 Nucleic Acids Research, 1999, Vol. 27, No. 12 **2529–2537**

An element in the 5' common exon of the NCAM alternative splicing unit interacts with SR proteins and modulates 5' splice site selection

Jocelyn Côté, Martin J. Simard and Benoit Chabot*

Département de Microbiologie et d'Infectiologie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada

Received January 29, 1999; Revised April 8, 1999; Accepted April 21, 1999

ABSTRACT

The neural cell adhesion molecule (NCAM) gene contains an 801 nt exon that is included preferentially in neuronal cells. We have set up an *in vitro* splicing system that mimics the neuro-specific alternative splicing profile of NCAM exon 18. Splicing regulation is observed using model pre-mRNAs that contain competing 5' or 3' splice sites, suggesting that distinct pathways regulate NCAM 5' and 3' splice site selection. While inclusion of exon 18 is the predominant choice in neuronal cells, an element in the 5' common exon 17 improves exon 17/exon 19 splicing in a neuronal cell line. A similar behavior is observed *in vitro* as the element can stimulate the 5' splice site of exon 17 or a heterologous 5' splice site. The minimal 32 nt sequence of the exon 17 enhancer consists of purine stretches and A/C motifs. Mutations in the purine stretches compromise the binding of SR proteins and decreases splicing stimulation *in vitro*. Mutations in the A/C motifs do not affect SR protein binding but reduce enhancing activity. Our results suggest that the assembly of an enhancer complex containing SR proteins in a 5' common exon ensures that NCAM mRNAs lacking exon 18 are made in neuronal cells.

INTRODUCTION

Although alternative splicing is implicated in the expression of a large number of genes, understanding how the splicing machinery selectively chooses a pair of splice sites in a given cell type or tissue remains a major challenge. In mammals, consensus splicing signals cannot by themselves account for the high specificity associated with splice site selection and pairing in complex pre-mRNAs (reviewed in 1). Indeed, recent studies have identified other *cis*-acting elements and associated *trans*-acting factors that modulate the interaction of U1 snRNP and U2AF with the 5' and the 3' splice sites, respectively (reviewed in 2). SR proteins can influence alternative splicing by promoting proximal 5' or 3' splice site selection on a variety of pre-mRNAs (3–5). Many vertebrate alternative and constitutive exons contain purine-rich splicing enhancers bound by subsets of SR proteins (1). In some situations, this interaction can stabilize U2AF binding to an upstream 3' splice site through RS domain interactions with $U2AF^{35}$ (6–8), while in other cases, interactions with the U1 snRNP 70K protein may improve downstream 5' splice site recognition (7,9,10). Given that the position and sequence of splicing enhancers are important for SR protein activity, tissue-specific differences in the relative abundance of each SR protein may form the basis of alternative splicing regulation in some situations. For example HRS/SRp40 can mediate, in an enhancer-dependent fashion, the specific inclusion of a rat fibronectin EIIIB exon in proliferating liver (11). A similar role for SRp20 was reported in the alternative splicing of its own pre-mRNA (12,13). However, at least for some pre-mRNAs, additional RNA-binding proteins are implicated either in stabilizing the interaction of SR proteins (14,15) or in preventing the interaction of constitutive splicing factors (16–21). In the neuro-specific c-*src* exon N1, the small size of this exon as well as flanking intronic negative elements are responsible for its exclusion in non-neuronal cells (22,23). In this case, PTB binds specifically to the upstream control element to repress splicing of the intron downstream of exon N1 (24). In neuronal cells, inclusion of N1 is mediated through the binding of several proteins, including hnRNP F and KSRP, to an intronic splicing enhancer (25,26).

We have chosen the mouse NCAM pre-mRNA as a model system to study the regulation of alternative splicing. More specifically, we focused on the neuro-specific inclusion of an 801 nt exon (E18) located in the 3'-region of the NCAM pre-mRNA (27). E18 encodes an intracellular cytoplasmic domain of 267 amino acids that is important for adhesion-mediated events (28,29). Analysis of the alternative splicing of E18 has highlighted the importance of the E18 5' splice site sequence in the establishment of regulated splicing profiles (30). In a previous study, we reported that nucleotides +5 to +8 of the E18 intron are involved in base pairing interactions with nucleotides upstream of the major branch site (31). In this case, duplex formation negatively influences NCAM 5' splice site selection in NIH 3T3 fibroblast extracts by interfering with the assembly of early splicing complexes. Here we report the establishment of an *in vitro* splicing system which allows

*To whom correspondence should be addressed. Tel: +1 819 564 5295; Fax: +1 819 564 5392; Email: b.chabot@courrier.usherb.ca
Present address:
Jocelyn Côté, Department of Pediatrics, Division of Genetics, Washington University School of Medicine, 1 Children's Place, Box 8116, St Louis, MO 63110, USA

model NCAM pre-mRNAs containing competing 5' splice sites or competing 3' splice sites to reproduce the regulated splicing profile of E18. In addition, we have identified a 61 nt region in the 5' common exon E17 which stimulates E17/E19 splicing in a neuronal cell line. Using a neuronal extract, we show that a 32 nt portion of this element is bound by SR proteins and can stimulate a nearby downstream 5' splice site. Our findings suggest that the assembly of an enhancer complex in the 5' common E17 is an important event that helps establish the delicate balance between NCAM E18 skipping and inclusion in neuronal cells.

MATERIALS AND METHODS

Cells and transfection procedure

The pCMVdef vector was generated by sequentially inserting into the previously described pCMVSV (32) a *Hin*fI–*Msc*I fragment containing E17 with its 5' splice site and downstream intron sequences, a *Sph*I–*Taq*I fragment containing the complete E18 with portions of flanking introns and a *Sac*I–*Bgl*II fragment containing the 3' splice site of E19. The pCMVdefΔD derivative was produced by removing an *Eco*RI–*Ava*I fragment containing 61 nt of the 5' portion of E17 and upstream vector sequences. Mouse N2a neuroblastoma cells were maintained in DMEM containing 10% fetal bovine serum. Transfection of N2a was accomplished using 15 and 20 μg of Dosper Liposomal Transfection Reagent® (Boehringer Mannheim), respectively. Cells were harvested after 48 h and the RNA extracted using the guanidine-HCl procedure (33). RNA samples were subjected to DNase I digestion followed by heat inactivation and proteinase K treatment. RT–PCR analysis was carried out as previously described (31) with the exception that a modified three-step touchdown PCR protocol (34) was used for the first 12 cycles (hybridization temperatures from 71 to 50°C). Amplification was then carried out for 18 cycles (94°C for 1 min, 50°C for 2 min, 72°C for 3 min). The oligonucleotides used in RT–PCR assays were CMV1 (35), Ne2 (CCTGCAG-ACAGCGCTGTGCC), Ne8 (GCAGTGGCAAAGGTTT), Nf1 (TTGCTTGGTACCCATCATGC), CA$^-$ (GGGGGGAAGTAT-ATAGAGCTCATCGAGATCAC) Pu$^-$ (GTGGTGGATGCAC-ACAGTG-CCCAACGTGACCAC), Nd2 and Nf2 (31). Amplified products were separated in 6% acrylamide gels.

In vitro transcription

pS1 and pC553 were previously described in Chabot *et al*. (35) and Côté and Chabot (31). pC553$^{\Delta D}$ was produced by substituting a *Pst*I fragment of pC553 for an *Ava*I–*Pst*I fragment containing the last 4 nt of E17 with its associated 5' splice site and intron sequences. pAd553 was produced by inserting a *Hin*dIII–*Bgl*II fragment containing the E18–E19 region from pSPE at the *Hin*dIII–*Bam*HI sites in pSPAd (36). pd/Ad553 was generated by substituting the E17-containing *Ava*I fragment of pSPD for a *Pvu*II fragment of pAd553 (37). pC533 consists of a *Sph*I–*Hpa*II fragment containing the 3' splice site of E18 with 150 nt of exon sequences, flanked by the E17 5' splice site and the E19 3' splice site fragments described above for pCMVdef, inserted into the pGEM2 plasmid vector. The structure of all recombinant molecules was verified by extensive restriction analysis and was confirmed by DNA sequencing. All plasmids were linearized at the *Ava*II site prior to *in vitro* transcription. Splicing substrates were synthesized using SP6 RNA polymerase (Pharmacia Biotech) from corresponding linear templates, in the presence of CAP analog (Pharmacia Biotech) and [α-^{32}P]UTP (Amersham Life Sciences). Full-length transcripts were gel purified before use (33).

Splicing extracts and *in vitro* splicing

HeLa, Y79, N2a and NIH 3T3 cell nuclear extracts were prepared as described (38). Splicing reactions were set up and processed as previously described (39), except for Y79, N2a and NIH 3T3 extracts, which were supplemented with 1 U of creatine kinase. Splicing products were either separated on polyacrylamide/urea gels (acrylamide:bisacrylamide 38:2) directly or submitted to RT–PCR amplification, as described (31) with the exception that 9 U of AMV reverse transcriptase (Promega) were used. The oligonucleotides used in the RT–PCR assay were $SP6_{1-20}$ (GAATACAAGCTTGGGCTGCA) and Nf2. The identity of the distal and proximal lariat bands was confirmed by performing a debranching reaction in a S100 extract followed by gel migration next to molecular weight markers.

Gel retardation assays

These assays were carried out as described previously (39), except that heparin was added to a final concentration of 5 mg/ml before loading onto a 6% polyacrylamide gel (acrylamide/bisacrylamide 29:1) containing 5% glycerol. SR proteins were prepared from calf thymus as described (40). Following incubation with nuclear extracts or SR proteins, 1 or 2 μl of concentrated mAb104 (40) were added to the mix prior to gel electrophoresis as before. For these experiments, the mAb104 hybridoma supernatant was concentrated 10-fold using Microcon 3 microconcentrators (Amicon). Templates for the preparation of RNA competitors were prepared by linearizing pSPD (37) at the *Nle*III, *Ban*II or *Ava*I restriction sites, prior to *in vitro* transcription. Cold RNA competitors were synthesized as above except that the relative amount of [α-^{32}P]UTP was reduced 2000-fold. In all cases, competitors were pre-incubated for 10 min with nuclear extracts under gel shift conditions prior to the addition of the respective radiolabeled RNA substrate.

RESULTS

Splicing of a NCAM mini-gene *in vivo*

To identify cell lines that could be used to dissect NCAM alternative splicing, we screened mouse cell lines for their endogenous NCAM splicing profile. A RT–PCR assay using three oligonucleotides was performed to measure the relative levels of mRNAs containing or lacking E18 in NIH 3T3 fibroblasts and in the mouse neuroblastoma cell line N2a. We observed that the prevalence of the E18-containing mRNA species was reduced in NIH 3T3 cells compared to N2a cells (Fig. 1B, lanes 1 and 2). A mini-gene carrying the genomic portion of the NCAM alternative splicing unit from E17, E18 and E19 but lacking the majority of intron sequences (Fig. 1A) was transiently expressed in NIH 3T3 and N2a cells. RT–PCR analysis indicated that most of the amplified products derived from the expression of the mini-gene in NIH 3T3 cells lacked E18, while E18-containing products represented approx-

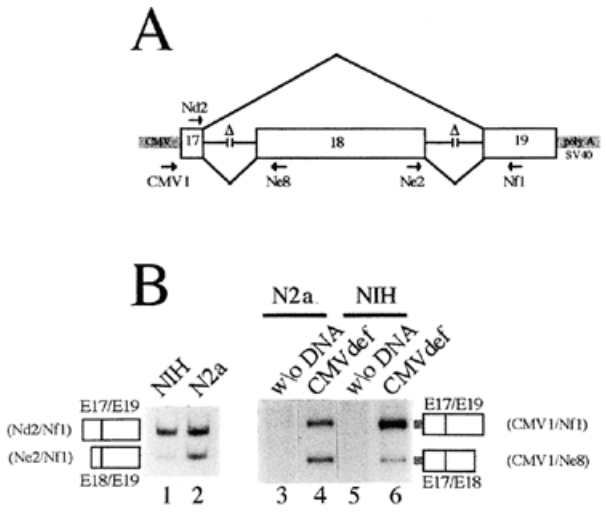

Figure 1. Alternative splicing of NCAM E18 *in vivo*. (**A**) Structure of the NCAM mini-gene. The positions of oligonucleotides used in RT–PCR assays are shown. (**B**) RT–PCR assays using total RNA from N2a and NIH 3T3 cells. Lanes 1 and 2 represent an assay performed with oligonucleotides Nd2, Ne2 and Nf1 to monitor endogenous E18 inclusion levels. Lanes 3–6 represent an assay performed with oligonucleotides CMV1, Ne8 and Nf1 to monitor mRNA species expressed from mock-transfected cells (lanes 3 and 5) and cells transfected with the mini-gene (lanes 4 and 6). Note that the CMV1/Nf1 and Nd2/Nf1 oligonucleotide combinations did not generate the larger E18-containing product in the PCR conditions used (not shown).

imately half of the amplified material in N2a cells (Fig. 1B, lanes 6 and 4, respectively). Thus, the NCAM mini-gene contains all the information necessary to promote cell-specific splicing regulation of NCAM E18.

In vitro splicing of model NCAM pre-mRNAs

Next, we prepared nuclear extracts from N2a and NIH 3T3 cells to determine whether NCAM cell-specific splicing could be reproduced *in vitro*. Extracts from HeLa cells and the human neuronally-derived Y79 cell line were also tested. To facilitate the analysis of this large alternative splicing unit, we constructed a shortened derivative that lacked the 3' splice site of E18 as well as the majority of E18 sequences. This minimal substrate (C553) contains the 5' splice sites of E17 and E18 in competition for the unique 3' splice site of E19 (Fig. 2A). When C553 RNA was incubated in HeLa, N2a or Y79 extracts, the proximal 5' splice site of E18 was selected preferentially, as judged by monitoring lariat splicing intermediates and products on a denaturing polyacrylamide gel (Fig. 2B, lanes 1, 2 and 4, respectively). In contrast, the NIH 3T3 extract was more permissive to distal E17/E19 splicing (Fig. 2B, lanes 3 and 5). Because splicing is relatively inefficient and pre-mRNA degradation obscures direct visualization of mRNAs, we relied on a RT–PCR assay to verify whether the same splicing profile could be seen with the mRNA products (Fig. 2C). The results clearly show that amplified products derived from E18/E19 splicing were predominant in the Y79 extract (Fig. 2C, lane 3), whereas mRNA-derived products arising from E17/E19 splicing were more abundant in the NIH 3T3 extract (lane 6). To determine whether the 5' splice site selection profile observed in NIH 3T3 and Y79 extracts was specific for the NCAM substrate, we tested 5' splice site selection on a non-NCAM pre-mRNA (S1 RNA; see 35). In contrast to the NCAM pre-mRNA, the analysis of lariat molecules derived from S1 RNA indicated that the proximal 5' splice site was used preferentially in both Y79 and NIH 3T3 extracts (Fig. 2D). Thus, NCAM splicing regulation can be reproduced *in vitro* using a

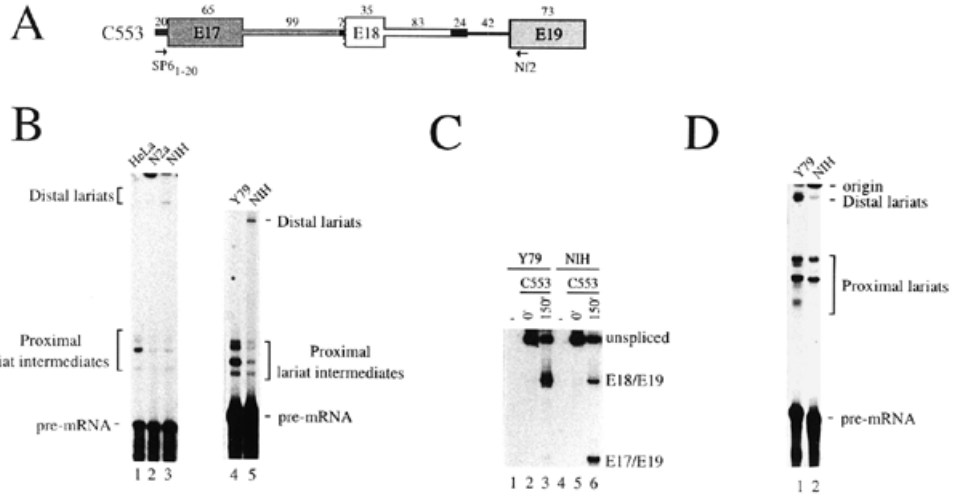

Figure 2. Regulated 5' splice site selection. (**A**) Structure of the model C553 pre-mRNA. The length (in nucleotides) of exon (large rectangles) or intron sequences (thin rectangles) are indicated. The positions of the oligonucleotides $SP6_{1–20}$ and Nf2 used in the RT–PCR assays are also indicated. (**B**) *In vitro* splicing of labeled C553 pre-mRNA. Following incubation in nuclear extracts, labeled RNAs were fractionated in 9% acrylamide/7 M urea gels. Only pre-mRNAs and lariat species corresponding to distal and proximal 5' splice site utilization are shown. Proximal lariat introns migrated below the pre-mRNAs and were obscured by RNA degradation. Lariat molecules are produced through the use of several branch sites (31). (**C**) Monitoring *in vitro* splicing by RT–PCR. To visualize mRNA species derived from C553 splicing *in vitro*, reactions prepared as in (B) but containing unlabeled pre-mRNA were analyzed by RT–PCR using oligonucleotides $SP6_{1–20}$ and Nf2. Labeled amplified products were separated on a 6% acrylamide gel. (**D**) *In vitro* splicing of an unrelated pre-mRNA. A labeled pre-mRNA carrying the 5' splice sites of hnRNP A1 exons 7 and 7B in competition for the adenovirus L2 3' splice site (35) was used in NIH 3T3 and Y79 extracts. RNA products were fractionated in an 8.5% acrylamide/8 M urea gel.

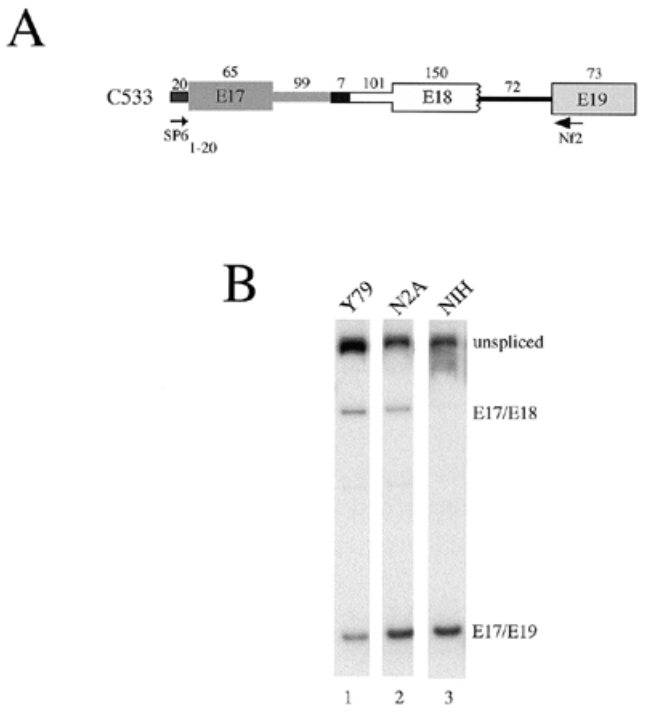

Figure 3. Regulated 3' splice site selection. (**A**) Structure of the model C533 pre-mRNA. The length of exon and intron fragments and the positions of oligonucleotides $SP6_{1–20}$ and Nf2 used in the RT–PCR assays are indicated. (**B**) *In vitro* splicing of C533 pre-mRNA detected by RT–PCR. The mRNAs derived from unlabeled C533 splicing *in vitro* were analyzed by RT–PCR using oligonucleotides $SP6_{1–20}$ and Nf2. Labeled amplified products were separated on a 6% acrylamide gel.

model NCAM pre-mRNA substrate lacking most of E18 sequences and its associated 3' splice site.

A previous analysis by Tacke and Goridis (30) had indicated that although the 5' splice site of E18 was important, E18 sequences and the 3' splice site might also play a role in splicing regulation in N2a cells. To verify whether regulation could be observed in the absence of the E18 5' splice site, we used a model pre-mRNA which contains the 5' splice site of E17 and the 3' splice sites of both E18 and E19 (C533; Fig. 3A). Following incubation of C533 RNA in nuclear extracts, the profile of mRNA-derived products revealed that the E19 3' splice site was selected preferentially in the NIH 3T3 extract (Fig. 3B, lane 3). In contrast, proximal E17/E18 splicing was more efficient in the Y79 and N2a extracts (Fig. 3B, lanes 1 and 2). A similar result was observed when monitoring lariat products (data not shown). Again, this splicing profile was specific for the NCAM substrate (data not shown). The fact that NCAM E18 splicing regulation can be reproduced *in vitro* using pre-mRNAs carrying either competing 5' splice sites or competing 3' splice sites suggests that the regulated splicing of NCAM E18 involves independent control of 5' and 3' splice site selection.

An element in exon 17 modulates NCAM 5' splice site selection

Most studies that have addressed the control of alternative splicing have focused on elements present in alternative exons or in flanking intron sequences. Although purine-rich enhancers have also been uncovered in the 5' or 3' common exon of a few alternative splicing units (39,41), their role in alternative splicing remains unclear. To ascertain whether the E17 common exon in the NCAM alternative splicing unit contained such an element, we constructed a derivative of the NCAM mini-gene lacking 61 bp in E17 (Fig. 4A) and tested the effects of this deletion on NCAM splicing *in vivo* and *in vitro*. *In vivo* this deletion decreased the relative frequency of E17/E19 splicing in N2a cells (Fig. 4A, compare lane 2 with lane 3). Because E18 inclusion is up-regulated in N2a cells, the E17 element may be important to maintain the production of NCAM mRNAs lacking E18. To verify whether the E17 element stimulated E17/E19 splicing, we tested the effect of the deletion *in vitro* using a derivative of the C553 pre-mRNA (Fig. 4B). Deletion of the 61 nt region eliminated E17/E19 splicing in a HeLa extract, as judged by the disappearance of lariat species diagnostic of distal 5' splice site utilization ($C553^{\Delta D}$; Fig. 4C, lane 2). The same result was obtained in a Y79 extract (data not shown). As we could not eliminate the possibility that the E17 5' splice site was used poorly because of the small size of the distal exon, we relied on a heterologous substrate to verify the enhancing activity of the E17 sequences. Grafting the E17 element upstream of the adenovirus L1 5' donor site (d/Ad553; Fig. 4B) stimulated distal 5' splice site utilization in HeLa, Y79 and N2a extracts (Fig. 4D, lanes 1–6, and data not shown). These results suggest that an element within the 61 nt sequence of E17 stimulates the use of an adjacent downstream 5' splice site in neuronal and HeLa extracts. Moreover, the ability of the E17 element to activate a heterologous 5' splice site suggests that the effect is mediated by a *trans*-acting factor.

SR proteins interact with the E17 element

The 61 nt coding region of NCAM E17 is conserved in several vertebrate species, the similarity with the mouse sequence varying between 77% for *Xenopus* and 92% for the rat (Fig. 5A). The most conserved region in E17 is a 17 nt sequence located within a 37 nt stretch lacking thymidines and beginning with nine consecutive purines (GGGGGGAAG in mouse). To address whether cellular factors associate with the 61 nt region in E17, we performed gel shift assays (Fig. 5A and B). Transcripts made by terminating transcription at the *Nle*III site did not assemble into complexes following incubation in a Y79 extract, in a mixture of SR proteins purified from calf thymus or in a HeLa S100 fraction which lacks SR proteins (data not shown). Likewise, an RNA fragment containing the highly conserved 17 nt sequence (boxed in Fig. 5A) did not stably interact with nuclear factors (data not shown). In contrast, a transcript terminating at the *Ava*I site formed complexes in the Y79 extract, in the SR protein mixture but not in the HeLa S100 fraction (Fig. 5B, lanes 6, 15 and 16, respectively). SR proteins in the Y79 extract contributed to the shift of the E17 element since co-incubation with the anti-SR antibody mAb104 disrupted complex formation (Fig. 5C, lanes 2–4). The addition of a heat-inactivated mAb104 had no effect on the migration of the complex (data not shown). A transcript terminating at the *Ban*II site also formed complexes, but with reduced efficiency compared to the complete 61 nt element (Fig. 5B, lanes 1–4). These results suggest that SR proteins interact with sequences in E17 and that sequences downstream from the *Ban*II site are required for stable

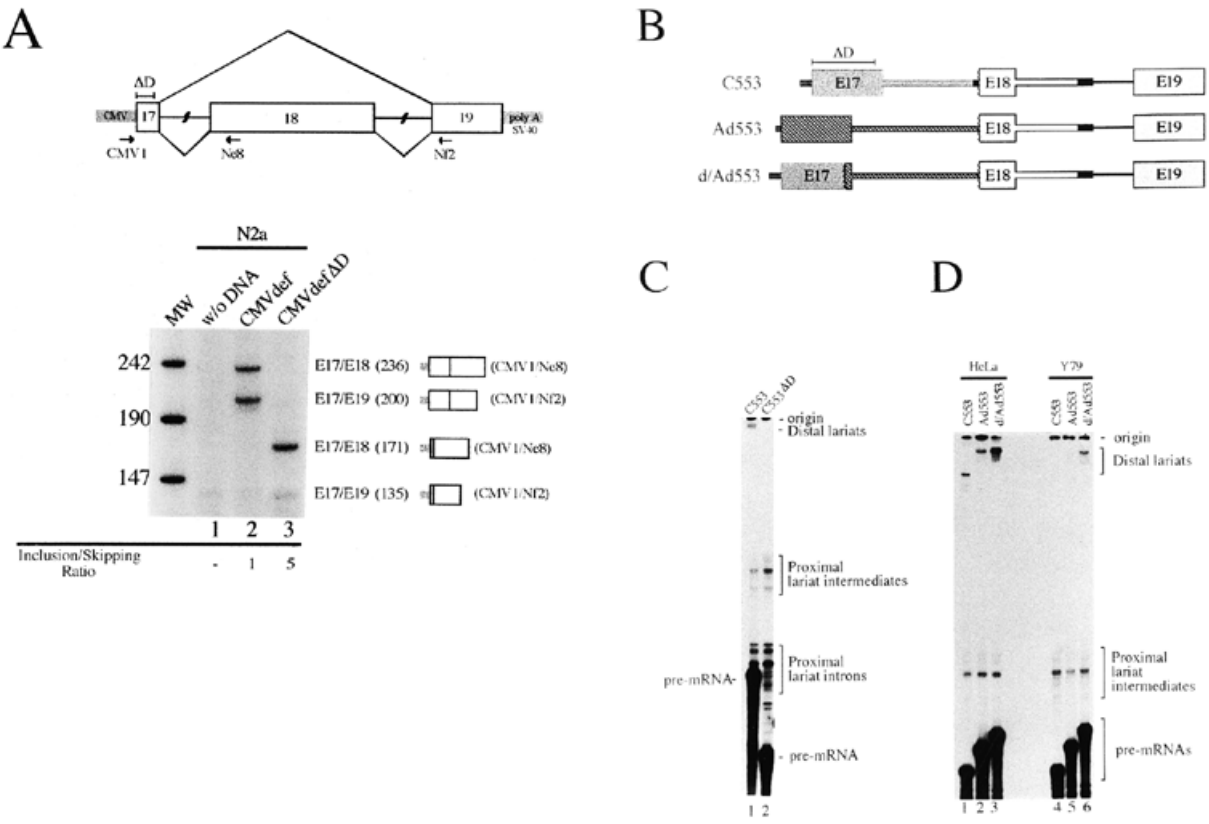

Figure 4. Sequences within constitutive E17 improve E17/E19 splicing in neuronal cells and extracts. (**A**) RT–PCR analysis using oligonucleotides CMV1, Ne8 and Nf2 was performed on total RNA from cells transfected with the intact mini-gene or the mini-gene carrying a 61 nt deletion in E17. The sizes of the RT–PCR fragments in base pairs are indicated and their migration can be compared to molecular weight markers (MW). (**B**) Structure of C553 and derivatives. (**C**) C553 and a derivative that carries a 61 nt deletion in E17 ($C553^{\Delta D}$) were incubated in a HeLa extract. Proximal and distal lariat molecules were fractionated in a 9.5% acrylamide/8 M urea gel. (**D**) The E17 sequences promote distal 5' splice site selection on a pre-mRNA substrate carrying a heterologous 5' splice site. The 61 nt sequence in E17 was inserted in a C553 derivative in which the E17 5' splice site region was substituted for the adenovirus L1 5' splice site and its associated intron region. Splicing products were fractionated on an acrylamide/urea gel.

complex formation. This conclusion was supported by performing the assay in the presence of excess of cold competitor. In contrast to a competitor containing the complete 61 nt region, a competitor RNA containing the first 34 nt of the E17 region was unable to inhibit complex formation on the full-length element (Fig. 5B, lanes 7–14).

To further delineate the sequences required for stable complex formation, we tested a central 32 nt region of the E17 element which begins with nine consecutive purines and is followed by three A/C motifs separated by two short purine stretches (Fig. 6A, WT). Incubation of wild-type RNA in a Y79 extract or in the SR proteins mixture led to efficient complex formation (Fig. 6B, lanes 7–9). Insertion of this element upstream of the adenovirus L1 5' splice site stimulated splicing at that site (Fig. 6C, lane 3). To address the contribution of the purine elements and the A/C motifs, we tested mutated versions of the 32 nt element. Although substituting selected nucleotides throughout the A/C motifs for uridines did not affect complex formation (CA^-; Fig. 6A and B, lanes 10–12), stimulation of the distal 5' splice site was compromised (Fig. 6C, lane 4). In contrast, replacing selected nucleotides in the purine stretches for uridines nearly abolished complex formation in the Y79 extract and in the mixture of SR proteins (Pu^-; Fig. 6A and B, lanes 13–15). Compared to the wild-type element, inserting the mutated Pu^- element upstream of the L1 5' splice site reduced the relative level of distal 5' splice site utilization (Fig. 6C, lane 5). Thus, although intact purine stretches are sufficient for complex formation, both the purine elements and the A/C motifs contribute to stimulation of an adjacent downstream 5' splice site.

Further evidence for the role of SR proteins in the activity of the E17 element was obtained by monitoring the effect of adding SR proteins to a Y79 extract. The addition of a mixture of SR proteins from calf thymus greatly stimulated proximal 5' splice site usage when the Pu^- element (not bound by SR proteins) was present upstream of the distal site (Pu^-/Ad553; Fig. 7B, lanes 4–6). In contrast, when the pre-mRNA included the CA^- element (bound by SR proteins) upstream of the distal site (CA^-/Ad533; lanes 1–3), distal 5' splice site selection remained efficient despite the addition of SR proteins. These results suggest that the binding of SR proteins upstream of the distal 5' splice site helps maintain selection at this site and are consistent with a role for SR proteins in mediating the enhancing activity of the E17 element.

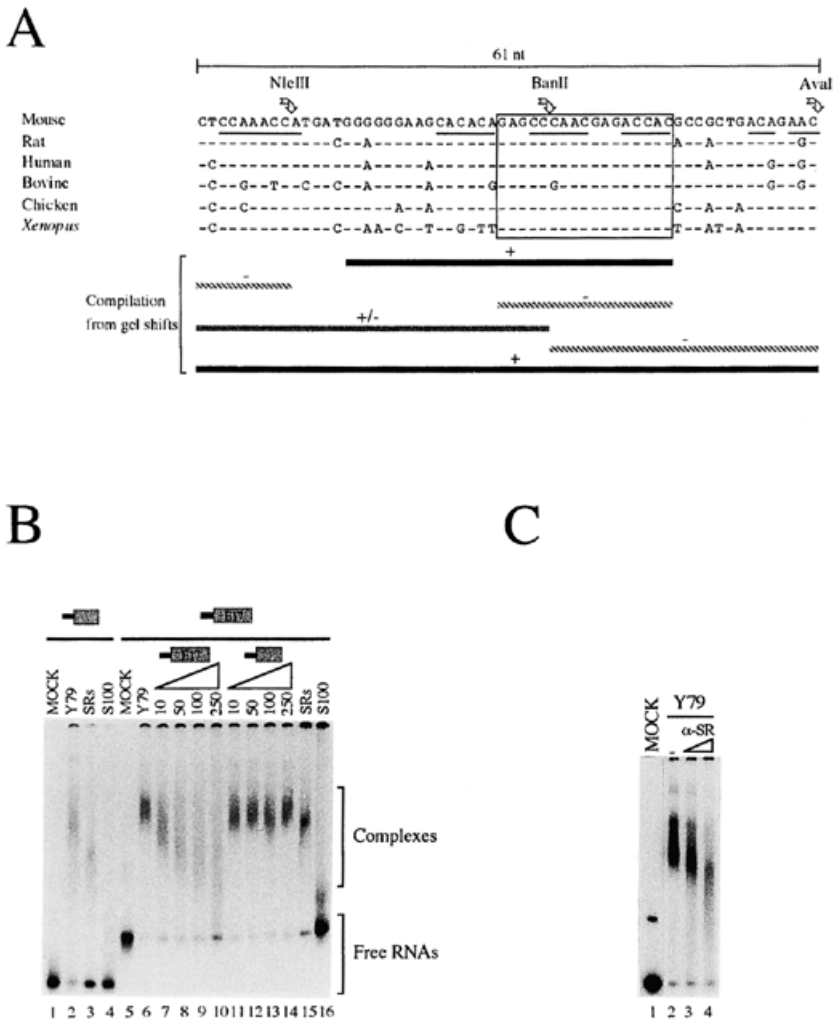

Figure 5. SR proteins bind to the E17 element. (**A**) Comparison of E17 DNA sequences in various vertebrate species. The sites of cleavage for some restriction enzymes are shown. The stretch of well-conserved nucleotides is boxed and the A/C motifs are underlined. Results of gel shift experiments in neuronal extracts are summarized below; a black bar indicates a strong shift (+), a dark gray bar indicates a weaker shift (+/–) and a dashed gray bar indicates no shift (–). (**B**) Gel shift assays in a Y79 extract, a HeLa S100 fraction extract and purified SR proteins from calf thymus. RNA extending to the *Ban*II (lanes 1–4) and *Ava*I (lanes 5–16) sites were tested. The position of free RNAs and bound complexes are shown. Increasing amounts of cold competitor RNAs terminating at the *Ava*I site (lanes 7–10) or the *Ban*II site (lanes 11–14) were added to the nuclear extract prior to the addition of radiolabeled substrates. The numbers above these lanes correspond to molar excesses of competitor. (**C**) SR proteins are present in neuronal complexes assembled on E17. An antibody that recognizes phosphorylated SR proteins (mAb104) was added to the Y79 extract during incubation.

DISCUSSION

A mini-gene construct containing the alternative splicing unit of murine NCAM E18 and lacking the majority of flanking intron sequences reproduced the neuro-specific splicing profile of the endogenous NCAM pre-mRNA. Extracts prepared from N2a and NIH 3T3 cells mimicked the *in vivo* splicing profiles. *In vitro* cell-specific splicing was observed with two sets of pre-mRNAs lacking most of the E18 sequences and containing either the 5' splice site or the 3' splice site of E18. The ability to achieve cell-specific regulation of splicing in the absence of the complete set of E18 splice sites suggests that distinct mechanisms control the selection of the 5' and the 3' splice sites of E18. In small exons, splice site selection can be coordinated through exon bridging interactions (6,42). In the case of NCAM, regulatory events targeting the individual splice sites of E18 may be required, given that E18 is an unusually large exon that may be refractory to communication between the E18 splice sites through the exon definition pathway.

Our first attempt at investigating the contribution of elements involved in NCAM E18 splicing in neuronal cells has led to the identification of a small 61 nt exon element located in the 5' common exon E17. This element is required for E17/E19 splicing in neuronal cells and extracts. Because the splice sites of common exons in alternative splicing units are often suboptimal, stimulation of E17/E19 splicing may compensate for the activation of E18/E19 splicing in neuronal cells, allowing maintenance of an adequate balance between the two forms

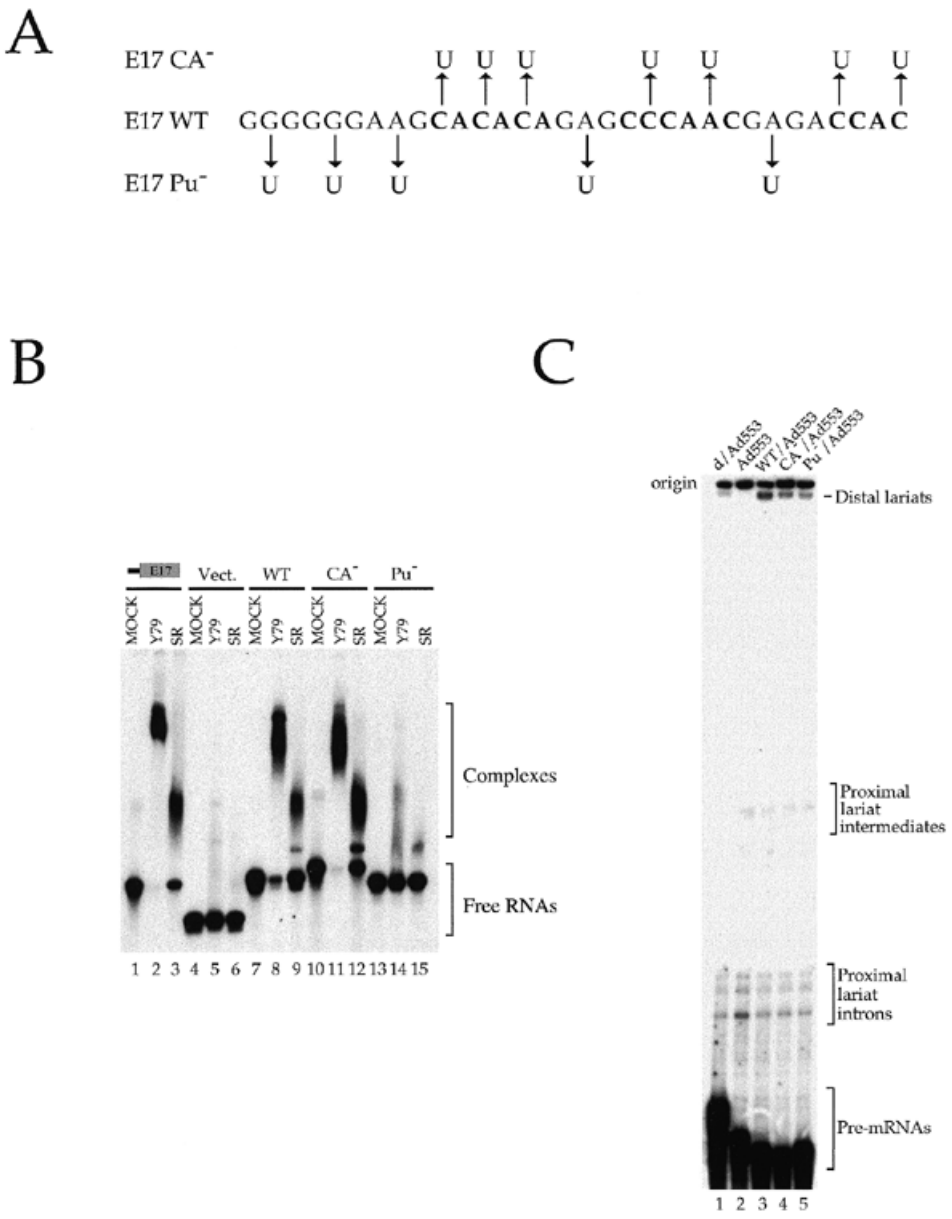

Figure 6. The purine elements are required for complex formation and 5' splice site stimulation. (**A**) Sequence of the minimal 32 nt element from E17. The A/C motifs interspersed between the purine elements are shown in bold. Positions mutated to U are indicated above and below the sequence. (**B**) Gel shift assays in a Y79 extract and purified SR proteins from calf thymus. E17 RNA terminating at the *Ava*I site was tested as before (lanes 1–3). Vect. represents an RNA transcribed from pBluescript KS using T3 RNA polymerase and terminating at the *Eco*RI site (lanes 4–6). WT, CA^- and Pu^- consist of labeled RNAs derived from oligonucleotides inserted at the *Eco*RV site of pBluescript KS, transcribed from the T3 promoter and terminating at the *Eco*RI site (lanes 7–15). The position of free RNAs and complexes are shown. (**C**) The minimal E17 sequence promotes distal 5' splice site utilization in a Y79 extract. Oligonucleotides carrying the wild-type 32 nt sequence of E17 or mutated derivatives were inserted at the *Pvu*II site of pAd533, upstream of the distal 5' splice site. Splicing products were fractionated on an 11% acrylamide/8 M urea gel. Splicing intermediates and products are indicated on the right.

of NCAM proteins involved in cellular adhesion events. The fact that E17/E19 splicing must be accomplished in an environment where E18 inclusion is activated suggests that modulating the binding of factors and the interaction between factors bound at all splice sites, including the constitutive ones, involves a complex and tightly coordinated series of events in neuronal cells.

The participation of *trans*-acting factors mediating the activity of the E17 element is suggested by the observation that the element can also stimulate a heterologous 5' splice site. In a neuronal extract, the E17 element assembles into a stable complex containing SR proteins. We have shown that a 32 nt sequence within E17 is sufficient for distal 5' splice site activation, complex formation and the binding of SR proteins. This element contains three purine stretches and three A/C motifs. The replacement of selected purines by uridines in the purine elements compromises the binding of SR proteins and weakens enhancing activity. Consistent with a role for SR

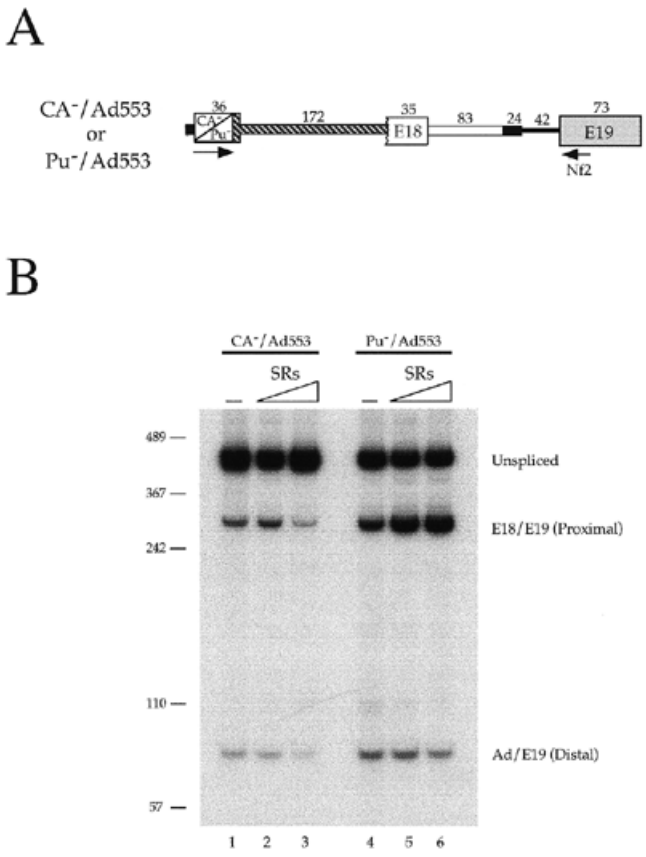

Figure 7. Involvement of SR proteins in the activity of the E17 element. (**A**) Structure of the pre-mRNAs. The positions of the oligonucleotides used in the RT–PCR assay are indicated. (**B**) *In vitro* splicing assays in a Y79 extract were performed in the presence of increasing amounts (0, 70 or 140 ng) of purified SR proteins from calf thymus. The mRNAs derived from *in vitro* splicing of CA⁻/Ad553 and Pu⁻/Ad553 RNA were analyzed by RT–PCR using oligonucleotides CA⁻/Pu⁻ and Nf2, respectively. Labeled amplified products were separated on a 6% acrylamide gel. The positions of some molecular weight markers are indicated on the left.

proteins in mediating the activity of the element, grafting a mutated element (CA⁻) still bound by SR proteins offered resistance against the loss of distal 5' splice site selection when the concentration of SR proteins is increased. However, because a mutated element defective in SR protein binding (Pu⁻) still displayed a moderate level of enhancing activity, SR proteins are likely not entirely responsible for the activity of the E17 enhancer element. The contribution of additional factors is further suggested by the effect of mutations in the A/C motifs which, although not contributing to complex formation and the binding of SR proteins, reduced the enhancing activity. A/C-rich motifs are present in a variety of enhancer elements (43) and have recently been identified as part of a bipartite enhancer in an alternative exon of the minute virus of mice (44). A/C-rich motifs in the repeated elements of the *dsx* enhancer in *Drosophila* are bound by Tra2 (14,15,45,46) and human homologs of Tra2 bind to and mediate the enhancing activity of A-rich elements *in vitro* (47). The role of mammalian Tra2 proteins in the activity of the NCAM E17 element remains to be verified. However, whereas Tra2 collaborates with Tra to recruit or stabilize SR protein binding to the *dsx* element, an E17 element with mutated A/C motifs is still efficiently assembled into complexes containing SR proteins.

Whether the E17 element and associated factors also contribute to efficient E18 skipping in non-neuronal cells is currently being investigated. Although SR proteins may play a role in regulation, western analyses suggest that the differences in splicing between neuronal and NIH 3T3 cells are not due to quantitative or qualitative differences in phosphorylated SR proteins (data not shown). Moreover, cell-specific differences in NCAM splice site usage *in vitro* are not seen when using unrelated pre-mRNAs. Thus, NCAM splicing regulation cannot be explained solely in terms of the previously documented effects of SR proteins on splice site selection.

While exon splicing enhancers have initially been identified by their ability to activate 3' splice site utilization, a growing number of exon elements have been shown to influence 5' splice site selection. In the caldesmon pre-mRNA, a purine-rich splicing enhancer favors the use of the upstream 5' splice site (48), while its substitution for the troponin T enhancer activates the downstream site (49). In the *fruitless* pre-mRNA of *Drosophila*, activation of the downstream female-specific 5' splice site requires a *cis*-acting element bound by Tra and Tra2 (10). The NCAM E17 element therefore represents the first documented mammalian enhancer that can activate a downstream 5' splice site in a natural context. The assembly of an enhancer complex bound to E17 may stimulate U1 snRNP binding to the flanking 5' splice site through an interaction between the RS domains of SR and the U1 snRNP 70K proteins (8,9).

Specific profiles of alternative splicing likely reflect subtle differences in the interactions of splicing factors bound at competing pairs of splice sites. While the participation of a variety of elements located in alternative exons and in flanking introns is becoming apparent, the contribution of enhancer elements located in flanking common exons has received little attention (6,39,41,50). Purine-rich elements with enhancer activity have been uncovered in the 3' common exon of the fibronectin ED1 alternative splicing unit (39) and the 5' common exon of the calcitonin/CGRP alternative splicing unit (41). To our knowledge, the NCAM E17 element is the first example of a modulating element located in a 5' common exon that can promote a decrease in the inclusion frequency of a nearby alternative exon in its natural setting. Because the presence of an enhancer element in a common exon may help offset the activity of other elements that increase the inclusion of the alternative exon, it would not be surprising if enhancer elements in the common exons of other alternative splicing units also contribute to the delicate control of splice site selection.

ACKNOWLEDGEMENTS

We thank Marco Blanchette for helpful discussions, Johanne Toutant for transfections and help in the preparation of nuclear extracts and Alain Lavigueur, Raymund Wellinger and Jane Wu for comments on the manuscript. J.C. and M.J.S. are recipients of a studentship from the FCAR. This work was supported by a grant from the Medical Research Council of Canada. B.C. is a Chercheur Boursier Senior from the FRSQ.

REFERENCES

1. Black,D.L. (1995) *RNA*, **1**, 763–771.
2. Chabot,B. (1996) *Trends Genet.*, **12**, 472–478.
3. Fu,X.D. (1995) *RNA*, **1**, 663–680.
4. Horowitz,D.S. and Krainer,A.R. (1994) *Trends Genet.*, **10**, 100–106.
5. Manley,J.L. and Tacke,R. (1996) *Genes Dev.*, **10**, 1569–1579.
6. Wang,Z., Hoffmann,H.M. and Grabowski,P.J. (1995) *RNA*, **1**, 21–35.
7. Wu,J.Y. and Maniatis,T. (1993) *Cell*, **75**, 1061–1070.
8. Zuo,P. and Manley,J.L. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3363–3367.
9. Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) *Nature*, **368**, 119–124.
10. Heinrichs,V., Ryner,L.C. and Baker,B.S. (1998) *Mol. Cell. Biol.*, **18**, 450–458.
11. Du,K., Peng,Y., Greenbaum,L.E., Haber,B.A. and Taub,R. (1997) *Mol. Cell. Biol.*, **17**, 4096–4104.
12. Jumaa,H. and Nielsen,P.J. (1997) *EMBO J.*, **16**, 5077–5085.
13. Jumaa,H., Guenet,J.L. and Nielsen,P.J. (1997) *Mol. Cell. Biol.*, **17**, 3116–3124.
14. Lynch,K.W. and Maniatis,T. (1996) *Genes Dev.*, **10**, 2089–2101.
15. Tian,M. and Maniatis,T. (1993) *Cell*, **74**, 105–114.
16. Adams,M.D., Tarng,R.S. and Rio,D.C. (1997) *Genes Dev.*, **11**, 129–138.
17. Ashiya,M. and Grabowski,P.J. (1997) *RNA*, **3**, 996–1015.
18. Lin,C.H. and Patton,J.G. (1995) *RNA*, **1**, 234–245.
19. Singh,R., Valcárcel,J. and Green,M.R. (1995) *Science*, **268**, 1173–1176.
20. Pérez,I., Lin,C.H., McAfee,J.G. and Patton,J.G. (1997) *RNA*, **3**, 764–778.
21. Valcárcel,J., Singh,R., Zamore,P.D. and Green,M.R. (1993) *Nature*, **362**, 171–175.
22. Black,D.L. (1991) *Genes Dev.*, **5**, 389–402.
23. Chan,R.C. and Black,D.L. (1995) *Mol. Cell. Biol.*, **15**, 6377–6385.
24. Chan,R.C. and Black,D.L. (1997) *Mol. Cell. Biol.*, **17**, 4667–4676.
25. Min,H., Chan,R.C. and Black,D.L. (1995) *Genes Dev.*, **9**, 2659–2671.
26. Min,H., Turck,C.W., Nikolic,J.M. and Black,D.L. (1997) *Genes Dev.*, **11**, 1023–1036.
27. Barbas,J.A., Chaix,J.C., Steinmetz,M. and Goridis,C. (1988) *EMBO J.*, **7**, 625–632.
28. Pollerberg,G.E., Schachner,M. and Davoust,J. (1986) *Nature*, **324**, 462–465.
29. Pollerberg,G.E., Burridge,K., Krebs,K.E., Goodman,S.R. and Schachner,M. (1987) *Cell Tissue Res.*, **250**, 227–236.
30. Tacke,R. and Goridis,C. (1991) *Genes Dev.*, **5**, 1416–1429.
31. Côté,J. and Chabot,B. (1997) *RNA*, **3**, 1248–1261.
32. Yang,X., Bani,M.-R., Lu,S.-J., Rowan,S., Ben-David,Y. and Chabot,B. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 6924–6928.
33. Chabot,B. (1994) In Hames,D. and Higgins,S. (eds), *RNA Processing: A Practical Approach*. IRL Press, Oxford, UK, Vol. I, pp. 1–29.
34. Don,R.H., Cox,P.T., Wainwright,B.J., Baker,K. and Mattick,J.S. (1991) *Nucleic Acids Res.*, **19**, 4008.
35. Chabot,B., Blanchette,M., Lapierre,I. and La Branche,H. (1997) *Mol. Cell. Biol.*, **17**, 1776–1786.
36. Solnick,D. (1985) *Cell*, **42**, 157–164.
37. Côté,J., Beaudoin,J., Tacke,R. and Chabot,B. (1995) *J. Biol. Chem.*, **270**, 4031–4036.
38. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
39. Lavigueur,A., La Branche,H., Kornblihtt,A.R. and Chabot,B. (1993) *Genes Dev.*, **7**, 2405–2417.
40. Zahler,A.M., Lane,W.S., Stolk,J.A. and Roth,M.B. (1992) *Genes Dev.*, **6**, 837–847.
41. Yeakley,J.M., Hedjran,F., Morfin,J.P., Merillat,N., Rosenfeld,M.G. and Emeson,R.B. (1993) *Mol. Cell. Biol.*, **13**, 5999–6011.
42. Robberson,BL., Cote,G.J. and Berget,S.M. (1990) *Mol. Cell. Biol.*, **10**, 84–94.
43. Coulter,L.R., Landree,M.A. and Cooper,T.A. (1997) *Mol. Cell. Biol.*, **17**, 2143–2150.
44. Gersappe,A and Pintel,D.J. (1999) *Mol. Cell. Biol.*, **19**, 364–375.
45. Lynch,K.W. and Maniatis,T. (1995) *Genes Dev.*, **9**, 284–293.
46. Tian,M. and Maniatis,T. (1994) *Genes Dev.*, **8**, 1703–1712.
47. Tacke,R., Tohyama,M., Ogawa,S. and Manley,J.L. (1998) *Cell*, **93**, 139–148.
48. Humphrey,M.B., Bryan,J., Cooper,T.A. and Berget,S.M. (1995) *Mol. Cell. Biol.*, **15**, 3979–3988.
49. Elrick,L.L., Humphrey,M.B., Cooper,T.A. and Berget,S.M. (1998) *Mol. Cell. Biol.*, **18**, 343–52.
50. Tsukahara,T., Casciato,C. and Helfman,D.M. (1994) *Nucleic Acids Res.*, **22**, 2318–2325.