In vitro splicing of fibronectin pre-mRNAs

Pamela A. Norton* and Richard O. Hynes¹

Center for Cancer Research and ¹Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received May 14, 1990; Accepted June 22, 1990

ABSTRACT

We have investigated the alternative splicing of the EIIIB exon of the rat fibronectin gene. Mini-gene constructs containing this exon and portions of adjacent introns and exons, when transfected into HeLa cells, are transcribed and spliced, but omit the EIIIB exon. *In vitro*, HeLa nuclear extracts similarly splice out (skip) the EIIIB exon from similarly structured transcripts. Therefore, the HeLa splicing apparatus recognizes as atypical the EIIIB exon and its flanking intron sequences, both *in vivo* and *in vitro*. We also report that alterations in the ionic conditions of the *in vitro* splicing reaction can promote the initiation of EIIIB exon inclusion, as reflected by the formation of intermediate and product RNAs related to the removal of the intron upstream of EIIIB. Processing of this intron correlates with the formation of complexes resembling intermediates in spliceosome assembly. The branch sites involved in this alternative processing pathway are rather distant from the EIIIB 3' splice site, and lie within a region which is well conserved in the fibronectin genes of other species. Thus, the intron upstream of EIIIB shows singular structure and behavior which probably have a bearing on the regulated alternative splicing of this exon.

INTRODUCTION

Splicing of eukaryotic pre-mRNAs is generally precise and invariant. Intron excision and exon ligation occur via two coordinated trans-esterification reactions, involving the formation of a branched lariat intron (reviewed in 1–3). The splicing reactions require conserved sequences at both the 5' and 3' ends of introns as well as at the site of the lariat branch which is typically 18–37 nucleotides upstream of the 3' splice site. The use of *in vitro* splicing systems has led to the identification of a number of cellular factors essential for splicing, which include several abundant snRNPs and a number of protein factors (4,5). Formation of large macromolecular complexes called spliceosomes has been detected *in vitro* (6–8). It is presumed that the assembly of pre-RNAs into such ordered structures provides the basis for the specificity of the ultimate splicing reactions.

An increasing number of genes have been identified which

produce multiple mRNAs by alternative RNA splicing. Several patterns of alternative splicing have been described (9,10). Multiple 5' or 3' splice sites can be paired with a common partner, and events such as exon skipping involve the failure to recognize both ends of an exon, but little is known about how such choices are made. Features of primary sequence alone are insufficient to account for alternative processing, as most alternatively used splice sites resemble normal junctions. Secondary structures are capable of sequestering splice sites from the splicing machinery under certain circumstances (11,12), but may not be involved in all alternative reactions (13). Trans-acting factors seem to be involved in alternative splice site selection, as many such events are cell-type-specific. The nature of such factors is largely unknown, although it has recently been suggested that a number of RNA binding proteins and U snRNPs may act as regulatory splicing factors (14,15).

We have been characterizing the alternative splicing of the rat fibronectin (FN) pre-mRNA. The large (>70 kb) rat gene produces a single pre-mRNA (16) which is subject to alternative splicing at three positions (17); similar variation is seen in human, chicken and frog FN mRNAs (18–21, and D. DeSimone, P. Norton and R.O. Hynes, manuscript in preparation). The rat and human FN genes are similarly organized (22,23), and the chicken gene appears to have the same intron-exon structure as the mammalian genes (24). Splicing at two positions involves the inclusion or skipping of single exons, termed EIIIA and EIIIB, which encode discrete protein segments. The third position, V, involves variable subdivision of a large exon by pairing multiple 3' splice sites with a common 5' site. All these splicing variations have been shown to be cell-type-specific (see 25 and references therein), from which the existence of cell-type-specific regulatory splicing factors has been inferred.

Transfection studies have shown that small fragments of FN genomic DNA can support alternative processing (26–30), indicating that the necessary cis-acting elements are located within or near the regulated exons. However, this experimental strategy is not suited to examination of trans-acting regulatory factors. *In vitro* splicing systems have proved highly successful in the elucidation of the essential components of the splicing machinery. Alternative splicing should be amenable to similar dissection, and we have thus begun to analyze the *in vitro* processing of the rat EIIIB exon in HeLa nuclear extract. EIIIB is largely excluded from the FN mRNAs of most cultured cell types including HeLa

* To whom correspondence should be addressed

cells (17, 29,30) but it is widely included in early embryos (21,31). We report that transfection of HeLa cells with a fragment of the rat FN gene lacking most of the introns flanking EIIIB gave rise to EIIIB⁻ transcripts. EIIIB skipping was also reproduced *in vitro* under standard splicing conditions. Altering the ionic conditions of the *in vitro* reaction activates a splicing pathway related to EIIIB inclusion, and we describe features of these alternative processing events.

MATERIALS AND METHODS

DNA manipulations

All DNA manipulations were according to standard procedures. Restriction and DNA modifying enzymes were from New

England Biolabs and Boehringer Mannheim. The transcription template 7B8 was assembled in pBS (Stratagene) from cloned fragments of the rat FN gene (17). Essentially, a blunt-ended XmnI (within exon III7b) to ApaI (within exon III8a) fragment was subcloned into blunt-ended BamHI and HincII sites. The central portion of IVS1 was removed as a 1.1 kb BstXI-BstEII fragment and the center of IVS2 was removed as a 0.8 kb BalI-XbaI fragment (Fig. 1A). A few additional nucleotides are present at each junction, as a result of intermediate manipulations. Linearization of 7B8 for transcription was at a PstI site within the vector sequences immediately following exon III8a and the protruding 3' end produced by PstI was removed enzymatically. The resulting T7 RNA polymerase transcript was 929 nucleotides long. The 7iB DNA template was generated by cleaving 7B8 with

Figure 1. Construction of 7B8 minigene and processing in HeLa cells. A. The EIIIB region of the rat FN gene (not to scale) and the two alternate processing pathways. Lines, introns; boxes, exons; hatched box, EIIIB; numbers are lengths in nucleotides. 7B8 was generated as described in Materials and Methods. The sizes of the resulting truncated introns IVS1 and IVS2 are indicated. 5' and 3' exons include vector-derived polylinker sequences present in the *in vitro* synthesized transcript. B. Northern blots of RNAs isolated from HeLa cells transfected with the 7B8 fragment in expression vector pLEN (lanes 1) or from control transfected HeLa cells (lanes 2). Duplicate filters were hybridized to probes specific for EIIIB or for exons III7b-III8a, as indicated. The B⁻ band migrated at ca. 500 nucleotides (markers not shown), close to the expected size of 475 nucleotides for the joining of III7b and III8a. The position predicted for EIIIB-containing RNA is indicated as B⁺; pre, position of unprocessed 1.2 kb precursor. The larger species seen in both lanes of panel C is due to non-specific hybridization of this probe to 18S RNA. Panel B was exposed over 10 times longer than Panel C, indicating that inclusion of EIIIB is rare. C. RNase protection analysis of 7B8-derived transcripts. The probe was complementary to an EIIIB- cDNA, spanning the III7b-III8a junction. Lanes 1 and 3 contained RNA from two independent transfections with 7B8 DNA; lane 2 is untransfected HeLa cell RNA and lanes 4 and 5 are controls of probe alone with (lane 4) or without (lane 5) RNase treatment. P, undigested probe; B⁻, protection of III7b-III8a junction; B⁺, position expected for protected III7b exon alone, as would result from EIIIB⁺ (or unprocessed) RNAs. The B⁻ species is running as a pair of bands; the large species (B⁻) is due to incomplete digestion of polylinker sequences flanking III7b. Predicted mobilities were assigned relative molecular size standards (not shown).

BamHI (within the EIIIB exon) and PstI (after exon III8a) and substituting a 1.2 kb fragment of rat FN cDNA extending from the BamHI site within EIIIB to a downstream PstI. Transcription templates were generated by linearizing this plasmid with BamHI, EagI, RsrII, or BglII (Fig. 3C).

For expression in eukaryotic cells, the insert of 7B8 was excised with EcoRI and PstI which removes the FN insert with a small amount of flanking vector sequence and SalI linkers were added. A derivative of the expression vector pLEN (a gift of Tyler White, California Biotechnology Co.; 32) with an XhoI site inserted (33) was linearized with XhoI and ligated to 7B8 insert. The vector contributes about 70 nucleotides of 5' sequence from the human metallothionein gene and about 100 nucleotides of 3' untranslated sequence from the human growth hormone gene to the resulting transcript.

Cell Culture

HeLa S3 cells (obtained from the laboratory of P. A. Sharp) were maintained in spinner culture in Joklik's MEM plus 5% horse serum for preparation of nuclear extract. For transfection, cells were maintained as adherent cells in DME plus 5% horse serum.

Transfection and RNA analyses

HeLa cells were plated at about 20% confluence on the day before transfection. Plasmid DNA was precipitated with calcium phosphate (34), and RNA was isolated 48 hours post-transfection by guanidinium isothiocyanate lysis and centrifugation through a CsCl cushion (35).

For Northern blot analysis, total cellular RNAs were separated through a 1% agarose gel containing 6% formaldehyde, 50 mM Hepes, 5 mM EDTA; molecular size markers (Boehringer Mannheim) were electrophoresed in parallel. RNA was transferred in 20×SSC to nitrocellulose (0.2 μm; Schleicher and Schuell). Pre-hybridization was at 42° in 25% formamide using a commercially prepared Northern pre-hybridization mix (5 prime-3 prime, Inc). Hybridization was at the same temperature using hybridization mix from the same manufacturer plus 25% formamide. Probes were prepared by randomly primed DNA synthesis using purified cDNA fragments containing either EIIIB or III7b-III8a. Following overnight hybridization, filters were washed in 2×SSC, 0.2% SDS at 52°, then autoradiographed with an intensifying screen.

For RNase protection studies, total cellular RNA was treated with DNase prior to hybridization with uniformly labelled RNA probes. An RNA probe complementary to exons III7b-III8a was prepared by *in vitro* transcription (see below) of the corresponding cDNA (17) with SP6 polymerase, and hybridizations were performed as described (21). Unhybridized RNA was digested with 20 μg/ml RNaseA, 1 μg/ml RNase T1 at 30°C for 30 minutes. Nucleases were inactivated by proteinase K treatment and phenol extraction and recovered by precipitation with ethanol. Electrophoresis was through a 6% polyacrylamide denaturing gel.

In vitro transcription

Linearized DNA templates were phenol-extracted, precipitated with ethanol, and resuspended at approximately 1 μg/μl in water. Transcription was with T7 RNA polymerase (Promega Biotech) in the presence of 0.5 mM capping nucleotide (New England Biolabs), 0.5mM all four NTPs, 100 μCi α-[³²P]-UTP (800 Ci/mmol) and 1–2 μg DNA/50 μl reaction. Reactions were incubated 1 hr at 37°, then extracted with phenol and precipitated with ethanol prior to electrophoresis through a denaturing

polyacrylamide gel. Transcripts were located in the gel by autoradiography, and eluted from gel slices in 0.5% SDS, 0.3M NaHOAc, 1.0 mM EDTA.

In vitro splicing reactions

Nuclear extract was prepared from HeLa cells essentially as described (36), but MgCl₂ was omitted from Buffer D. Immediately before use, extract was thawed on ice, clarified by brief centrifugation, then pre-incubated for 10–20 minutes at 30° before use. Eluted transcripts were precipitated with ethanol and resuspended in water at a concentration of 0.1–0.2 pmol/μl.

Standard splicing conditions included 10 μl of nuclear extract and ca. 0.2 pmol labelled precursor RNA in a 25 μl reaction volume containing 3.0 mM MgCl₂, 5.0 mM creatine phosphate and 1.6 mM ATP (except where specified). In other experiments, nuclear extract was increased to 11–15 μl/25 μl reaction, or polyvinyl alcohol was included at 2% final concentration. These variations did not greatly alter the pattern or efficiency of RNA processing. Samples were incubated at 30°C for the indicated times, then digested with an equal volume of 2× proteinase K mix (2× proteinase K mix = 50mM Tris, pH 7.5, 100 mM NaCl, 10 mM EDTA, 1% SDS, 3mg/ml Proteinase K) for 10 minutes. 0.3 ml of 0.3 M Na acetate, 1mM EDTA, 10 μg/ml yeast tRNA was added, extracted with phenol/chloroform, and the RNA was recovered by ethanol precipitation. RNAs were analyzed by denaturing polyacrylamide gel electrophoresis, at the indicated acrylamide concentration (acrylamide:bis = 20:1). Dried gels were exposed at –80°C to Kodak XAR-5 film with an intensifying screen.

Debranching and Primer extension analyses

Splicing reactions and electrophoresis were performed as above; RNAs were located by autoradiography of the wet gel and recovered by elution and ethanol precipitation in the presence of carrier RNA. Debranching reactions were performed as described (37). RNAs were again recovered by precipitation with ethanol, and re-electrophoresed through denaturing polyacrylamide, alongside untreated lariats.

Labelled primer was prepared by incubation in the presence of γ-³²P-ATP and T4 polynucleotide kinase (35). Primer (2 pmol) was annealed to the isolated lariats or to debranched lariats by heating to 70°C for 3 minutes followed by incubation at 30°C for several hours in 50mM Tris-HCl, pH 8.5, 8 mM MgCl₂, 30 mM KCl, 0.01mM DTT. AMV reverse transcriptase (Boehringer Mannheim) and dNTPs were then added and incubated at 42° for 1 hour. Reactions were heated to 95°C for 5 minutes, RNase-treated, precipitated and analyzed by denaturing polyacrylamide gel electrophoresis. Sequencing reactions were performed using 7iB DNA as template, labelled primer and Sequenase (United States Biochemical Co.).

DNA amplification and sequencing

0.5 μg of chicken genomic DNA was amplified (38) for 40 cycles (92°, 2 minutes; 50°, 2 minutes; 72°, 8 minutes) by Taq DNA polymerase (Cetus-Perkin Elmer) using a forward primer corresponding to III7b (5'-TGGTGTGGAGTACAATGT-CAG-3') and a reverse primer complementary to EIIIB (5'-GTGATGCGGTACCCAATAATG-3'), based on chicken FN cDNA sequence (21). The single amplified product was purified from an agarose gel, cut with HindIII (which cuts within EIIIB) and XbaI (which cuts within the intron), and subcloned

Figure 2. Processing of 7B8 RNA *in vitro*. A. Radiolabelled 7B8 pre-RNA was incubated (11 ml extract/25 ml reaction) for the indicated times (hours) under standard splicing conditions plus 2% polyvinyl alcohol. RNAs were separated by electrophoresis through a denaturing 4% polyacryl-amide gel. Lane -, control reaction without ATP and creatine phosphate (3 hours). Identification of the bands is discussed in the text; conventions are as in Figure 1A. In addition to the lariat forms, there is a broad, slowly migrating ATP-independent band, which we presume is unrelated to splicing. In the left panel, lanes a and c contain the isolated lariats and lanes b and d, the debranched lariats. The marks to the left show the position of molecular weight markers; from the top: 948, 832, 623, 528, 405, 310, 243, 239, 218, 202, 191, 181, and 161 nt. B. 7B8 RNA was incubated for 3 hours under conditions as in A, with 3 mM $MgCl_2$ (lane 3) or 5 mM $MgCl_2$ (lane 5). Electrophoresis was through a denaturing 5% polyacrylamide gel to resolve the additional lariat species. Curved bracket indicates a triplet of lariats identified as intermediates in IVS1 processing (see text). Asterisk, as in A, predicted mobility of ligated three exon product.

into pGEM-3. Chicken and rat sequence was determined by the dideoxy chain termination method (39).

Native gel electrophoresis

Splicing reactions were incubated as usual, then placed on ice to stop the reaction. 5 μ l of each sample was immediately loaded onto a Tris-glycine 4% polyacrylamide gel (40); electrophoresis was for 4–5 hours at 15 V/cm at room temperature.

RESULTS

An Optional Exon is Excluded *In Vivo* and *In Vitro*

Our studies have focused on EIIIB, an exon which can be included or skipped, as diagrammed in Fig. 1A. Transfected HeLa cells largely exclude EIIIB when processing transcripts derived from 'mini-genes' containing the human exon and its flanking introns

and exons (30) and the corresponding fragment of the rat gene is similarly processed (G. Huh and R.O.H., unpublished data). However, as the introns flanking EIIIB are each over 1 kb (17, and our unpublished data), we deleted most of the two introns, leaving about 100 nucleotides adjacent to each splice site; this 'mini-gene' is termed 7B8 (Fig. 1A). To test the effect of these deletions on EIIIB splicing, the 7B8 fragment was inserted into the expression vector pLEN (see Materials and Methods). HeLa cells were transfected with this plasmid, and RNA was prepared 48 hours later and analyzed via Northern blotting (Fig. 1B). The majority of the plasmid-derived transcripts failed to hybridize with the EIIIB probe, and represented joining of exons III7b to III8a (B^-). A low level of EIIIB-containing product was detected (B^+), but the majority of EIIIB-containing RNAs appear to be unprocessed precursor (pre). The fidelity of III7b-III8a joining was demonstrated by protection of a ca. 190

Figure 3. Structure and *in vitro* processing of 7iB RNAs. A. Transcripts B, E, R, and G (see section C) were incubated for 3 hours in the presence of 15 ml nuclear extract/ 25 ml reaction at the indicated concentration of MgCl₂. Electrophoresis was through a denaturing 4% polyacrylamide gel. Pre-RNA processing occurs at 5 mM MgCl₂, as indicated by free 5' exon intermediate (same size for all four precursors) as well as ligated III7b-EIIB product (arrowheads; size varies with the length of the precursor). In addition, in panel G lane 5 a faint set of bands was observed (bracket), probably representing free IVS1 lariats. Also in this lane, the slowly migrating lariat intermediates are visible (curved bracket). Inset, a shorter exposure of part of panel G; arrowhead indicates spliced product. B. Duplicate samples to those in 4B were electrophoresed through a denaturing 6% gel; only the 5 mM reactions are shown. Curved bracket indicates the G lariat triplet, and similar sets of bands migrating more slowly than each of the precursors can be seen in the other lanes. C. The construction of 7iB is described in Materials and Methods; conventions are as in Figure 1A. 7iB was linearized at the indicated restriction sites to produce transcripts of 369 (H), 473 (B), 617 (E), 657 (R), and 1,179 (G) nucleotides in length. Transcripts are referred to by these letters. Splicing of the H transcript is not shown; it was not detectably processed, probably because the short length of the 3' exon decreased IVS1 splicing efficiency.

nucleotide probe fragment derived from a cDNA corresponding to EIIB⁻ mRNA (Fig 1C, lanes 1 and 3). Use of a probe specific for EIIB joined to III8a also failed to reveal any significant amount of EIIB⁺ RNA (data not shown). Thus, the ability of HeLa cells to skip EIIB is preserved in the absence of most of its flanking introns.

When the corresponding 7B8 precursor RNA was synthesized and processed *in vitro*, we also detected EIIB skipping. Several species accumulated in an ATP-dependent fashion (Fig. 2A). The lower molecular weight species corresponded to the predicted mobilities of free 5' exon and joined exon 1 and exon 3, as diagrammed. In addition, we detected two species of reduced mobility, indicating lariat forms. Their identities were confirmed by debranching the isolated RNAs (37); lanes b and d show that each species was converted to a single linear form. The mobilities

of these debranched RNAs correspond precisely to the lariat intermediate and product expected for the EIIB skipping reaction. There was no evidence for forms related to EIIB inclusion (495 nt, asterisk), although less abundant species might have been obscured by background. Thus, the dominant spliced product seen *in vitro* results from exon skipping, as does the major *in vivo* product. It should be noted that the central exon of a pre-mRNA derived from a constitutively processed portion of the FN precursor RNA is always included (data not shown), indicating that exon skipping is related to EIIB.

Ionic conditions influence EIIB skipping

As other workers have reported that altering the conditions of the *in vitro* reaction could change patterns of alternative splicing (43,44), we tested whether altering cation concentration could

promote EIIIB inclusion. Reducing Mg^{++} concentration to 2 mM slightly decreased 7B8 processing efficiency, with the skipping reaction the only one detected, and further reduction to 1 mM resulted in essentially no processing of any FN precursor (data not shown). However, raising Mg^{++} concentration to 5 mM revealed additional species (Fig 2B). A doublet of reduced mobility relative to the precursor corresponded to the lariats from the EIIIB skipping event, and these forms were less abundant when reactions were performed at 5 mM Mg^{++} . A slightly more rapidly migrating group of three bands was prominent at the higher Mg^{++} level. The same lariats were detected at low abundance in 3 mM Mg^{++} reactions when the RNAs are separated on a 5% polyacrylamide gel (Fig. 2B) instead of 4%, as in Fig. 2A.

The joined III7b-III8a product which results from EIIIB skipping was readily detected in the 3 mM reactions, but was not observed in those performed at increased Mg^{++} concentration. In addition, the ratio of lariat intermediate to product was different, suggesting that the second step of splicing was less efficient at higher Mg^{++} . Background prevented unambiguous detection of the predicted product, III7b joined to EIIIB-IVS2-III8a (midway between asterisk and precursor). Thus, increasing Mg^{++} concentration from 3 to 5 mM decreased the efficiency of EIIIB skipping, and increased the efficiency of an alternative reaction. The reaction seen at the higher cation concentration could relate to processing of either IVS1 or IVS2. However, the lariat triplet correlated with the appearance of free 5' exon III7b, suggesting that the triplet results from processing of IVS1.

IVS1 is recognized in a Mg^{++} dependent fashion

We wished to test the ability of the extract to remove IVS1 from a simpler pre-RNA containing only a single intron. The second intron was replaced by a fragment of rat fibronectin cDNA generating the plasmid 7iB (Fig 3C). A series of transcripts extending through and beyond EIIIB was derived by linearization of the template DNA at the indicated restriction sites. Each transcript was incubated with extract at the indicated Mg^{++} concentration. While no processing was detected at 2 or 3 mM Mg^{++} (Fig. 3A, lanes 2 and 3), incubation in the presence of 5 mM Mg^{++} resulted in the appearance of bands representing free 5' exon intermediate and joined III7b-EIIIB product (lanes 5). The ligated exon product from the G precursor is more readily seen on a shorter exposure (arrowhead in inset). IVS1 processing was more efficient with larger precursors; however, we have not established whether this effect is simply due to lengthening of the 3' exon, or is specific for the sequences inserted. The H precursor, with only 27 nucleotides of EIIIB, was not detectably processed under any condition (data not shown). Thus, IVS1 processing occurs in a Mg^{++} -sensitive fashion, and is not dependent on sequences downstream of EIIIB.

When identical samples were run on a higher percentage polyacrylamide gel several lariats with reduced mobility relative to the unprocessed precursor were detected (Fig 3B). The pattern of these lariats was remarkably similar to the triplet seen with 7B8 (Fig 2B), suggesting that these bands result from processing of IVS1. As the mobility of the triplet correlated with the length of the RNA precursor (and with the length of the 3' exon), it seemed likely that these species were lariat intermediates, containing both IVS1 and 3' exon sequences. A set of faint bands (Fig 3A, panel G, lane 5, square bracket) were not detected on

a higher percentage gel (see lane G in Fig. 3B). It is likely that these smaller RNAs represent the lariat products corresponding to the more prominent intermediate forms.

IVS1 is a truncated version of the wild type intron. The IVS1 5' splice site functions at lower Mg^{++} , as it participates in the skipping reaction seen with 7B8. It was more likely that the cation-sensitivity involved the 3' splice site, where only 84 nucleotides of intron has been retained. Reinsertion of 400 nucleotides upstream of this splice site did not alter the Mg^{++} sensitivity or increase the efficiency of IVS1 removal (data not shown). Thus, the Mg^{++} sensitivity was probably not a consequence of the IVS1 deletion but was a property of sequences near the 3' splice site; this result focused attention on the region upstream of EIIIB.

Lariat heterogeneity is due to selection of multiple, distant branch sites

The cluster of three lariats which appeared when precursors were processed at high Mg^{++} could reflect several types of heterogeneity. Since the free 5' exon appeared as a single species of the predicted size, 5' splice site cleavage was probably correct and invariant. Heterogeneity could result from the use of multiple branch points, or from multiple sites of 3' splice site cleavage. The E precursor was chosen for further analysis. The putative III7b-EIIIB product was excised from a gel, and served as template for cDNA synthesis, using a primer complementary to EIIIB. This material was then subjected to amplification by PCR (38) using the same EIIIB-related primer and one corresponding to a region of III7b. A species of ca. 200 nucleotides was purified, and direct sequencing demonstrated precise III7b-EIIIB joining (data not shown). Thus, both 5' and 3' splice site cleavages occurred correctly.

Primer extension analysis using an oligonucleotide complementary to EIIIB and the isolated set of lariats as templates identified three reverse transcriptase stop sites (Fig. 4, lane 2). These bands were not detected when unprocessed precursor RNA was used as template (not shown) or when lariats were debranched prior to primer annealing (lane 1). In the latter case, a single species appears, with mobility corresponding to linear intron-3' exon (286 nt). Based on this experiment and others using different primers, we have identified three A residues used as branch points during the processing of IVS1. All lie rather far upstream of the 3' splice site (62, 70, and 76 nucleotides 5' of the cleavage site), and only the furthest upstream corresponds well with the branch site consensus. However, there are no other good branch candidates proximal to EIIIB, apart from a segment which lies immediately adjacent to the 3' splice site (A at -9, see Figure 5), and is presumably too close to be utilized.

In addition to the absence of a candidate branch point adenine at the usual position, there is an exceptionally long pyrimidine rich stretch extending nearly to the 3' splice site (Fig. 5, top line). This unusual sequence prompted a phylogenetic comparison of this region from other vertebrate FN genes. The sequence from this region of the human gene has been published (29), and we obtained sequence from the relevant region of the chicken gene. The intron was amplified using primers based on the III7b and EIIIB exons. A fragment of the 1.2 kb amplification product was subcloned and the sequence immediately upstream of the EIIIB exon was determined. The ca. 100 nucleotides upstream of EIIIB from the three vertebrate species are aligned in Fig. 5. The long pyrimidine-rich stretch is conserved, as is the fairly A-rich 3'

Figure 4. Primer extension mapping of IVS1 branch points. The E RNA derived from 7iB was processed for 2.5 hours as in the 5mM Mg^{++} reactions in Figure 3. The lariat triplet was gel purified, and half was subjected to *in vitro* debranching. Both lariats and debranched RNAs were hybridized to a 5' end-labelled oligonucleotide complementary to EIIB, and primers were extended with reverse transcriptase. Extension products were separated through a denaturing 6% polyacrylamide gel, with sequencing tracks loaded in parallel. Product derived from the debranched material (arrowhead) was sized relative to MspI cut pBR322 DNA (lines to left of gel represent 310, 243, 241, and 221 nts); the sequence of the region of IVS1 including the deduced branch points (asterisks) is shown to the right (this sequence is the complement of the sequence read directly from the gel). Each stop site is seen as a doublet, due to heterogeneity during transcription (data not shown; also see 41). The branch points were assigned as one nucleotide 5' of the reverse transcriptase stop sites (66).

splice site. In addition, all three of the utilized branch points are conserved in the human gene; two remain in the chicken gene. The conservation of these residues is consistent with their being functional in all three species.

Figure 5. Phylogenetic comparison of sequences from the branch site/ 3' splice site region upstream of EIIB. The sequence of this intron from the rat gene has been published (29), the sequence from the human FN gene has been published (29), the sequence from the rat gene was obtained from subclones of rat genomic DNA clones (17), and the sequence from the chicken gene was obtained from polymerase chain amplification of the corresponding intron from chicken genomic DNA. Sequences were aligned using the program Bestfit (67). Numbers refer to the rat sequence; -1, last nucleotide of the intron. Overlined sequences have been deleted from the truncated IVS1 of 7B8 and 7iB. The mapped branch points of IVS1 are indicated (asterisks).

Mg^{++} -sensitive complexes form on IVS1-containing RNA

As the utilization of the EIIB 3' splice site and/or branch sites appeared to be Mg^{++} -sensitive, we wished to determine whether this was due to a failure to assemble pre-RNA into splicing complexes at low Mg^{++} , or to an inability to perform the appropriate cleavage reactions. The E precursor was incubated under splicing conditions in either 3 mM or 5 mM Mg^{++} for 90 minutes. Samples were assessed for the presence of splicing complexes by electrophoresis through a 4% non-denaturing polyacrylamide gel. Parallel lanes included a duplicate reaction at 5 mM Mg^{++} in the absence of ATP, as well as naked pre-RNA (Fig. 6). Incubation in nuclear extract reduced the mobility of all the input pre-RNA relative to naked RNA (lane RNA; 45). Migration of the 'non-specific' complex was somewhat different in the presence or absence of ATP, but did not differ with Mg^{++} concentration (data not shown). An abundant, ATP-dependent complex forms at 5 mM Mg^{++} (arrow) which is not detected at 3 mM Mg^{++} . This species resembles an ATP-dependent complex which forms rapidly on an adenovirus precursor at either Mg^{++} concentration (unpublished data). It thus appears that the Mg^{++} -sensitive step in the splicing of IVS1 involves an early step of complex assembly.

DISCUSSION

Alternative splicing events can be reproduced *in vitro* by manipulating divalent cation concentration

A rat FN precursor RNA consisting of three exons and two introns which are constitutively spliced *in vivo* is similarly processed *in vitro*, with no detectable skipping of the central exon (data not shown). In contrast, when exon 2 of the precursor is EIIB, which is largely skipped in HeLa cells (Figure 1), the exon was also skipped *in vitro* (Figure 2). Under standard splicing conditions (i.e., 2–3 mM Mg^{++}) EIIB was largely excluded but, at higher Mg^{++} , exon skipping was decreased. The decline could be due to competition with the IVS1 reaction or to decreased recognition of the 3' splice site of the III8a exon. Deletion of EIIB and its flanking intron regions from 7B8 does not noticeably improve the efficiency of III7b-III8a joining (data not shown) suggesting that competition alone cannot explain the favored pairing.

Figure 6. Native gel electrophoresis of splicing reactions. The E RNA derived from 7iB was incubated for 90 minutes under the indicated conditions in the presence of 11 ml nuclear extract/ 25 ml reaction. 5 ml of each sample was loaded directly onto a Tris-glycine gel. Approximately twice as much naked RNA was loaded in parallel. Wide arrow indicates the ATP-dependent, Mg^{++} -sensitive complex.

Analyses of certain other alternatively spliced exons have also shown that splicing pattern was dependent on cation concentration. Usage of the three alternative 5' splice sites of the adenovirus E1A gene could be manipulated *in vitro* by changes in K^+ or Mg^{++} (43). Additionally, one of the mutually exclusive exons of rat tropomyosin1 has been found to be included in a cation-sensitive fashion (44). Ionic changes could modulate either RNA secondary structure or RNA-factor interactions. We have carried out sequence analyses in an attempt to predict stable secondary structures in the vicinity of the E1iB 3' splice site or branch site, but no single structure has been strongly predicted (unpublished data).

The Mg^{++} -sensitive appearance of RNA processing intermediates and products could be due either to a block in pre-spliceosome assembly or to a defect in RNA cleavage. Native gel electrophoresis demonstrated that a complex which migrates more slowly than the non-specific complex fails to form at the lower Mg^{++} concentration. This complex probably represents complex A (following the nomenclature of Konarska and Sharp; 40,45), based on kinetics of appearance and mobility relative to complexes formed by a similarly sized adenovirus precursor (unpublished results). However, as the putative pre-splicing complex detected at higher Mg^{++} is much more abundant than the level of splicing intermediates cleaved, later steps in IVS1 processing may also be inefficient. Other precursors which are very poorly processed *in vitro* in fact undergo efficient assembly into pre-splicing complexes (46–50), suggesting that individual

introns might evade the splicing machinery by different mechanisms. We are currently investigating the protein and snRNP components of the Mg^{++} -dependent complex which forms on IVS1.

We detected no evidence for IVS2 removal during 7B8 processing, although a very inefficient reaction could have been missed. Using a precursor extending from the center of E1iB to III8a, we have intermittently detected a low level of processing of IVS2 (not shown), which may be extract-specific. Alternative 5' splice sites of the immunoglobulin k light chain gene were not selected properly *in vitro* by HeLa extract (51,52) and it has been suggested that improper splice site selection could be due to leakage of factors during extract preparation (53). We are attempting to determine whether failure to process IVS2 is due to a defect in the HeLa extract.

Branch site selection and models of alternative splicing

The sequences immediately flanking E1iB conform reasonably well to consensus splice site motifs (1–3). However, comparison of the IVS1 sequences upstream of E1iB in three vertebrate fibronectin genes indicated that this region was very well conserved, and that all three introns lacked conventionally placed branch point sites. This suggested that an unusual site or sites might be used and, indeed, IVS1 processing is associated with the use of three distant branch points. Only the furthest upstream closely resembles a 'good' branch site, YURAC, as defined by mutagenesis studies (47,54). *In vitro* studies of alternative splice sites have uncovered other instances where there are multiple and/or far upstream branch sites. The pattern of branch site utilization in the SV40 early intron varies for large T versus small t splicing (55); moreover, mutations in branch sites can influence the choice of 5' splice site (56). Three branch sites located over 50 nucleotides from the 3' splice site are used in an alternative processing reaction of the adenovirus E1A gene (57), and this is apparently facilitated by the presence of a stable hairpin between the branch points and the 3' splice site (58). As mentioned above, we have been unable to predict such a potential structure within IVS1. A branch site located over 100 nucleotides upstream of an alternative exon of α -tropomyosin has been identified (49) and three similarly distant sites have been mapped within an intron of the β -tropomyosin gene (59). In both the latter cases, the branch points are followed by a lengthy polypyrimidine tract, and the UAG at the 3' splice site is the first downstream occurrence of the 'YAG' motif commonly found at 3' splice sites (1–3). Similar sequence features are present in IVS1, and may participate in splicing regulation. It is anticipated that signals necessary for the alternative splicing of E1iB will be phylogenetically conserved, as is the pattern of splicing regulation and we are currently testing the role of the conserved sequence features.

As other workers have noted, branch site selection is an attractive mechanism for regulated splicing, as U2 recognition of this site seems to be an early step in spliceosome formation (45,46,60–63). E1iB exon inclusion could thus be determined by tissue-specific factors which mediate the recognition of the 'poor' IVS1 branch sites. If this is the case, elevated cation levels apparently circumvent the factor requirement, possibly by altering the specificity of the splicing apparatus. Alternatively, the IVS1 branch points may serve as targets for negative regulation. We cannot currently distinguish between these two possibilities, but the latter would require that HeLa cells (which do not transcribe the fibronectin gene to a significant extent; 27) contain the

inhibitory factor(s). Also, it should be noted that IVS1 removal does not require either the presence or the prior removal of IVS2, at least *in vitro*. This is not easily explained by a regulatory mechanism based solely on competition between the two 3' splice sites or branch sites.

As IVS1, EIIB and IVS2 are removed from 7B8 as a single intron, exon skipping requires that both the 5' and 3' splice sites flanking EIIB are ignored. It has been suggested that in an intron is not committed to splicing until the 3' splice site has been synthesized (65). If the IVS1 3' splice site or branch site were recognized prior to the completion of IVS2 synthesis, then the transcript could be committed to IVS1 removal and EIIB inclusion. If, however, recognition of the 3' end of IVS1 was inefficient (or negatively regulated), synthesis of the IIIa 3' splice site would occur, and this site could pair with either 5' splice site. Recognition of the EIIB 5' splice site could thus 'define' the exon (66), permitting the subsequent removal of IVS1. Thus, one model for regulation of EIIB splicing includes 5' splice site choice by cell-specific factors.

ACKNOWLEDGEMENTS

The authors are grateful to P. Grabowski, A. Lamond and P. Sharp for being generous with advice and reagents during the early phases of this work. We would like to thank E. George, G. Huh and M. Garcia-Blanco for helpful criticisms of the manuscript. Technical assistance was capably provided by D. Fonda, S. Francis, M. Jalenak and R. Patel-King.

This work has been supported by a grant from the USPHS National Cancer Institute (CA26712). P.A. Norton was a post-doctoral fellow of the National Cancer Institute (CA08027) and R.O. Hynes is an investigator of the Howard Hughes Medical Institute.

REFERENCES

- Green, M.R. (1986) *Ann. Rev. Genet.* **20**:671-708.
- Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) *Ann. Rev. Biochem.* **55**:1119-1150.
- Krainer, A.R. and Maniatis, T. (1988) In Hames, B.D. and Glover, D.M. (eds.), *Transcription and Splicing*. IRL Press, Oxford, pp.131-206.
- Dreyfuss, G. (1986) *Ann. Rev. Cell Biol.* **2**:459-498.
- Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Kramer, A., Frendewey, D. and Keller, W. (1988) In Birnstiel, M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, Springer-Verlag, Berlin, pp.115-154.
- Grabowski, P.J., Seiler, S.R. and Sharp, P.A. (1985) *Cell* **42**:345-353.
- Frendewey, D. and Keller, W. (1985) *Cell* **42**:355-367.
- Brody, E. and Abelson, J. (1985) *Science* **228**:963-967.
- Leff, S.E., Rosenfeld, M.G. and Evans, R.M. (1986) *Ann. Rev. Biochem.* **55**:1091-1117.
- Andreadis, A., Gallego, M. and Nadal-Ginard, B. (1987) *Ann. Rev. Cell Biol.* **3**:207-242.
- Solnick, D. and Lee, S.I. (1987) *Mol. Cell. Biol.* **7**:3194-3198.
- Eperon, L.P., Graham, I.R., Griffiths, A.D. and Eperon, I.C. (1988) *Cell* **54**:393-401.
- Cooper, T.A., Cardone, M.H. and Ordahl, C.P. (1988) *Nucleic Acids Res.* **16**:8443-8465.
- Bandziulis, R.J., Swanson, M.S. and Dreyfuss, G. (1989) *Genes Dev.* **3**:431-437.
- Mattaj, I.W. and Hamm, J. (1989) *Development* **105**:183-189.
- Patel, R.S., Odermatt, E., Schwarzbauer, J.E. and Hynes, R.O. (1987) *EMBO J.* **6**:2565-2572.
- Schwarzbauer, J.E., Patel, R.S., Fonda, D., and Hynes, R.O. (1987) *EMBO J.* **6**:2573-2580.
- Kornblihtt, A.R., Vibe-Pedersen, K. and Baralle, F.E. (1984) *Nucleic Acids Res.* **12**:5853-5868.
- Sekiguchi, K., Klos, A.M., Kurachi, K., Yoshitake, S. and Hakomori, S. (1986) *Biochemistry* **25**:4936-4941.
- Gutman, A. and Kornblihtt, A.R. (1987) *Proc. Natl. Acad. Sci. USA* **84**:7179-7182.
- Norton, P.A. and Hynes, R.O. (1987) *Mol. Cell Biol.* **7**:4297-4307.
- Hynes, R.O. (1985) *Ann. Rev. Cell Biol.* **1**:67-90.
- Owens, R.J., Kornblihtt, A.R. and Baralle, F.E. (1986) *Oxf. Surv. Eucaryot. Genes* **3**:141-160.
- Kubomura, S., Obara, M., Karasaki, Y., Taniguchi, H., Gotoh, S., Tsuda, T., Higashi, K., Ohsato, K., Hirano, H. (1987) *Biochim. Biophys. Acta* **910**:171-181.
- French-Constant, C. and Hynes, R.O. (1989) *Development* **106**:375-388.
- Schwarzbauer, J.E., Mulligan, R.C. and Hynes, R.O. (1987) *Proc. Natl. Acad. Sci. USA* **84**:754-758.
- Vibe-Pedersen, K., Kornblihtt, A.R., and Petersen, T.E. (1984) *EMBO J.* **3**:2511-2516.
- Mardon, H.J., Sebastio, G. and Baralle, F.E. (1987) *Nucleic Acids Res.* **15**:7725-7733.
- Paoletta, G., Henchcliffe, C., Sebastio, G. and Baralle, F.E. (1988) *Nucleic Acids Res.* **16**:3545-3557.
- Barone, M.V., Henchcliffe, C., Baralle, F.E. and Paoletta, G. (1989) *EMBO J.* **8**:1079-1085.
- French-Constant, C. and Hynes, R.O. (1988) *Development* **104**:369-382.
- Neufeld, G., Mitchell, R., Ponte, P. and Gospodarowicz, D. (1988) *J. Cell Biol.* **106**:1385-1394.
- Guan, J.-L., Trevithick, J.E. and Hynes, R.O. (1990) *J. Cell Biol.* **110**:833-847.
- Graham, F.L. and van der Eb, A.J. (1973) *Virology* **52**:456-461.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**:1475-1489.
- Ruskin, B. and Green, M.R. (1985) *Science* **229**:135-140.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science* **239**:487-491.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
- Konarska, M.M. and Sharp, P.A. (1987) *Cell* **49**:763-774.
- Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984) *Cell* **37**:415-427.
- Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) *Cell* **38**:317-331.
- Schmitt, P., Gattoni, R., Keohavong, P. and Stevenin, J. (1987) *Cell* **50**:31-39.
- Helfman, D., Ricci, W.M. and Finn, L.A. (1988) *Genes Dev.* **2**:1627-1638.
- Konarska, M.M. and Sharp, P.A. (1986) *Cell* **46**:845-855.
- Lamond, A.I., Konarska, M.M. and Sharp, P.A. (1988) *Genes Dev.* **1**:532.
- Reed, R. and Maniatis, T. (1988) *Genes Dev.* **2**:1268-1276.
- Nelson, K.K. and Green, M.R. (1988) *Genes Dev.* **2**:319-329.
- Smith, C.W.J. and Nadal-Ginard, B. (1989) *Cell* **56**:749-758.
- Agris, C.H., Nemeroff, M.E. and Krug, M.E. (1989) *Mol. Cell. Biol.* **9**:259-267.
- Lowery, D.E. and VanNess, B.G. (1987) *Mol. Cell. Biol.* **7**:1346-1351.
- Kedes, D.H. and Steitz, J.A. (1987) *Proc. Natl. Acad. Sci. USA* **84**:7928-7932.
- Kedes, D.H. and Steitz, J.A. (1988) *Genes Dev.* **2**:1448-1459.
- Zhuang, Y., Goldstein, A.M. and Weiner, A.M. (1989) *Proc. Natl. Acad. Sci. USA* **86**:2752-2756.
- Noble, J.C.S., Pan, Z.-Q., Prives, C. and Manley, J.L. (1987) *Cell* **50**:227-236.
- Noble, J.C.S., Prives, C. and Manley, J.L. (1988) *Genes Dev.* **2**:1460-1475.
- Gattoni, R., Schmitt, P. and Stevenin, J. (1988) *Nucleic Acids Res.* **16**:2389-2409.
- Cheblis, K., Gattoni, R., Schmitt, P., Hildwein, G. and Stevenin, J. (1989) *Mol. Cell. Biol.* **9**:4852-4861.
- Helfman, D.M. and Ricci, W.M. (1989) *Nucleic Acids Res.* **17**:5633-5650.
- Grabowski, P.J. and P.A. Sharp. (1986) *Science* **233**:1294-1299.
- Bindereif, A. and Green, M.R. (1987) *EMBO J.* **6**:2415-2424.
- Zillman, M., Zapp, M.L. and Berget, S.M. (1988) *Mol. Cell. Biol.* **8**:814-821.
- Kramer, A. (1988) *Genes Dev.* **2**:1155-1167.
- Beyer, A.L. and Osheim, Y.N. (1988) *Genes Dev.* **2**:754-765.
- Robberson, B.L., Cote, G.J. and Berget, S.M. (1990) *Mol. Cell. Biol.* **10**:84-94.
- Jacquier, A. and Rosbash, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**:5835-5839.
- Devereux, J., Haeblerli, P. and Smithies, O. (1984) *Nucl. Acids Res.* **12**:387-395.