An Intronic Downstream Enhancer Promotes 3′ Splice Site Usage of a Neural Cell-specific Exon

Neng-hua Guo and Sachiyo Kawamoto

From the Laboratory of Molecular Cardiology, NHLBI, National Institutes of Health, Bethesda, Maryland 20892

The human nonmuscle myosin heavy chain B gene contains a 30-nucleotide alternative exon, N30, that is included in the mRNA from neural cells but is skipped in all other cells. We have previously identified an intronic distal downstream enhancer (IDDE) region that is required for neural cell-specific inclusion of N30. In this study, we investigated the mechanism by which the IDDE promotes N30 exon usage. In vitro splicing analysis using neural cell nuclear extracts and two-exon pre-mRNA substrates, which consist of the N30 exon and either the upstream (E5) or downstream (E6) exon, demonstrates that the IDDE activates upstream E5-N30 splicing by facilitating early prespliceosome complex formation. The IDDE has no effect on N30-E6 splicing where the IDDE resides. Inspection of splice site consensus sequences shows that a polypyrimidine (Py) tract preceding N30 is suboptimal for U2AF binding. Optimizing the Py tract completely relieves the requirement for the IDDE in E5-N30 splicing in vitro. In transfected cells, the wild-type minigene transcripts, which consist of three exons, E5, N30, and E6, undergo neural cell-specific and IDDE-dependent alternative splicing of N30. Optimizing the Py tract in minigenes also completely relieves the requirement for the IDDE in N30 inclusion. Furthermore, overexpression of the truncated U2AF65, which contains the arginine and serine dipeptide-rich domain and linker domain, but lacks the RNA binding domain, selectively inhibits the IDDE-mediated N30 inclusion in mRNA from the wild-type minigene in a dominant negative fashion. These results support the hypothesis that the IDDE facilitates the recognition of the 3′ splice site preceding N30 by a network of protein-protein interactions implicated in the recruitment of U2AF to a suboptimal Py tract.

Alternative splicing of pre-mRNA is a fundamental mechanism for regulating eukaryotic gene expression. In many cases, alternative RNA splicing contributes to developmentally regulated and cell type-specific patterns of gene expression. Although a great deal of information is available concerning the general constitutive splicing reactions of simple splicing units, the molecular basis for alternative splice site selection is not well understood (for review see Refs. 1–3). Considerable insight into the regulation of alternative splicing has been gained from studies of genes identified in Drosophila. The most extensively characterized of these involve Drosophila sex determination, for which examples of both positive and negative regulation of splicing were found (for review see Refs. 3 and 4). In vertebrates, however, much less is known about the mechanisms and cellular factors involved in regulated alternative splicing.

Studies of vertebrate genes aimed at understanding the regulation of alternative splicing have led to identification of a number of pre-mRNA features that influence alternative splice site selection. These include the sequence of the 5′ and 3′ splice sites, branch point sequence and location, exon size, intron size, and specific RNA sequences (enhancers or repressors) located in exons or introns. Many parameters can affect splicing in ways that are complex and not readily predictable.

cis-Acting RNA sequences can enhance or repress utilization of alternative splice sites. Elements that promote splicing of an adjacent splice site are known collectively as splicing enhancers and are classified by their location as exonic or intronic enhancers. Exonic enhancers are often rich in purine residues (5–7). These exonic splicing enhancers bind to members of the SR protein family, which contain a characteristic arginine and serine dipeptide repeat (RS) domain and RNA binding domains (for review see Refs. 8 and 9). For exonic enhancers, which can recruit SR proteins, there is evidence that the enhancer-protein complex functions by stimulating splicesosome assembly at the upstream 3′ splice site. A large number of exons have this enhancer element. Only a few of them, however, appear to be a target for cell type-specific regulation. In addition to exonic enhancers, a number of splicing events are controlled by intronic splicing enhancers (10–19). These sequences have been found in the introns downstream of small and/or cell type-specific exons, and their presence is required for the splicing of these exons. However, with a few exceptions (16, 20–23), the proteins that mediate enhancer effects remain unidentified. Moreover, how these enhancers affect splicesosome assembly is not known.

We have been using the human nonmuscle myosin heavy chain (NMHC)-B gene as a model system to study the regulatory mechanisms responsible for neural cell-specific alternative splicing of pre-mRNA. The NMHC-B gene encodes a polypeptide of approximately 200 kDa. A dimer of this gene product, together with two pairs of light chains (20 and 17 kDa), constitutes a myosin molecule that demonstrates force-generating ATPase activity when it interacts with actin. The NMHC-B gene is expressed ubiquitously in most cell types (24); however, in neural cells, specific forms of NMHC-B are generated by
cassette-type alternative splicing. NMHC-B has been shown to play an important role in neural cell migration and adhesion in NMHC-B-depleted mice (25). Alternative splicing occurs at two different locations in the pre-mRNA (26). One alternative exon (N30), consisting of a 30-nt coding sequence, is located between the constitutive exons E5 and E6. The 10 amino acids encoded by the N30 exon is located near the ATP-binding region of the molecule and includes a serine residue that has been shown to be phosphorylated by mitogen-activated protein kinase and brain-specific cyclin-dependent protein kinase 5 (27), suggesting a specific role of this isoform in a signal transduction pathway. The other alternative exon, consisting of a 63-nt coding sequence, is located between constitutive exons E15 and E16, which is near the actin-binding region. Inclusion of these two alternative exons is restricted to neural cells in mammals and birds. The two alternative splicing events, however, appear to be regulated differentially by agonist stimulation in cultured neural cells as well as during brain development (28).

We have focused on N30 regulation, since some culture systems maintain regulation of N30 splicing (14, 28). Previously, we demonstrated that a minigene encompassing E5-N30-E6 produced cell type-specific and differentiation state-dependent regulation of N30 inclusion, using a transient minigene transfection system (14). In addition, the previous study established that neural cell-specific N30 inclusion requires a cis-acting intronic enhancer sequence that is located 1.5 kb downstream of exon N30. This intronic enhancer region, 142 nucleotides (nts) in length, is now designated as an intronic distal downstream enhancer (IDDE) region (corresponding to fragment d in Ref. 14). In the present study, we attempted to find out how the IDDE promotes N30 exon usage during the splicing reaction using in vitro splicing and minigene transfection systems. The experimental data shown here support the hypothesis that the IDDE facilitates recognition of the 3′ splice site preceding the N30 exon by a network of protein–protein interactions implicated in the recruitment of U2AF to a suboptimal polyypuridine (Py) tract.

**EXPERIMENTAL PROCEDURES**

**Construction of Plasmids—** For in vitro splicing substrates, defined portions of the genomic DNA (see Fig. 1C) were amplified and connected by recombinant PCR using the appropriate synthetic primers and were introduced into pBluescriptIIISK (+) (Stratagene), which contains the T7 promoter. Unique restriction enzyme sites were engineered at the junctions between the different portions of the genomic DNA. Construction of a wild-type minigene (Fig. 1B) has been described previously (corresponding to minigene C in Ref. 14). The constructs for in vitro splicing substrates and minigenes bearing the mutations m1 and m2 were generated by recombinant PCR.

For U2AF65 expression, a plasmid pC3S+ MT (29), which contains the cytomegalovirus promoter/enhancer, Myc epitope tags, and SV40 polyA signal, was used as a host vector. This vector was a gift from Dr. Yongsook Kim (NHLBI, National Institutes of Health). The nuclear localization signal of the SV40 large T antigen was introduced following the correct amber codon at the amino terminus. All of the constructs were verified by DNA sequencing.

**Nuclear Extract Preparation—** Nuclear extracts from Y79 cells were prepared according to protocols previously described (30). Cells are washed in phosphate-buffered saline and pelleted by centrifugation for 10 min at 1500 × g. The packed cell volume of the hypotonic buffer (10 mM HEPES, pH 7.9 at 4 °C, 10 mM KCl, 1.5 mM MgCl2, 0.4 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 mM dithiothreitol (DTT)) and are centrifuged for 5 min at 1850 × g. The cells are resuspended in the hypotonic buffer to a final volume of 3× the original packed cell volume and allowed to swell for 10–15 min on ice. The swollen cells are homogenized with 10–15 strokes in a glass Dounce homogenizer to obtain >70% cell lysis and the nuclei are centrifuged for 15 min at 3300 × g. The nuclear pellets are resuspended in the low salt buffer (20 mM HEPES, pH 7.9, 2 mM KCl, 1.5 mM MgCl2, 0.2 mM EDTA, 25% glycerol, 0.4 mM PMSF, and 0.5 mM DTT) with a volume equal to 0.33–0.5× the packed nuclear volume. High salt buffer (0.33–0.5× the original packed cell volume) consisting of 20 mM HEPES, pH 7.9, 1.2 mM KCl, 1.5 mM MgCl2, 0.2 mM EDTA, 25% glycerol, 0.4 mM PMSF, and 0.5 mM DTT, is added to the suspended nuclei, and then the entire suspension is homogenized with 3–5 strokes in the Dounce homogenizer. Following extraction for 30 min with stirring, the nuclear suspension is centrifuged for 30 min at 3300 × g. The resulting supernatant is dialyzed against a buffer consisting of 20 mM HEPES, pH 7.9, 100 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.4 mM PMSF, and 0.5 mM DTT for 2–4 h. The insoluble materials are removed by centrifugation. The aliquoted nuclear extracts are quickly frozen in liquid N2 and stored at −70 °C. All procedures were performed at 0–4 °C. Protein concentrations were determined by the Bradford method (Bio-Rad) using bovine serum albumin as a standard. Typically, protein concentrations of 7–10 mg/ml were obtained.

**In Vitro Splicing and Analysis of Splicingosome Assembly—** Capped splicing RNA substrates, uniformly labeled with [α-32P]UTP (800 Ci/mmol, Amersham Pharmacia Biotech), were synthesized with T7 RNA polymerase (Stratagene) from the template plasmids linearized by the appropriate restriction enzymes. In vitro splicing reactions performed in 15-μl volumes containing 5 μl of nuclear extracts, 1 μl of a 15× concentrated splicing mix (4.5 mM MgCl2, 22.5 mM ATP, 75 mM creatine phosphate, and 75 mM dithiothreitol), 40 units of RNasin (Promega), and 2 ng of capped and labeled pre-mRNA substrates. Reaction mixtures were incubated at 30 °C for 0.5–4 h. The reaction was stopped by addition of proteinase K, followed by phenol/chloroform extraction and ethanol precipitation. The resulting RNAs were electro- phoresed in denaturing urea-polyacrylamide (4.5, 6, and 10%) gels and autoradiographed. The branched reaction was performed as described previously (31).

For analysis of spliceosome complexes, the splicing reaction was performed under the same conditions as described above. The reaction was started by addition of the labeled pre-mRNA and incubated at 30 °C for the indicated times. The reaction was stopped by addition of heparin (final concentration 5 mg/ml) and placed on ice for 10 min prior to electrophoresis. Samples for zero time incubation were assembled by first adding heparin to the reaction mixture without the pre-mRNA, and then the pre-mRNAs were added on ice. The reaction mixture was electrophoresed in a native polyacrylamide gel (4% acrylamide with 0.5% bisacrylamide) using a running buffer containing 25 mM Tris and 192 mM glycine at 120 V/10 cm for 2–5 h. The dried gels were autoradiographed.

**Determination of Branch Sites—** The spliced RNA products were reverse-transcribed by Superscript II RNase H− transcribe transcriptase (Life Technologies, Inc.) using random primers. The resulting cDNAs were amplified by AmpliTag (PerkinElmer Life Sciences) using primers 5′-GTGTTTGAAGATGATTGTC-3′ and 5′-GGTGGTTTGGGA-3′, followed by the second amplification using primers 5′-GTTGTTGAGATTACGGTCTCC-3′ and 5′-TGGAAAATGGCT- TACTGCTAGGCA-3′. The products were subcloned in pGEM-T (Promega). DNA sequencing was performed using dRhodamine termination cycle sequencing reagents (PE Applied Biosystems).

**Minigene Transfection and Analysis of mRNA—** The minigenes were transfected into Y79 cells using Lipofectin and into HeLa cells by the calcium-phosphate coprecipitation method as described previously (14). For cotransfection of a minigene with the U2AF65 expression constructs, indicated amounts of the U2AF65 construct and the empty host vector (total 1.25 μg) and the wild-type minigene containing IDDE (1 μg) were cotransfected into 2×106 Y79 cells with 7–9 μl of Lipofectin (Life Technologies, Inc.). After transfection, Y79 cells were replated on poly-L-lysine-coated plates and maintained in 2 mM sodium butyrate containing medium for 4–6 days before being harvested. Preparation of total RNA from transfected cells, analysis of mRNAs derived from the minigenes by RT-PCR, and quantitation of RT-PCR products by using 5′ end 32P-labeled primers was performed as described previously (14).

**Immunoblot—** Experiments that required both protein and mRNA analysis were done using immunoblot analysis. Minigenes were cotransfected into Y79 cells using Lipofectin and into HeLa cells by the calcium-phosphate coprecipitation method as described previously (14). To cotransfect a minigene with the U2AF65 expression constructs, indicated amounts of the U2AF65 construct and the empty host vector (total 1.25 μg) and the wild-type minigene containing IDDE (1 μg) were cotransfected into 2×106 Y79 cells with 7–9 μl of Lipofectin (Life Technologies, Inc.). After transfection, Y79 cells were replated on poly-L-lysine-coated plates and maintained in 2 mM sodium butyrate containing medium for 4–6 days before being harvested. Preparation of total RNA from transfected cells, analysis of mRNAs derived from the minigenes by RT-PCR, and quantitation of RT-PCR products by using 5′ end 32P-labeled primers was performed as described previously (14).
In a previous study, we have defined an intronic enhancer region, IDDE, which is required for neural cell-specific alternative splicing of the cassette-type exon, N30, in human NMHC-B pre-mRNA (see Fig. 1). In the presence of the IDDE, the mRNA from a minigenic transfected into human neural retinoblastoma Y79 cells can include the N30 exon, similar to the endogenous NMHC-B mRNA, whereas in the absence of the IDDE, the minigenic mRNA excludes exon N30. Exon N30 is skipped in the mRNAs from minigenes both with and without the IDDE in transfected non-neural cell lines, such as HeLa and NIH3T3 cells (14) (Fig. 6). Therefore, the IDDE confers neural cell specificity on N30 inclusion. To investigate how the IDDE enhances N30 recognition and promotes removal of the upstream and downstream introns, we first established an in vitro splicing assay system for NMHC-B pre-mRNA using Y79 nuclear extracts. We chose Y79 cells as a source of nuclear extracts since endogenous Y79 mRNA in-fuences the time course of their appearances, their electrophoretic mobility in different polyacrylamide concentration gels, and the effect of treating the products with S100 fractions containing debranching activity (data not shown). Relative amounts of each of the four products from the pre-mRNA containing the IDDE are roughly equal to those from pre-mRNA lacking the IDDE, indicating that there is no effect of the IDDE on the splicing efficiency of N30-E6. In contrast, pre-mRNA consisting of the E5 and N30 exons, in the absence of the IDDE, is poorly spliced. Neither the expected intermediates nor the final products are detectable (lane 2 in Fig. 2). However, including the IDDE downstream of N30 in the pre-mRNA causes marked activation of E5-N30 splicing. All four intermediate and final products are readily observed (lanes 4 and 6 in Fig. 2). Thus, the IDDE promotes removal of the intron between E5 and N30. We also analyzed pre-mRNA consisting of the two constitutive exons E5 and E6. In the absence of the IDDE, this pre-mRNA substrate is efficiently spliced (lane 10 in Fig. 2). The 5′ splice donor sequence following E5 is readily used in the context of the E5-E6 pre-mRNA. Since all these splicing reactions are carried out using two-exon pre-mRNA, without a competing exon, poor efficiency of the E5-N30 splicing in the absence of the IDDE is likely due to the intrinsically poor recognition signals of the 3′ splice site preceding the N30 exon. Thus, the IDDE appears to promote the recognition of the 3′ splice signal of the N30 exon.

The IDDE Facilitates Early Prespliceosome Complex Formation—Pre-mRNA splicing is a multistage, dynamic biochemical reaction taking place within the spliceosome. It has been well documented that assembly of the spliceosome proceeds through a highly ordered pathway of complex formation, involving 5 snRNAs and over 50 distinct protein components. A series of complexes are distinguished at different stages of spliceosome assembly in the order E → A → B → C (32).

---

* S. Kawamoto, unpublished observations.

---

![Diagram of the human NMHC-B gene surrounding alternative exon N30 and the constructs used in this study. A. native gene. Rectangles and horizontal lines in the diagram indicate exons and introns, respectively. E5 and E6 are constitutive exons, and N30 and R18 are alternative exons. Inclusion of R18 in the mRNA is always associated with N30 inclusion. However, N30 inclusion occurs without R18 inclusion, and inclusion of N30 alone is the dominant form of neural mRNA. Inclusion of both N30 and R18 can be found in a small degree, limited number of neural cell lines but not in neural tissues of animals. The IDDE is located between nts 1544 and 1685 downstream of N30 in the native gene. Exon size and the IDDE are not drawn to scale. B. minigene. Minigenes are constructed by inserting the human NMHC-B genomic DNA fragments indicated into the intron between exons E2 and E3 of the rat preproinsulin gene (PPI). Transcription of the minigene is driven by the Rous sarcoma virus long terminal repeat (RSLLTR). Arrows above E5 and PPIE3 indicate the location of the primers used for RT-PCR. C. in vitro splicing substrates. A blank space between the horizontal lines in E5-N30 and E5-E6 indicates a deletion.**
To determine at what stage during the splicing reaction the IDDE functions, we analyzed spliceosome assembly using the E5-N30 pre-mRNAs with and without the IDDE. The splicing reaction mixture at different time points was resolved by native polyacrylamide gel electrophoresis. Both pre-mRNAs form large quantities of complex H in which hnRNP proteins bind nonspecifically to the pre-mRNAs (Fig. 3). Complex H is not considered a functional intermediate in the spliceosome assembly pathway. The initial prespliceosome complex E, in which U1snRNP and U2AF are recruited to the 5′ and 3′ splice sites, respectively, cannot be separated out from complex H in this assay system. By 15 min, the pre-mRNA containing the IDDE can form prespliceosome complex A in which U2snRNP is recruited to the 3′ splice site (lanes 2–4 in Fig. 3). The observation that this slower migrating complex appears in an ATP-dependent manner (lanes 9 and 10 in Fig. 3) meets a criterion for complex A. Upon removal of the IDDE, however, formation of the A complex is greatly diminished at all time points analyzed (lanes 5–8 in Fig. 3). The subsequent complexes B and C can hardly be seen at this time. This may be due to the overall lower efficiency of splicing of NMHC-B pre-mRNA compared with widely used in vitro splicing substrates such as β-globin pre-mRNA. The splicing efficiency of NMHC-B pre-mRNAs is severalfold lower compared with that of β-globin pre-mRNA (data not shown). Although whether pre-mRNA lacking the IDDE can recruit U2AF, whose binding is a prerequisite for efficient U2snRNP binding, has not been determined, it does appear that the IDDE facilitates spliceosome assembly at an early stage before formation of complex A.

Substitution of a Py Tract with Optimal Sequences Relieves Splicing Regulation by an Intronic Enhancer

The above data indicate that the 3′ splice site region preceding exon N30 is poorly recognized by spliceosome components in the absence of the IDDE. Therefore, we focused on analyzing this 3′ splice region in detail. For this purpose, the location of the branch site was determined by RT-PCR. It has been reported that some reverse transcriptases pause at the 2′–5′ linkage position of a template RNA but, eventually, are capable of reading through this backbone (33). The lariat-containing RNA mixtures were reverse-transcribed, and the resulting cDNAs were amplified by PCR using the primer set diagrammed in Fig. 4A. This primer set selectively permits the cDNA of the lariat RNA to be amplified, but fails to amplify the cDNA of the linear RNA (Fig. 4A right versus Fig. 4A, left). The sequences of the PCR products are shown in Fig. 4, B1 and 2. The arrows shown are at the junction between the end of the 5′ splice site nucleotide and the branch site. In the first case, reverse transcriptase misincorporates adenosine instead of thymidine at the branch point (Fig. 4B1), as described previously (33). In the second case, this enzyme skips the branch point nucleotide (Fig. 4B2). In either case, the branch site adenine is located –31 nts relative to the first nucleotide of the N30 exon (Fig. 4C). This location of the branch site is typical. The surrounding nucleotide sequence of this A, TGCTAA*C (where A* is the branch site), is well matched to the consensus sequence for the mammalian branch region sequence, YNYTRA*Y (where Y is C or T, R is A or G, and N is A, T, G, or O). The trinucleotide of the 3′ splice donor sequence (CAG) also agrees with the consensus sequence (YAG). However, the sequence between the branch site and the 3′ splice junction, which corresponds to a Py tract, is not typical, and a number of purine residues interrupt the stretch of pyrimidine residues (Fig. 4C, 34). Therefore, the essential splicing factor, U2AF, which binds to the Py tract during E complex formation at the time of spliceosome assembly, would bind to this region inefficiently.

The sequence of these events suggests that the failure of A complex formation in the E5-N30 pre-mRNA lacking the IDDE is likely due to a failure of U2AF binding to the Py tract, rather...
than due to a failure of U2snRNP binding at the branch site. Therefore, we hypothesized that the IDDE may function by facilitating U2AF binding to the suboptimal 3’ splice site preceding the N30 exon to initiate a splicing reaction between exons E5 and N30. If this is the case, changing the Py tract sequence to the optimal sequence should relieve the IDDE requirement for E5-N30 splicing. We, therefore, tested the effect of mutating the Py tract on E5-N30 splicing in vitro. The purine residues located in the Py tract were replaced with pyrimidine residues as shown in Fig. 4 C, m1 or m2 used for in vitro splicing substrates and minigenes are shown. A* represents a branch site. ● indicates unchanged nucleotide.

**FIG. 4.** Determination of the branch site in the E5-N30 intron. A, schematic diagrams of exons E5 and N30 (rectangles) and introns (solid lines). Two arrows, a and b, represent the position and orientation of the primers used for PCR. The exons, introns, and arrows are not drawn to scale. B, DNA sequences of the PCR products. Arrows indicate the junctions between the 5’ end of the intron and the branch site (1) or the nucleotide just to the 5’ side of the branch site (2). Note that the reverse transcriptase misincorporated (1) or skipped (2) the branch site nucleotide. C, genomic DNA sequences flanking the N30 exon. Sequences of the wild-type (wt) and mutants m1 and m2 used for in vitro splicing substrates and minigenes are shown. A* represents a branch site. ● indicates unchanged nucleotide.

**FIG. 5.** Substitution of a Py tract with optimal sequences relieves the requirement for IDDE in E5-N30 splicing in vitro. The E5-N30 substrate RNAs with substitution of the 3’ splice site of N30 with the m1 sequence (see Fig. 4C, Py m1 (optimal)) which includes (+) or lacks (−) the IDDE were incubated with (+) or without (−) nuclear extract (NE) for 2 h. The resulting RNAs were resolved in denaturing urea-polyacrylamide (4.5%) gels. The precursor, intermediate, and final product RNAs are indicated by diagrams at either side of the gels. Asterisks indicate the RNA fragments produced independently of the splicing reaction. The appearance of these fragments varied among different preparations of nuclear extracts.
model by analyzing the effect of overexpression of a truncated U2AF65 fragment on N30 inclusion in transfected cells. The truncated U2AF65 fragment (65 RSL, Fig. 7A) includes the amino-terminal RS domain and the linker region that binds to the U2AF35 subunit and a number of other proteins (36–38) but lacks the carboxyl-terminal RNA binding domain. Overexpression of 65RS-L might sequester U2AF35 and/or other interacting proteins from endogenous intact U2AF65 and, thus, act as a dominant negative mutant. As controls, we also expressed the full-length U2AF65 (65 Full) and another mutant (65 RBD, Fig. 7A) which includes the RNA binding domain but lacks the RS and linker domains. The exogenously expressed proteins contain Myc tag sequences for antibody detection, as well as a nuclear localization signal from the SV40 large T antigen, to ensure that the truncated proteins localized to the nucleus. These expression constructs were cotransfected with the wild-type minigene containing IDDE (Hwt(+)) into Y79 cells, and the splicing patterns of the minigene were analyzed. Dose-dependent expression of the proteins was verified by immunoblot analysis using the mouse anti-Myc antibody (Fig. 7B). The splicing patterns of the minigene transcript in each transfection are shown in Fig. 7C and quantitated in 7D. It is evident that cotransfection of the mutant 65RS-L causes a decrease in the IDDE-mediated N30 inclusion in a dose-dependent manner. This occurred despite lower expression of the 65RS-L protein compared with the other two proteins (Fig. 7B). The full-length U2AF65 shows a slight decrease in N30 inclusion at the highest dosage, but the mutant 65RBD has no effect. The selective inhibition of N30 recognition by overexpressing the RS and linker domains of U2AF65 suggests that IDDE-dependent N30 inclusion is mediated via a protein-protein interaction involving U2AF. Thus, the experiments shown here provide support for a model in which the affinity of U2AF for the 3′ splice site is increased by IDDE binding components via the RS domains and/or the linker region of U2AF65 and U2AF35.

**Discussion**

The cassette-type exon N30 of the NMHC-B gene is subject to neural cell-specific and developmental splicing regulation in mammals and birds (26, 28). Inclusion of the N30 exon is restricted to neural tissues, including retina, although a detailed distribution of N30 expression within neural tissues and the precise identities of N30 expressing neural cells have not yet been characterized. Reflecting the fact that retina mRNA includes the N30 exon to a high degree (included in over 90% of the mRNAs), we utilized human retinoblastoma Y79 cells that are also capable of including the N30 exon in over 80% of the transcripts during the postmitotic and differentiated states triggered by butyrate treatment (14, 28). In a previous report, we described a minigene that reproduces regulative splicing of N30 using Y79 and non-neural cells (14). Thus, the N30 exon in the NMHC-B gene provides an excellent model for studying the molecular basis of neural cell-specific splicing events involving cassette-type exon selection, especially using cultured cells. We previously defined an IDDE in the intron downstream of N30. The majority of this intron is dispensable for N30 inclusion, but the IDDE of 142 nts located 1.5 kb downstream of exon N30 is indispensable for N30 inclusion. In this study, we have focused on determining the target of the IDDE action in relation to the general splicing components.

Both the 3′ and 5′ splice site signals adjacent to the N30 exon deviate from the consensus sequences. The Py tract at the 3′ splice site is suboptimal for U2AF binding. Indeed, this 3′ splice signal is unable to recruit general splicing factors to assemble the spliceosome. The IDDE located downstream of exon N30 facilitates spliceosome assembly in the intron be-
between the upstream exons, E5 and N30. In the splicing reaction in vitro, the IDDE facilitates removal of the intron upstream of N30 but not the intron downstream, where the IDDE resides. Optimization of the Py tract sequence relieves the requirement for the IDDE both in in vitro and in situ splicing. All of the data shown here are consistent with the idea that the IDDE action is targeted primarily at the 3' splice site preceding N30. By using natural and synthetic alternative cassette-type exons, subtle changes in sequences at the 3' and 5' splice sites adjacent to the alternative exon, as well as exon sizes, have been shown to affect alternative exon usage (13, 39–44). Changes in nucleotide sequence to those optimal for snRNA or U2AF binding often enhance alternative exon usage, similar to the effects of mutations m1 and m2 shown here. However, how suboptimal splice signals can be overcome in the native context is not well understood. We have, in addition, shown that the suboptimal nature of the 3' splice signal preceding N30 is overcome by the presence of the IDDE.

How does the IDDE promote 3' splice site recognition? One possible model is that the IDDE recruits one or more factors, and these factors can interact with the U2AF, either the 65- or 35-kDa subunit. The overexpression of the amino-terminal RS and linker domains of U2AF65 (65RS-L), which have been shown to interact with a number of proteins including the U2AF35 subunit and UAP56 (36–38), selectively inhibits splicing of the N30 exon. We interpret this observation as being a consequence of U2AF65-interacting proteins being sequestered from endogenous U2AF65 by the mutant 65RS-L. This result provides indirect support for the above model and suggests the existence of factors that bridge U2AF and the IDDE. These factors are required for N30 inclusion but not for constitutive E5-E6 splicing. The SR proteins, which bind to the exonic enhancers, have been demonstrated to facilitate U2AF binding to the suboptimal 3' splice site via the RS domain of U2AF35 in a number of alternative splicing systems (45). Analogous to the exonic enhancer and SR protein complex, the IDDE and its interacting complex might facilitate recruitment of U2AF to the 3' splice site beyond the upstream alternative exon. Possible interaction of the factors that might bind to exon N30 or the 5' splice site with the putative IDDE-interacting factors is not excluded in this model. Alternatively, or in addition to the above model, the IDDE or the factors that bind to the IDDE might help to release a repressor molecule that occupies the Py tract allowing U2AF to then bind to the Py tract. The best example of such a repressor molecule is sex-lethal in the Drosophila sex determination pathway (46, 47). In a mammalian system, polyypyrimidine tract binding proteins have been shown to be implicated in the 3' splice site inhibition in tissue-specific alternative splicing of several pre-mRNAs including muscle α- and β-tropomyosins, neural Src and γ2 of the GABA\textsubscript{A} receptor (34, 48–51). However, how repression by polyypirimidine tract binding proteins can be released is unknown.

The N30 exon is followed by the unusual 5' splice donor sequence GC instead of GT. The GC splice donor sequence has been found in a number of exons and this suboptimal sequence is targeted primarily at the 3' splice site preceding N30. By using natural and synthetic alternative cassette-type exons, subtle changes in sequences at the 3' and 5' splice sites adjacent to the alternative exon, as well as exon sizes, have been shown to affect alternative exon usage (13, 39–44). Changes in nucleotide sequence to those optimal for snRNA or U2AF binding often enhance alternative exon usage, similar to the effects of mutations m1 and m2 shown here. However, how suboptimal splice signals can be overcome in the native context is not well understood. We have, in addition, shown that the suboptimal nature of the 3' splice signal preceding N30 is overcome by the presence of the IDDE.

How does the IDDE promote 3' splice site recognition? One possible model is that the IDDE recruits one or more factors, and these factors can interact with the U2AF, either the 65- or 35-kDa subunit. The overexpression of the amino-terminal RS and linker domains of U2AF65 (65RS-L), which have been shown to interact with a number of proteins including the U2AF35 subunit and UAP56 (36–38), selectively inhibits splicing of the N30 exon. We interpret this observation as being a consequence of U2AF65-interacting proteins being sequestered from endogenous U2AF65 by the mutant 65RS-L. This result provides indirect support for the above model and suggests the existence of factors that bridge U2AF and the IDDE. These factors are required for N30 inclusion but not for constitutive E5-E6 splicing. The SR proteins, which bind to the exonic enhancers, have been demonstrated to facilitate U2AF binding to the suboptimal 3' splice site via the RS domain of U2AF35 in a number of alternative splicing systems (45). Analogous to the exonic enhancer and SR protein complex, the IDDE and its interacting complex might facilitate recruitment of U2AF to the 3' splice site beyond the upstream alternative exon. Possible interaction of the factors that might bind to exon N30 or the 5' splice site with the putative IDDE-interacting factors is not excluded in this model. Alternatively, or in addition to the above model, the IDDE or the factors that bind to the IDDE might help to release a repressor molecule that occupies the Py tract allowing U2AF to then bind to the Py tract. The best example of such a repressor molecule is sex-lethal in the Drosophila sex determination pathway (46, 47). In a mammalian system, polyypyrimidine tract binding proteins have been shown to be implicated in the 3' splice site inhibition in tissue-specific alternative splicing of several pre-mRNAs including muscle α- and β-tropomyosins, neural Src and γ2 of the GABA\textsubscript{A} receptor (34, 48–51). However, how repression by polyypirimidine tract binding proteins can be released is unknown.

The N30 exon is followed by the unusual 5' splice donor sequence GC instead of GT. The GC splice donor sequence has been found in a number of exons and this suboptimal sequence for U1snRNP binding is implicated in the regulation of some cases of alternative splicing (43, 52, 53). However, the 5' splice site of N30 does not appear to be a direct target of IDDE function, since the splicing of the N30-E6 exons in vitro is not affected by the IDDE. Moreover, a mutation that improves the base complementation of the N30 5' splice site and U1snRNA in the minigenes does not completely eliminate the requirement for the IDDE in N30 inclusion. The IDDE is still required for maximum inclusion of the N30 exon. However, these data do not preclude the possibility that the effect of the IDDE at the 5' splice site might collaborate with or enhance its effect at the 3' splice site.

The mutant minigene bearing the optimal 5' splice site provides useful information for understanding the regulation of N30 splicing. The mutant Hm2(−) lacking the IDDE can in-
clude the N30 exon to a significant extent in Y79 cells, whereas the same minigene does not result in any increase in N30 inclusion in HeLa cells. These results may best be explained by the fact that the interaction of the 5’ splice site and U1snRNP promotes the association of U2AF with the suboptimal Py tract, as demonstrated with the E4 exon of the preprotachykinin gene (54). Component(s) in addition to U1snRNP were suggested to be required for this mechanism (55), and Y79 cells may contain high levels of these factors. HeLa cells with only low levels of these factors and/or an additional repressor mechanism that has not been identified are less sensitive to these changes. These findings with the mutant Hm2 also raise the possibility that components that are involved in the exon spanning between U2AF and U1snRNP might be shared or interact with components that bridge between the IDDE and the 3’ splice site of N30. Moreover, the data support and are consistent with the exon definition model, proposed by Berget (56), in which the exon is first defined as a unit by U2AF/U2snRNP and U1snRNP before the intron is defined as a unit. Interestingly, in the context of the 5’ splice site mutation, the IDDE is capable of promoting the usage of the N30 exon to a small but significant extent, even in non-neural HeLa cells. This observation suggests that non-neural cells contain either a low level or at least some of the factors that can associate with the IDDE. In non-neural cells, the IDDE-protein complex is not sufficient to recruit U2AF at the upstream 3’ splice site, but in collaboration with the U1snRNP-recruited factor(s), it is able to recruit U2AF to initiate the splicing reaction. This observation reinforces the possibility of an interaction of factor(s) bridging the 3’ and 5’ splice sites of N30 with those bridging the IDDE and the 3’ splice site of N30. The finding that the IDDE, in the context of the mutant pre-mRNA, can function in non-neural cells is similar to that with the downstream control sequence (DCS) for Src N1 regulation. The DCS is an intronic downstream enhancer that is required for neural cell-specific inclusion of an 18-nt cassette alternative exon, N1, in the mouse src gene. DCS activates N1 exon inclusion in the native pre-mRNA context in a neural cell-specific manner (10). However, DCS is also capable of activating the inclusion of the alternative exon in the heterologous pre-mRNA context in non-neural cells (57). Thus, similar to DCS, the function of the IDDE itself is not sufficient to determine absolute neural cell specificity for N30 splicing. It appears that the suboptimal 5’ splice site following the N30 exon in NMHC-B pre-mRNA functions as a repressor to ensure the exclusion of N30 in non-neural cells.

In a previous study, the most critical sequence within the IDDE has been identified as the 19-nt sequence, UGCAUGUACGUACUGCAUGU, although just these 19 nts are not sufficient to activate N30 inclusion in the minigene context. This 19-nt sequence includes two copies of a sequence motif UGCAUG. At least one intact copy of this element is indispensable for N30 recognition (14). The importance of this hexanucleotide for the intronic enhancer mechanism was first recognized by a study of alternative splicing of the fibronectin gene (11). This regulatory sequence has been found to play a role in the regulation of alternative splicing of at least five different genes, but these are not restricted to neural cell-specific splicing (10, 11, 14, 15, 17, 18). In the case of fibronectin EIIIB splicing, the action of this enhancer element is targeted at the 5’ splice site (11), but whether or not the upstream 3’ splice site is regulated via this element has not been analyzed.

Neural tissue is one of the tissues in which alternative splicing takes place most frequently. Many examples of neural cell-specific alternative splicing have been reported (58), yet the underlying mechanism(s) is still poorly understood. Only a few of the examples have been characterized in any detail. Of particular note is that the 33-nt DCS is required for neural cell-specific N1 exon recognition for the mouse src gene (10). As noted above, the IDDE of the NMHC-B gene and the DCS of the src gene share a number of common features. They are both required for the neural cell-specific inclusion of a short cassette exon. They are both located in the downstream intron of the alternative exon, although the distances from the exons are quite different. The DCS is 37 nts downstream from the 5’ end of the intron, whereas the IDDE is 1.5 kb downstream. They both contain the element, UGCAUG, which is also found in other intronic enhancers, as mentioned above. Despite these similarities, the IDDE described here activates upstream 3’ splice site recognition and promotes removal of the upstream intron, whereas the DCS has been shown to promote removal of a downstream intron. Therefore, the mechanism by which the IDDE activates N30 usage appears to be different from the mechanism by which the DCS activates N1 usage. The DCS consists of multiple elements which recognize at least six proteins including hnRNP F, hnRNP H, and KSRP (20–22). Each of the factors, which can bind to the DCS, has been demonstrated to be required for N1 inclusion in an in vitro splicing system. However, how these factors affect the general splicing machinery to facilitate N1 exon usage has not been described.

For the last several years, biochemical studies from many laboratories have identified a number of RNA-binding proteins as positive or negative regulators for alternative splicing in vertebrates. These include SR proteins and hnRNP proteins, most of which are expressed ubiquitously. In addition to these generally expressed factors, a recent genetic approach has demonstrated that a neuron-specific RNA binding protein, Nova 1, is involved in the neuron-specific alternative splicing mechanism for some pre-mRNAs (59). The identification of the factor(s) that can bind to the IDDE and how these factors interact with the general splicing component at the 3’ splice site are the next interesting questions.

Acknowledgments—We thank Dr. Robert S. Adelstein (NHLBI) for the continued encouragement andcritical reading of the manuscript. We also thank Drs. Zhi-Ming Zheng and Carl C. Baker for helpful suggestions. We acknowledge the DNA synthesis and DNA sequencing facilities of George Poy (NIDDK) and the excellent editorial assistance of Catherine S. Magruder.

REFERENCES
Splicing Regulation by an Intron Enhancer
