of the

RRM domains to optimize recognition specificity in different sequence contexts. Still, interpretation of RRM mutant studies is limited because charge differences in the linker region between RRMs 2 and 3 may alter binding characteristics. Furthermore, individual RRMs may act differently when taken out of the wild type context. In the future, structural studies will be required to completely understand the topology of binding associated with CUGBP2's silencing function.

5.4 AUTOREGULATION BY CUGBP2

Here I reveal a novel autoregulatory role for CUGBP2. I demonstrate that overexpression of CUGBP2 causes an increase in exon 6 skipping in HEK293T cells (Figure 24). Furthermore, the exon 6 skipped product is detected in mRNA prepared from the rat cerebral cortex but not cerebellum, consistent with the brain region-specific expression of CUGBP2 (Figure 25). I verify that the exon 6 skipped product exactly matched the exon 5/7 junction and was not a non-specific PCR band. Moreover, in the context of a cloned splicing reporter, CUGBP2 overexpression caused increased skipping of exon 6, with a dependence on GU-rich motifs flanking the predicted branch sites (Figure 26). Finally, I demonstrate that CUGBP2 physically contacts GU- and U-rich motifs at the perimeters of the exon 6 branch sites, in support of the perimeter-binding model described for the NI exon.

Autoregulation has been shown for a growing number of splicing factors, including PTB, Fox2, Nova1, SRp20, SC35, TIA1, and TIAR (Wollerton et al., 2004; Dredge et al., 2005; Baraniak et al., 2006; Ni et al., 2007). Here I dissect the mechanism of CUGBP2 autoregulation in great detail, making this one of the best-characterized examples of autoregulation. Conceptual translation reveals that skipping of exon 6 causes a shift in the reading frame, which introduces a PTC in the exon 7 region of the transcript. In this way, the resulting transcript could

be targeted for NMD. Conversely, translation of the exon 6 skipped transcript could generate a truncated protein ending within RRM2. The advantage of a tight motif arrangement around the branch site region would be to dynamically adjust exon 6 inclusion based on fluctuations in the levels of CUGBP2 protein.

The observation that CUGBP2 can cross-regulate exons in the CUGBP1 and Fox2 transcripts (Figure 24), and that CUGBP1 can silence CUGBP2 exon 6 to a lesser extent (Figure 27), implicates CUGBP2 in a network of splicing factor regulation. I speculate that this may be important in specifying neural cell identity and for fine-tuning of neural exon splicing. In the future, it would be of interest to study the differences in binding specificities and target exon selection by CUGBP1 and CUGBP2.

5.5 FINE-TUNING OF SPLICING IN THE NERVOUS SYSTEM BY CUGBP2

The binding of CUGBP2 to the perimeters of the branch sites allows for sensitive gradations specifying the levels of NI exon inclusion. Because the peptide region encoded by the NI exon modulates sensitivity of the NMDA receptor to zinc ions, protons, and polyamines, such a mechanism would be advantageous for fine-tuning this modular property of receptor function in different regions of the brain or during development (Lipscombe, 2005). The biochemical functions of NMDA receptors are of fundamental importance in synaptic plasticity where deficits in this subunit are associated with altered brain function in the context of Alzheimer's disease. The CI cassette exon, which is regulated by the enhancing face of CUGBP2, encodes a functionally important region of the receptor involved in membrane trafficking and signaling to the nucleus. Additional CUGBP2 target exons function in processes such as alternative splicing (CUGBP2, CUGBP1, FOX2), transcription (NFAT5, MLLT10, SMARCE1, CTBP1), microtubule assembly (MAP4, MAPT), axon guidance (PPFIBP1), and insulin signaling (SORBS1, insulin

receptor). These processes have a general impact on neuronal function and fine-tuning of splicing of these and additional factors in the nervous system by CUGBP2 may be associated with specifying neural cell identity.

5.6 INSIGHTS INTO THE MECHANISMS OF SPLICING ENHANCEMENT BY CUGBP2

This study provides the starting point to investigate the broader roles of CUGBP2 in regulation of the CI cassette and additional exons throughout the transcriptome. How do RNA elements recognized by CUGBP2 for splicing enhancement differ from those recognized for splicing silencing? Are different nucleotides recognized for enhancement or are the positions of the motifs relative to the exon more important? Currently, cTNT exon 5 and the CI cassette are the only exons known to be enhanced by CUGBP2. In support of the positional significance of motifs for regulation, both of these exons contain GU-rich motifs in their downstream introns. Furthermore, several alternative splicing factors have been shown to bind near the 3´ splice site to silence an alternative exon and in the downstream intron for enhancement (Baraniak et al., 2006; Ule et al., 2006). Therefore, mechanisms of dual regulation by CUGBP2 may be similar to those used by other splicing factors.

I have footprinted CUGBP2 to a core UCUGU and downstream flanking UG dinucleotide in the intron downstream from the CI exon (data not shown). However, mutation of these sites only had a modest effect on splicing enhancement by CUGBP2. Therefore, downstream from the CI exon additional low affinity sites are likely involved in regulation. In fact, multiple GU-rich sequence elements are located near the 5´ splice site. In the future, it would be of interest to determine the effects of combinations of mutations to these GU-rich motifs for splicing enhancement by CUGBP2. Furthermore, what is the minimal sequence requirement for enhancement in a heterologous context? However, these studies are limited because CUGBP2

regulatory motifs may overlap with other control elements for alternative or constitutive splicing regulation. Additionally, if CUGBP2 functions to antagonize a splicing silencer, transfer assays would not be useful for understanding regulation. Still, a complete understanding of the sequence elements involved in splicing enhancement of the CI exon could provide insight into the mechanisms of dual regulation by CUGBP2 and other splicing factors.

A recent study has demonstrated that CUGBP2 can enhance inclusion of cTNT exon 5 by binding to GU-rich elements near the 5′ splice site in the downstream intron to recruit U2 snRNP to the branch site in the upstream intron (Goo and Cooper, 2009). This mechanism of enhancement is similar to that shown for SR proteins (reviewed by (Graveley, 2000)). Furthermore, CUGBP2 has also been shown to compete with PTB and MBNL splicing silencers for exon 5 enhancement (Ladd et al., 2005). It is likely that a combination of these effects are involved in splicing enhancement of cTNT exon 5 by CUGBP2.

Potential models for splicing enhancement of the CI exon are illustrated in Figure 29. These include a cross exon definition model (A) and antisilencing model (C) as have been demonstrated for cTNT exon 5. Additionally, I suggest an alternate, intron definition model (B). The intron definition model predicts that CUGBP2 at the 5′ splice site interacts with components of the U2 snRNP at the downstream 3′ splice site to facilitate the transition from an exon defined spliceosome to a catalytically active intron defined spliceosome. In support of this model, CUGBP2 interacts with two known helicases (Goo and Cooper, 2009). Therefore, CUGBP2 may recruit a helicase to assist in structural rearrangements of the spliceosome for release of U4 snRNP. Although, these models for CI splicing enhancement are currently only speculative, they provide a good starting point to study the mechanisms of regulation by this factor.

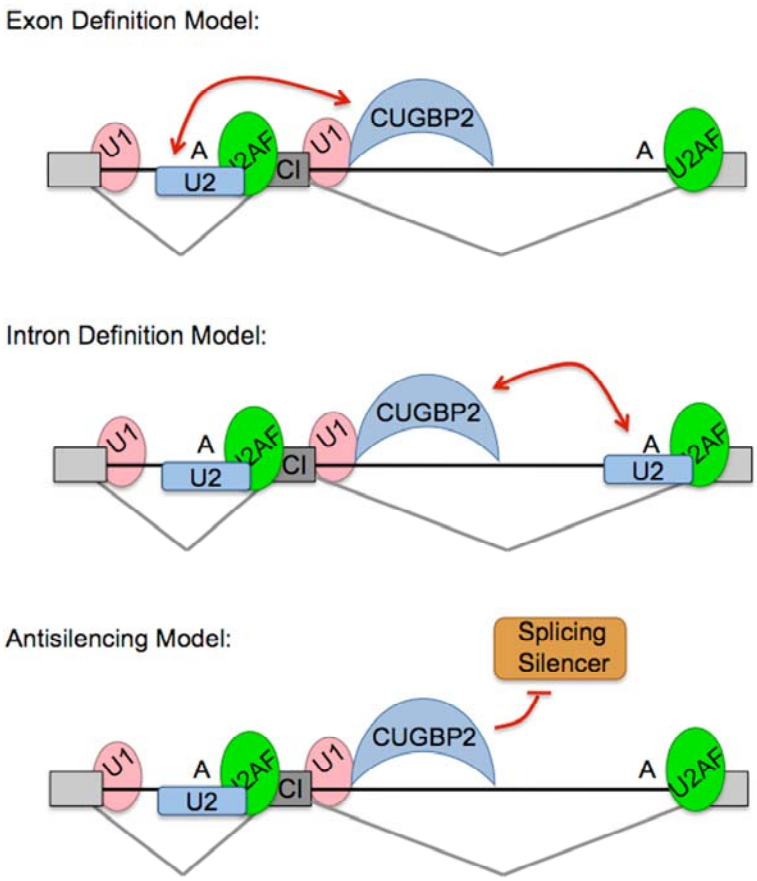


Figure 29. Models for splicing enhancement by CUGBP2. Exon Definition Model: CUGBP2 binds to the intron downstream of the CI exon. Here it may facilitate protein-protein interactions with U2 snRNP across the CI exon to aid in exon definition. Intron Definition Model: Alternately, CUGBP2 may interact with U2 snRNP at the downstream 3´ splice site, across the large downstream intron (>1900 nucleotides), to assist in the transition from an exon-defined spliceosome to an intron-defined spliceosome. Antisilencing Model: CUGBP2 may function to antagonize a splicing silencer in the 5´ splice site region of the CI exon to facilitate exon inclusion**.**

5.7 FUTURE EXPERIMENTS

Although this study has answered many questions about the specificity of binding and mechanism of splicing silencing by CUGBP2, it has raised even more questions. What is the positional significance of target motifs for silencing and enhancement? Does CUGBP2 recognize variations in the GU-rich RNA elements for regulation of a subset of target exons? If so, does CUGBP2 bind with different affinity to different sequence elements to specify splicing regulation based on fluctuations in protein levels? What additional target exons are regulated by CUGBP2? Do these target exons share common characteristics and does splicing of these exons tune modular protein functions in the cerebral cortex? What is the mechanism of splicing enhancement? Does CUGBP2 recruit factors associated with the general splicing machinery or compete with splicing silencers? What is the significance of the autoregulated splicing event? Does it serve as a feedback mechanism to regulate CUGBP2 protein levels? Moreover, how do binding differences of CUGBP1 and 2 direct target exon selection by these related family members? Answers to these questions will provide a more complete understanding of CUGBP2 function and will also offer insight into global mechanisms of alternative splicing regulation.

In the future, it would be of interested to refine the spicing code for CUGBP2 and use this information to generate a computer program to predict regulation by this factor. One way to do this would be to use splicing microarrays to compare splicing patterns in the presence and absence of endogenous CUGBP2 in the cerebral cortex. Target exons could be verified by RT-PCR and cloned into splicing reporters to determine the significance of the regulatory motifs by site-directed mutagenesis. These assays would provide insight into the network of exons regulated by CUGBP2, would help to further refine the splicing code, and could be used to generate a map to predict enhancement or silencing based on the position of target motifs.

Furthermore, structural studies could be used to completely understand the topology of binding by CUGBP2. In the past, such studies have been difficult because the inherent flexibility

of domains within the protein have made it difficult to crystallize. Additionally, attempts to crystallize individual RRMs and RRMs bound to an RNA ligand have been unsuccessful. Alternate approaches such as NMR or limited proteolysis may be more useful for structural analysis of this protein.

It would be of interest to dissect the mechanism of enhancement of the CI exon by CUGBP2. Mutational analysis of GU-rich motifs downstream from the CI exon in an *in* vivo splicing reporter system could be used to identify functionally significant motifs and *in vitro* splicing assays could be used to determine the effect of CUGBP2 on spliceosome assembly. It would also be interesting to follow up on the observation that the ΔDD mutant (containing RRMs 1, 2, and 3) has an enhancing effect on splicing of the NI exon while the RRM2_3 mutant (containing RRMs 2 and 3) has a silencing effect. Does RRM1 interact with proteins involved in splicing enhancement? Or do charge differences in the linker region between RRMs 2 and 3 of these two mutants account for this functional difference?

Finally, it would be interesting to determine how splicing patterns regulated by CUGBP2 are adjusted in response to cell depolarization. Is CUGBP2 modified in response to cell signaling or is this factor unresponsive to such events? Several potential phosphorylation sites are located in the divergent domain of CUGBP2. In the future, it would be interesting to determine whether these sites are used and how modification at these sites modulates protein activity.

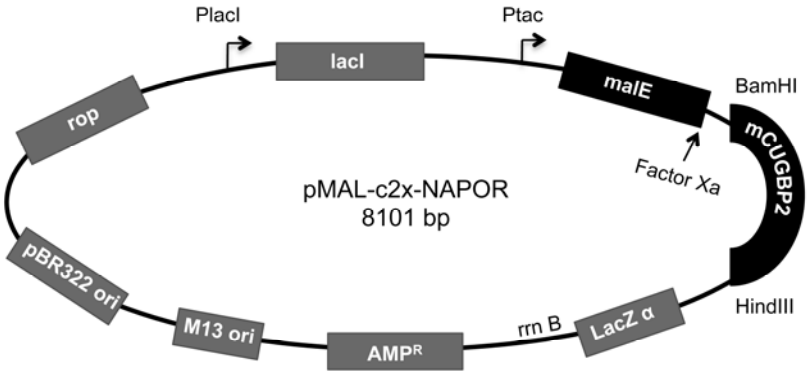
Taken together, I have provided insight into how splicing factors in general can recognize their cognate RNA substrates with high specificity by recognizing multiple contact sites on one pre-mRNA. Furthermore, I have characterized a novel mechanism of splicing silencing in great detail that is likely to be utilized by additional splicing silencers and dual functional regulators. This study also changes the way we think about splicing codes in that multiple recognition events and a functional reference point may be necessary to accurately predict splicing factor regulation of additional splicing factors. Moreover, cross-regulation and

autoregulation by splicing factors means that we need to be more careful about interpretation of results from overexpression and RNAi approaches in the future. Finally, the combination of approaches used in this study can be applied to characterize splicing codes and mechanisms of regulation by additional splicing factors that control alternative splicing events.

APPENDIX

PLASMID MAPS

Vectors for purification of recombinant proteins:



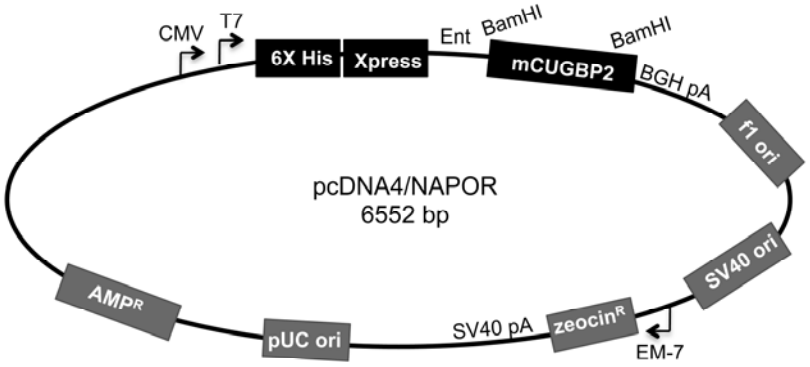
pMAL-c2x-NAPOR summary:

Nucleotide position:	Identity:
15-46	lacI promoter (PlacI)
81-1161	lacI open reading frame (ORF)
1406-1433	tac promoter (Ptac)
1528-2629	malE ORF – MBP tag
2629	Factor Xa cleavage site
2630-2702	BamHI restriction site
2702-4157	mouse (m) CUGBP2 ORF (amino acids 1-484)
4157	HindIII restriction site
4190-4369	lacZ α ORF
4369	start rrnB terminator
4873-5731	β-lactamase ORF – ampicillin resistance (AMP^R)
5775-6288	M13 origin of replication (ori)
6399-6987	pBR322 ori
7417-7606	rop ORF

CUGBP2 mutant vectors were generated from pMAL-c2x-NAPOR and contain CUGBP2 ORF fragments that code for the amino acids listed below.

Mutant:	Amino acids:
pMAL-c2x-RRM1_2	1-204
pMAL-c2x-RRM2_3	104-212, Xho1 (LE), 395-484
pMAL-c2x-ΔDD	1-204, Xba1 (SR), 392-484

Mammalian protein expression vectors:

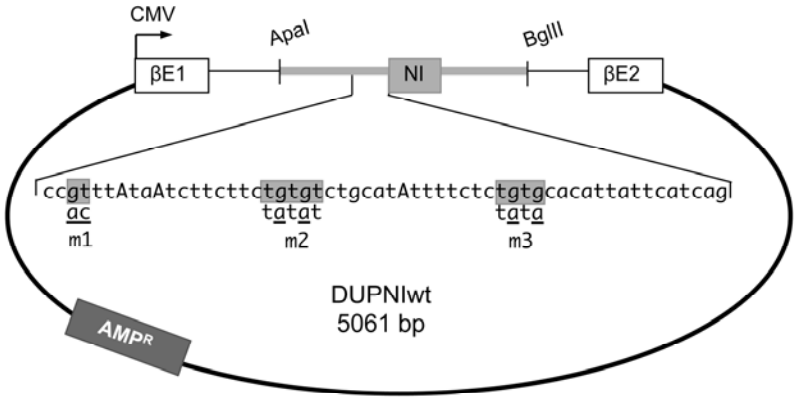


pcDNA4/NAPOR summary:

Nucleotide position:	Identity:
209-863	CMV promoter
863-882	T7 promoter
932-949	polyhistidine tag (6X His)
989-1012	Xpress epitope tag
998-1012	enterokinase recognition site (Ent)
1024	BamHI restriction site
1024-2479	mouse (m) CUGBP2 open reading frame (ORF)
2479	BamHI restriction site
2563-2790	BGH polyadenylation sequence (pA)
2836-3264	f1 origin of replication (ori)
3313-3599	SV40 ori
3647-3702	EM-7 promoter
3721-4095	zeocin resistance
4225-4355	SV40 pA
4738-5411	pUC ori
5556-6416	ampicillin resistance (AMP^R)

pcDNA4/PTB is essentially the same as pcDNA4/NAPOR except the mCUGBP2 ORF was replaced with the rat PTB ORF and the BamHI restriction sites were lost during cloning.

In vivo splicing reporters:

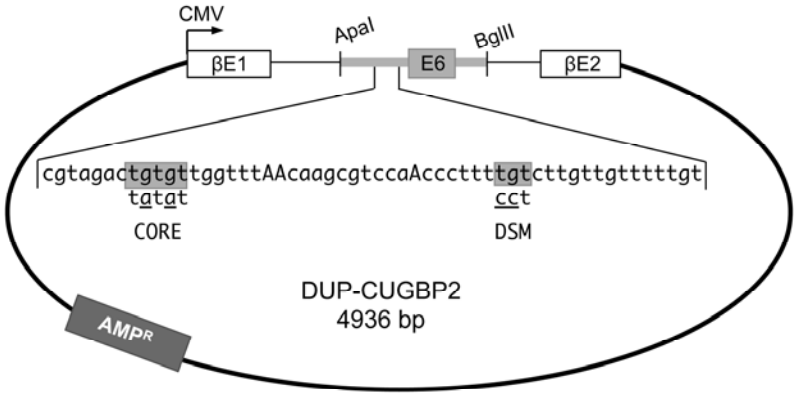


DUPNIwt summary:

Nucleotide position:	Identity:
unknown	CMV promoter
1-149	β-globin exon 1
150-235	β-globin intron 1 (5´ end)
236-241	ApaI restriction site
242-471	intron upstream from the NI exon
472-534	NI cassette exon of the GRIN1 gene
535-636	intron downstream from the NI exon
637-642	BglII restriction site
643-725	β-globin intron 1 (3´ end)
726-948	β-globin exon 2
unknown	ampicillin resistance (AMP^R)

Mutant plasmids were generated from DUPNIwt by site-directed mutagenesis and contained individual or combinations of mutations as indicated by underscores in the above sequence. The following mutants were made: DUPNIm1, DUPNIm2, DUPNIm3, DUPNIm1,2, DUPNIm2,3, DUPNIm1,3, DUPNIm1,2,3.

In vivo splicing reporters:

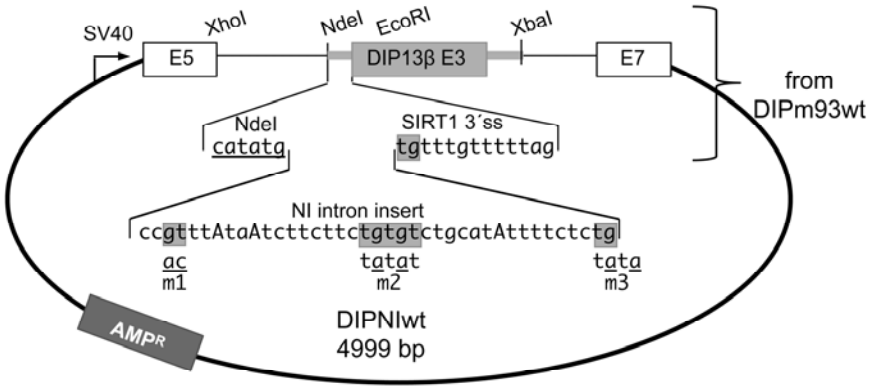


DUP-CUGBP2 summary:

Nucleotide position:	Identity:
unknown	CMV promoter
1-149	β-globin exon 1
150-235	β-globin intron 1 (5´ end)
236-241	ApaI restriction site
242-342	intron upstream from exon 6 (E6) containing E6
343-422	E6 of the CUGBP2 gene
423-511	intron downstream from E6
512-517	BglII restriction site
518-600	β-globin intron 1 (3´ end)
601-883	β-globin exon 2
unknown	ampicillin resistance (AMP^R)

Mutant plasmids were generated from DUP-CUGBP2 by site-directed mutagenesis and contained individual or combinations of mutations as indicated by underscores in the above sequence. The following mutants were made: DUP-CUGBP2-CORE mt, DUP-CUGBP2-DSM mt, DUP-CUGBP2-CORE/DSM mt.

In vivo splicing reporters:

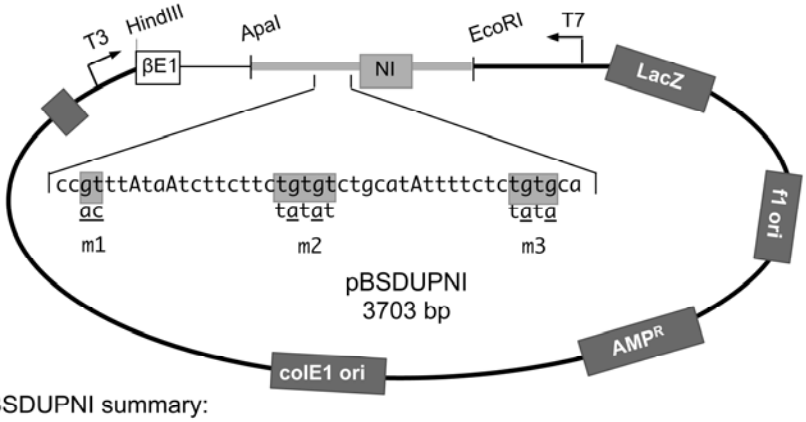


DIPNIwt summary:

Nucleotide position:	Identity:
unknown	SV40 promoter
1-308	SIRT1 exon 5
309-630	SIRT1 intron 5
315-320	XhoI restriction site
625-630	NdeI restriction site
631-669	39 nucleotide insert from NI branch site region
670-682	SIRT1 intron 5 3´ splice site (ss)
683-776	constitutive exon 3 of the DIP13β gene
721-726	EcoRI restriction site
777-1063	SIRT1 intron 6
787-792	XbaI restriction site
1064-1499	SIRT1 exon 7
unknown	ampicillin resistance (AMP^R)

Plasmids were generated from DIPm93wt by PCR amplification of the middle exon and flanking introns (shaded regions) with a 5´ primer containing a 5´ overhang of the wild type or mutant NI insert and the NdeI restriction site on the 5´ end. PCR fragments were reinserted into the DIPm93wt vector between the NdeI and XbaI restriction sites. Mutations are indicated by underscores in the above sequence. The following individual and combinations of mutations were made: DIPNIm1, DIPNIm2, DIPNIm3, DIPNIm1,2, DIPNIm2,3, DIPNIm1,3, DIPNIm1,2,3.

Plasmids for *in vitro* transcription:



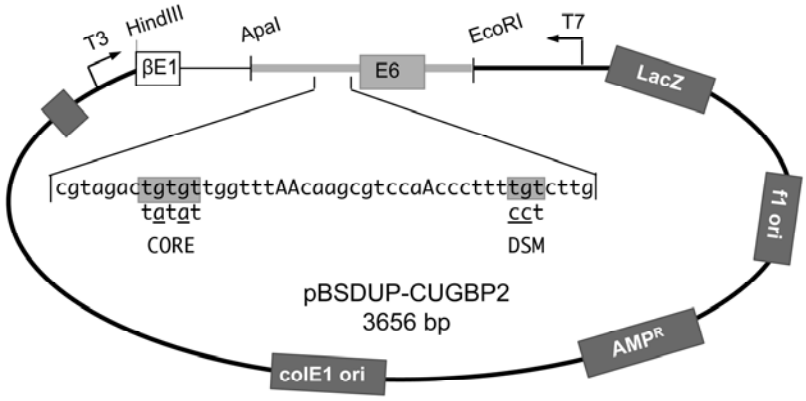
pBSDUPNI summary:

Nucleotide position:	Identity:
1-27	5′ end of the β-galactosidase gene (LacZ)
28-45	T3 promoter
46-50	region from pBS- polylinker
51-56	HindIII restriction site
57-146	3′ end of β-globin exon 1
147-238	β-globin intron 1 (5′ end)
239-244	ApaI restriction site
245-474	intron upstream from the NI exon
475-537	NI cassette exon of the GRIN1 gene
538-561	intron downstream from the NI exon
562-567	EcoRI restriction site
668-688	T7 promoter (in opposite direction)
689-841	3′ end of the the β-galactosidase gene (LacZ)
unknown	f1 origin of replication (ori)
unknown	ampicillin resistance (AMP^R)
unknown	colE1 ori

Mutant plasmids were generated from DUPNI mutants and contained individual or combinations of mutations as indicated by underscores in the above sequence. The following mutants were made: pBSDUPNIm1, pBSDUPNIm3, pBSDUPNIm1,2,3.

E5-8, E5-15, and E5-10 were similar to pBSDUPNI except the following GRIN1 gene regions were inserted between the HindIII and EcoRI restriction sites: E5-8, NI exon; E5-15, 78 nucleotides of the upstream intron; E5-10, 78 nucleotides of the upstream intron, the NI exon, and 24 nucleotides of the downstream intron.

Plasmids for *in vitro* transcription:



pBSDUP-CUGBP2 summary:

Nucleotide position:	Identity:
1-27	5´ end of the β-galactosidase gene (LacZ)
28-45	T3 promoter
46-50	region from pBS$^-$ polylinker
51-56	HindIII restriction site
57-146	3´ end of β-globin exon 1
147-238	β-globin intron 1 (5´ end)
239-244	ApaI restriction site
245-345	intron upstream from exon 6 (E6)
346-425	E6 of the CUGBP2 gene
426-514	intron downstream from E6
515-520	EcoRI restriction site
521-541	T7 promoter (in opposite direction)
542-694	3´ end of the the β-galactosidase gene (LacZ)
unknown	f1 origin of replication (ori)
unknown	ampicillin resistance (AMP^R)
unknown	colE1 ori

BIBLIOGRAPHY

Abovich N, Rosbash M (1997) Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. Cell 89:403-412.

Adema GJ, Baas PD (1991) Deregulation of alternative processing of Calcitonin/CGRP-I pre-mRNA by a single point mutation. Biochem Biophys Res Commun 178:985-992.

An P, Grabowski PJ (2007) Exon silencing by UAGG motifs in response to neuronal excitation. PLoS Biol 5:e36.

Anant S, Henderson JO, Mukhopadhyay D, Navaratnam N, Kennedy S, Min J, Davidson NO (2001) Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing. CUGBP2 modulates C to U editing of apolipoprotein B mRNA by interacting with apobec-1 and ACF, the apobec-1 complementation factor. J Biol Chem 276:47338-47351.

Ashiya M, Grabowski PJ (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 3:996-1015.

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

Auweter SD, Fasan R, Reymond L, Underwood JG, Black DL, Pitsch S, Allain FH (2006) Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. Embo J 25:163-173.

Baraniak AP, Chen JR, Garcia-Blanco MA (2006) Fox-2 mediates epithelial cell-specific fibroblast growth factor receptor 2 exon choice. Mol Cell Biol 26:1209-1222.

Bennett M, Michaud S, Kingston J, Reed R (1992) Protein components specifically associated with prespliceosome and spliceosome complexes. Genes Dev 6:1986-2000.

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

Berglund JA, Abovich N, Rosbash M (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. Genes Dev 12:858-867.

Berk AJ, Sharp PA (1978a) Structure of the adenovirus 2 early mRNAs. Cell 14:695-711.

Berk AJ, Sharp PA (1978b) Spliced early mRNAs of simian virus 40. Proc Natl Acad Sci U S A 75:1274-1278.

Bindereif A, Green MR (1987) An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. Embo J 6:2415-2424.

Black DL, Grabowski PJ (2003) Alternative pre-mRNA splicing and neuronal function. Prog Mol Subcell Biol 31:187-216.

Boutz PL, Chawla G, Stoilov P, Black DL (2007a) MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. Genes Dev 21:71-84.

Boutz PL, Stoilov P, Li Q, Lin CH, Chawla G, Ostrow K, Shiue L, Ares M, Jr., Black DL (2007b) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes Dev 21:1636-1652.

Bradley J, Carter SR, Rao VR, Wang J, Finkbeiner S (2006) Splice variants of the NR1 subunit differentially induce NMDA receptor-dependent gene expression. J Neurosci 26:1065-1076.

Brody E, Abelson J (1985) The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science 228:963-967.

Brunel C, Romby P (2000) Probing RNA structure and RNA-ligand complexes with chemical probes. Methods Enzymol 318:3-21.

Buckanovich RJ, Posner JB, Darnell RB (1993) Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. Neuron 11:657-672.

Burd CG, Swanson MS, Gorlach M, Dreyfuss G (1989) Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. Proc Natl Acad Sci U S A 86:9788-9792.

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31:3568-3571.

Cech TR, Zaug AJ, Grabowski PJ (1981) In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell 27:487-496.

Chang YF, Imam JS, Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem 76:51-74.

Charlet BN, Logan P, Singh G, Cooper TA (2002a) Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. Mol Cell 9:649-658.

Charlet BN, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA (2002b) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. Mol Cell 10:45-53.

Chiara MD, Gozani O, Bennett M, Champion-Arnaud P, Palandjian L, Reed R (1996) Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. Mol Cell Biol 16:3317-3326.

Choi DK, Ito T, Tsukahara F, Hirai M, Sakaki Y (1999) Developmentally-regulated expression of mNapor encoding an apoptosis-induced ELAV-type RNA binding protein. Gene 237:135-142.

Choi DK, Yoo KW, Hong SK, Rhee M, Sakaki Y, Kim CH (2003) Isolation and expression of Napor/CUG-BP2 in embryo development. Biochem Biophys Res Commun 305:448-454.

Choi YD, Grabowski PJ, Sharp PA, Dreyfuss G (1986) Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. Science 231:1534-1539.

Chow LT, Gelinas RE, Broker TR, Roberts RJ (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12:1-8.

Cull-Candy S, Brickley S, Farrant M (2001) NMDA receptor subunits: diversity, development and disease. Curr Opin Neurobiol 11:327-335.

Datta B, Weiner AM (1991) Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. Nature 352:821-824.

Dembowski JA, Grabowski PJ (2009) The CUGBP2 splicing factor regulates an ensemble of branchpoints from perimeter binding sites with implications for autoregulation. PLoS Genet 5:e1000595.

Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res 11:1475-1489.

Ding J, Hayashi MK, Zhang Y, Manche L, Krainer AR, Xu RM (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. Genes Dev 13:1102-1115.

Draper DE, Deckman IC, Vartikar JV (1988) Physical studies of ribosomal protein-RNA interactions. Methods Enzymol 164:203-220.

Dredge BK, Darnell RB (2003) Nova regulates GABA(A) receptor gamma2 alternative splicing via a distal downstream UCAU-rich intronic splicing enhancer. Mol Cell Biol 23:4687-4700.

Dredge BK, Stefani G, Engelhard CC, Darnell RB (2005) Nova autoregulation reveals dual functions in neuronal splicing. Embo J 24:1608-1620.

Ehlers MD, Tingley WG, Huganir RL (1995) Regulated subcellular distribution of the NR1 subunit of the NMDA receptor. Science 269:1734-1737.

Emeson RB, Hedjran F, Yeakley JM, Guise JW, Rosenfeld MG (1989) Alternative production of calcitonin and CGRP mRNA is regulated at the calcitonin-specific splice acceptor. Nature 341:76-80.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB (2002) Predictive identification of exonic splicing enhancers in human genes. Science 297:1007-1013.

Faustino NA, Cooper TA (2005) Identification of putative new splicing targets for ETR-3 using sequences identified by systematic evolution of ligands by exponential enrichment. Mol Cell Biol 25:879-887.

Fetzer S, Lauber J, Will CL, Luhrmann R (1997) The [U4/U6.U5] tri-snRNP-specific 27K protein is a novel SR protein that can be phosphorylated by the snRNP-associated protein kinase. RNA 3:344-355.

Fu XD (1995) The superfamily of arginine/serine-rich splicing factors. RNA 1:663-680.

Fu XD, Maniatis T (1992) The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. Proc Natl Acad Sci U S A 89:1725-1729.

Fukumura K, Kato A, Jin Y, Ideue T, Hirose T, Kataoka N, Fujiwara T, Sakamoto H, Inoue K (2007) Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1gamma gene. Nucleic Acids Res 35:5303-5311.

Gao K, Masuda A, Matsuura T, Ohno K (2008) Human branch point consensus sequence is yUnAy. Nucleic Acids Res 36:2257-2267.

Glanzer J, Miyashiro KY, Sul JY, Barrett L, Belt B, Haydon P, Eberwine J (2005) RNA splicing capability of live neuronal dendrites. Proc Natl Acad Sci U S A 102:16859-16864.

Goldstrohm AC, Greenleaf AL, Garcia-Blanco MA (2001) Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. Gene 277:31-47.

Goo YH, Cooper TA (2009) CUGBP2 directly interacts with U2 17S snRNP components and promotes U2 snRNA binding to cardiac troponin T pre-mRNA. Nucleic Acids Res.

Good PJ, Chen Q, Warner SJ, Herring DC (2000) A family of human RNA-binding proteins related to the Drosophila Bruno translational regulator. J Biol Chem 275:28583-28592.

Gozani O, Potashkin J, Reed R (1998) A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol Cell Biol 18:4752-4760.

Grabowski PJ, Sharp PA (1986) Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. Science 233:1294-1299.

Grabowski PJ, Zaug AJ, Cech TR (1981) The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of Tetrahymena. Cell 23:467-476.

Grabowski PJ, Seiler SR, Sharp PA (1985) A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell 42:345-353.

Grabowski PJ, Nasim FU, Kuo HC, Burch R (1991) Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. Mol Cell Biol 11:5919-5928.

Graveley BR (2000) Sorting out the complexity of SR protein functions. RNA 6:1197-1211.

Greer CL, Peebles CL, Gegenheimer P, Abelson J (1983) Mechanism of action of a yeast RNA ligase in tRNA splicing. Cell 32:537-546.

Gromak N, Matlin AJ, Cooper TA, Smith CW (2003) Antagonistic regulation of alpha-actinin alternative splicing by CELF proteins and polypyrimidine tract binding protein. RNA 9:443-456.

Hall SL, Padgett RA (1994) Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. J Mol Biol 239:357-365.

Hall SL, Padgett RA (1996) Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science 271:1716-1718.

Han J, Cooper TA (2005) Identification of CELF splicing activation and repression domains in vivo. Nucleic Acids Res 33:2769-2780.

Han K, Yeo G, An P, Burge CB, Grabowski PJ (2005) A combinatorial code for splicing silencing: UAGG and GGGG motifs. PLoS Biol 3:e158.

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

Hausner TP, Giglio LM, Weiner AM (1990) Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. Genes Dev 4:2146-2156.

He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5:522-531.

Helfman DM, Ricci WM (1989) Branch point selection in alternative splicing of tropomyosin pre-mRNAs. Nucleic Acids Res 17:5633-5650.

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

Hovhannisyan RH, Carstens RP (2005) A novel intronic cis element, ISE/ISS-3, regulates rat fibroblast growth factor receptor 2 splicing through activation of an upstream exon and repression of a downstream exon containing a noncanonical branch point sequence. Mol Cell Biol 25:250-263.

Hutton M (2001) Missense and splice site mutations in tau associated with FTDP-17: multiple pathogenic mechanisms. Neurology 56:S21-25.

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

Jackson IJ (1991) A reappraisal of non-consensus mRNA splice sites. Nucleic Acids Res 19:3795-3798.

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

Jamison SF, Crow A, Garcia-Blanco MA (1992) The spliceosome assembly pathway in mammalian extracts. Mol Cell Biol 12:4279-4287.

Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, Cooper TA (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. Proc Natl Acad Sci U S A 105:20333-20338.

Kohtz JD, Jamison SF, Will CL, Zuo P, Luhrmann R, Garcia-Blanco MA, Manley JL (1994) Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature 368:119-124.

Kol G, Lev-Maor G, Ast G (2005) Human-mouse comparative analysis reveals that branch-site plasticity contributes to splicing regulation. Hum Mol Genet 14:1559-1568.

Konarska MM, Sharp PA (1986) Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell 46:845-855.

Konarska MM, Sharp PA (1987) Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell 49:763-774.

Konarska MM, Grabowski PJ, Padgett RA, Sharp PA (1985) Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature 313:552-557.

Kralovicova J, Houngninou-Molango S, Kramer A, Vorechovsky I (2004) Branch site haplotypes that control alternative splicing. Hum Mol Genet 13:3189-3202.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31:147-157.

Kuo HC, Nasim FH, Grabowski PJ (1991) Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science 251:1045-1050.

Ladd AN, Cooper TA (2002) Finding signals that regulate alternative splicing in the post-genomic era. Genome Biol 3:reviews0008.

Ladd AN, Cooper TA (2004) Multiple domains control the subcellular localization and activity of ETR-3, a regulator of nuclear and cytoplasmic RNA processing events. J Cell Sci 117:3519-3529.

Ladd AN, Charlet N, Cooper TA (2001) The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. Mol Cell Biol 21:1285-1296.

Ladd AN, Stenberg MG, Swanson MS, Cooper TA (2005) Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development. Dev Dyn 233:783-793.

Lamond AI, Konarska MM, Grabowski PJ, Sharp PA (1988) Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proc Natl Acad Sci U S A 85:411-415.

Langford CJ, Gallwitz D (1983) Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33:519-527.

Lee JA, Xing Y, Nguyen D, Xie J, Lee CJ, Black DL (2007) Depolarization and CaM kinase IV modulate NMDA receptor splicing through two essential RNA elements. PLoS Biol 5:e40.

Leroy O, Dhaenens CM, Schraen-Maschke S, Belarbi K, Delacourte A, Andreadis A, Sablonniere B, Buee L, Sergeant N, Caillet-Boudin ML (2006) ETR-3 represses Tau exons 2/3 inclusion, a splicing event abnormally enhanced in myotonic dystrophy type I. J Neurosci Res 84:852-859.

Levers TE, Tait S, Birling MC, Brophy PJ, Price DJ (2002) Etr-r3/mNapor, encoding an ELAV-type RNA binding protein, is expressed in differentiating cells in the developing rodent forebrain. Mech Dev 112:191-193.

Li D, Bachinski LL, Roberts R (2001) Genomic organization and isoform-specific tissue expression of human NAPOR (CUGBP2) as a candidate gene for familial arrhythmogenic right ventricular dysplasia. Genomics 74:396-401.

Lim LP, Burge CB (2001) A computational analysis of sequence features involved in recognition of short introns. Proc Natl Acad Sci U S A 98:11193-11198.

Lipscombe D (2005) Neuronal proteins custom designed by alternative splicing. Curr Opin Neurobiol 15:358-363.

Liu Z, Luyten I, Bottomley MJ, Messias AC, Houngninou-Molango S, Sprangers R, Zanier K, Kramer A, Sattler M (2001) Structural basis for recognition of the intron branch site RNA by splicing factor 1. Science 294:1098-1102.

Long JC, Caceres JF (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochem J 417:15-27.

Long M, de Souza SJ, Gilbert W (1997) The yeast splice site revisited: new exon consensus from genomic analysis. Cell 91:739-740.

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

MacMillan AM, Query CC, Allerson CR, Chen S, Verdine GL, Sharp PA (1994) Dynamic association of proteins with the pre-mRNA branch region. Genes Dev 8:3008-3020.

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

Makeyev EV, Zhang J, Carrasco MA, Maniatis T (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell 27:435-448.

Maniatis T, Tasic B (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 418:236-243.

Maris C, Dominguez C, Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. Febs J 272:2118-2131.

Marquis J, Paillard L, Audic Y, Cosson B, Danos O, Le Bec C, Osborne HB (2006) CUG-BP1/CELF1 requires UGU-rich sequences for high-affinity binding. Biochem J 400:291-301.

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

Modafferi EF, Black DL (1997) A complex intronic splicing enhancer from the c-src pre-mRNA activates inclusion of a heterologous exon. Mol Cell Biol 17:6537-6545.

Mount SM (1982) A catalogue of splice junction sequences. Nucleic Acids Res 10:459-472.

Mu Y, Otsuka T, Horton AC, Scott DB, Ehlers MD (2003) Activity-dependent mRNA splicing controls ER export and synaptic delivery of NMDA receptors. Neuron 40:581-594.

Mukhopadhyay D, Houchen CW, Kennedy S, Dieckgraefe BK, Anant S (2003a) Coupled mRNA stabilization and translational silencing of cyclooxygenase-2 by a novel RNA binding protein, CUGBP2. Mol Cell 11:113-126.

Mukhopadhyay D, Jung J, Murmu N, Houchen CW, Dieckgraefe BK, Anant S (2003b) CUGBP2 plays a critical role in apoptosis of breast cancer cells in response to genotoxic injury. Ann N Y Acad Sci 1010:504-509.

Nakahata S, Kawamoto S (2005) Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. Nucleic Acids Res 33:2078-2089.

Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M, Jr. (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21:708-718.

Novoyatleva T, Tang Y, Rafalska I, Stamm S (2006) Pre-mRNA missplicing as a cause of human disease. Prog Mol Subcell Biol 44:27-46.

O'Neill JP, Rogan PK, Cariello N, Nicklas JA (1998) Mutations that alter RNA splicing of the human HPRT gene: a review of the spectrum. Mutat Res 411:179-214.

Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL, Allain FH (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science 309:2054-2057.

Okabe S, Miwa A, Okado H (1999) Alternative splicing of the C-terminal domain regulates cell surface expression of the NMDA receptor NR1 subunit. J Neurosci 19:7781-7792.

Padgett RA, Hardy SF, Sharp PA (1983) Splicing of adenovirus RNA in a cell-free transcription system. Proc Natl Acad Sci U S A 80:5230-5234.

Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA (1984) Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. Science 225:898-903.

Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA (1986) Splicing of messenger RNA precursors. Annu Rev Biochem 55:1119-1150.

Patel AA, McCarthy M, Steitz JA (2002) The splicing of U12-type introns can be a rate-limiting step in gene expression. Embo J 21:3804-3815.

Paul S, Dansithong W, Kim D, Rossi J, Webster NJ, Comai L, Reddy S (2006) Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. Embo J 25:4271-4283.

Peebles CL, Gegenheimer P, Abelson J (1983) Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease. Cell 32:525-536.

Philips AV, Timchenko LT, Cooper TA (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. Science 280:737-741.

Query CC, Moore MJ, Sharp PA (1994) Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes Dev 8:587-597.

Reed R (1990) Protein composition of mammalian spliceosomes assembled in vitro. Proc Natl Acad Sci U S A 87:8031-8035.

Reed R (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. Curr Opin Genet Dev 6:215-220.

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

Robberson BL, Cote GJ, Berget SM (1990) Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol 10:84-94.

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

Roscigno RF, Garcia-Blanco MA (1995) SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA 1:692-706.

Ruskin B, Green MR (1985) An RNA processing activity that debranches RNA lariats. Science 229:135-140.

Ruskin B, Krainer AR, Maniatis T, Green MR (1984) Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38:317-331.

Savkur RS, Philips AV, Cooper TA (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. Nat Genet 29:40-47.

Sawa H, Shimura Y (1992) Association of U6 snRNA with the 5'-splice site region of pre-mRNA in the spliceosome. Genes Dev 6:244-254.

Schmucker D, Clemens JC, Shu H, Worby CA, Xiao J, Muda M, Dixon JE, Zipursky SL (2000) Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell 101:671-684.

Sharma S, Falick AM, Black DL (2005) Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. Mol Cell 19:485-496.

Sharp PA, Burge CB (1997) Classification of introns: U2-type or U12-type. Cell 91:875-879.

Shen H, Green MR (2004) A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. Mol Cell 16:363-373.

Shen H, Kan JL, Green MR (2004) Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol Cell 13:367-376.

Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R (2006) Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res 34:3955-3967.

Sickmier EA, Frato KE, Shen H, Paranawithana SR, Green MR, Kielkopf CL (2006) Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. Mol Cell 23:49-59.

Singh G, Charlet BN, Han J, Cooper TA (2004) ETR-3 and CELF4 protein domains required for RNA binding and splicing activity in vivo. Nucleic Acids Res 32:1232-1241.

Singh R, Valcarcel J, Green MR (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science 268:1173-1176.

Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum Mol Genet 15:2490-2508.

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

Spellman R, Smith CW (2006) Novel modes of splicing repression by PTB. Trends Biochem Sci 31:73-76.

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

Staknis D, Reed R (1994b) Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. Mol Cell Biol 14:2994-3005.

Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 92:315-326.

Staley JP, Woolford JL, Jr. (2009) Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr Opin Cell Biol 21:109-118.

Stamm S (2007) Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem.

Sun H, Chasin LA (2000) Multiple splicing defects in an intronic false exon. Mol Cell Biol 20:6414-6425.

Sun JS, Manley JL (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev 9:843-854.

Suzuki H, Jin Y, Otani H, Yasuda K, Inoue K (2002) Regulation of alternative splicing of alpha-actinin transcript by Bruno-like proteins. Genes Cells 7:133-141.

Takahashi N, Sasagawa N, Suzuki K, Ishiura S (2000) The CUG-binding protein binds specifically to UG dinucleotide repeats in a yeast three-hybrid system. Biochem Biophys Res Commun 277:518-523.

Tarn WY, Steitz JA (1995) Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. Proc Natl Acad Sci U S A 92:2504-2508.

Tarn WY, Steitz JA (1996a) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science 273:1824-1832.

Tarn WY, Steitz JA (1996b) A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. Cell 84:801-811.

Timchenko LT, Miller JW, Timchenko NA, DeVore DR, Datar KV, Lin L, Roberts R, Caskey CT, Swanson MS (1996) Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. Nucleic Acids Res 24:4407-4414.

Timchenko NA, Iakova P, Cai ZJ, Smith JR, Timchenko LT (2001a) Molecular basis for impaired muscle differentiation in myotonic dystrophy. Mol Cell Biol 21:6927-6938.

Timchenko NA, Cai ZJ, Welm AL, Reddy S, Ashizawa T, Timchenko LT (2001b) RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. J Biol Chem 276:7820-7826.

Timchenko NA, Patel R, Iakova P, Cai ZJ, Quan L, Timchenko LT (2004) Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. J Biol Chem 279:13129-13139.

Treisman R, Orkin SH, Maniatis T (1983) Specific transcription and RNA splicing defects in five cloned beta-thalassaemia genes. Nature 302:591-596.

Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB (2003) CLIP identifies Nova-regulated RNA networks in the brain. Science 302:1212-1215.

Ule J, Stefani G, Mele A, Ruggiu M, Wang X, Taneri B, Gaasterland T, Blencowe BJ, Darnell RB (2006) An RNA map predicting Nova-dependent splicing regulation. Nature 444:580-586.

Ule J, Ule A, Spencer J, Williams A, Hu JS, Cline M, Wang H, Clark T, Fraser C, Ruggiu M, Zeeberg BR, Kane D, Weinstein JN, Blume J, Darnell RB (2005) Nova regulates brain-specific splicing to shape the synapse. Nat Genet 37:844-852.

Underwood JG, Boutz PL, Dougherty JD, Stoilov P, Black DL (2005) Homologues of the Caenorhabditis elegans Fox-1 protein are neuronal splicing regulators in mammals. Mol Cell Biol 25:10005-10016.

Utans U, Kramer A (1990) Splicing factor SF4 is dispensable for the assembly of a functional splicing complex and participates in the subsequent steps of the splicing reaction. Embo J 9:4119-4126.

Valadkhan S, Manley JL (2001) Splicing-related catalysis by protein-free snRNAs. Nature 413:701-707.

Valadkhan S, Mohammadi A, Wachtel C, Manley JL (2007) Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. RNA 13:2300-2311.

Valcarcel J, Gaur RK, Singh R, Green MR (1996) Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science 273:1706-1709.

Wachtel C, Manley JL (2009) Splicing of mRNA precursors: the role of RNAs and proteins in catalysis. Mol Biosyst 5:311-316.

Wagner EJ, Garcia-Blanco MA (2001) Polypyrimidine tract binding protein antagonizes exon definition. Mol Cell Biol 21:3281-3288.

Wang GS, Kearney DL, De Biasi M, Taffet G, Cooper TA (2007) Elevation of RNA-binding protein CUGBP1 is an early event in an inducible heart-specific mouse model of myotonic dystrophy. J Clin Invest 117:2802-2811.

Wang Z, Burge CB (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. RNA 14:802-813.

Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB (2004) Systematic identification and analysis of exonic splicing silencers. Cell 119:831-845.

Wassarman DA, Steitz JA (1992) Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. Science 257:1918-1925.

Will CL, Luhrmann R (1997) Protein functions in pre-mRNA splicing. Curr Opin Cell Biol 9:320-328.

Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CW (2004) Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell 13:91-100.

Wu JA, Manley JL (1991) Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. Nature 352:818-821.

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

Wyatt JR, Sontheimer EJ, Steitz JA (1992) Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes Dev 6:2542-2553.

Xie J, Jan C, Stoilov P, Park J, Black DL (2005) A consensus CaMK IV-responsive RNA sequence mediates regulation of alternative exons in neurons. RNA 11:1825-1834.

Yeo GW, Van Nostrand EL, Liang TY (2007) Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. PLoS Genet 3:e85.

Zamore PD, Patton JG, Green MR (1992) Cloning and domain structure of the mammalian splicing factor U2AF. Nature 355:609-614.

Zandberg H, Moen TC, Baas PD (1995) Cooperation of 5' and 3' processing sites as well as intron and exon sequences in calcitonin exon recognition. Nucleic Acids Res 23:248-255.

Zhang L, Liu W, Grabowski PJ (1999) Coordinate repression of a trio of neuron-specific splicing events by the splicing regulator PTB. RNA 5:117-130.

Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR (1992) Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. Proc Natl Acad Sci U S A 89:8769-8773.

Zhang W, Liu H, Han K, Grabowski PJ (2002) Region-specific alternative splicing in the nervous system: implications for regulation by the RNA-binding protein NAPOR. RNA 8:671-685.

Zhuang Y, Weiner AM (1986) A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell 46:827-835.

Zhuang YA, Goldstein AM, Weiner AM (1989) UACUAAC is the preferred branch site for mammalian mRNA splicing. Proc Natl Acad Sci U S A 86:2752-2756.

Zukin RS, Bennett MV (1995) Alternatively spliced isoforms of the NMDARI receptor subunit. Trends Neurosci 18:306-313.