

Polypyrimidine tract sequences direct selection of alternative branch sites and influence protein binding

Pamela A. Norton*

Department of Medicine, Division of Gastroenterology, Roger Williams Medical Center and Brown University, Providence, RI 02908, USA

Received July 11, 1994; Accepted August 15, 1994

ABSTRACT

IVS1, an intron derived from the rat fibronectin gene, is spliced inefficiently *in vitro*, involving the use of three alternative branch sites. Mutation of one branch point site, BP3, so as to increase complementarity to U2 snRNA resulted in exclusive use of that site and improved splicing efficiency, indicating that the wild type BP3 site is one determinant of poor IVS1 splicing. Deletions within the polypyrimidine tract had a variable effect on splicing efficiency and altered the pattern of branch site usage. Selection of each branch site was influenced negatively by purine substitutions ca. 20 nucleotides downstream. It is proposed that all three IVS1 branch sites are pyrimidine tract-dependent. Pyrimidine tract deletions also influenced the crosslinking of PTB (the polypyrimidine tract-binding protein), hnRNP C, and splicing factor U2AF65. All three proteins bound preferentially to distinct regions within the polypyrimidine tract and thus are candidates for mediating pyrimidine tract-dependent branch site selection. The findings indicate the complexity of the IVS1 polypyrimidine tract and suggest a crucial role for this region in modulating branch site selection and IVS1 splicing.

INTRODUCTION

Removal of introns from pre-mRNA involves two catalytic steps (1, 2). The products of the first reaction are intermediates consisting of a free 5' exon and an intron-3' exon species with a lariat structure. The lariat is formed by covalent linkage of the 5' phosphate of the residue at the newly cleaved end of the intron to a 2'-hydroxyl group of a residue which lies near the 3' end of the intron. This latter residue is referred to as the branch point site, as the linkage results in a branched nucleotide. In the second step, the lariat intermediate is cleaved, with concomitant ligation of the two exons and liberation of the free lariat intron.

The mechanism whereby the site of branch formation is chosen has been subjected to experimental scrutiny. Considerable

evidence points to base pairing of the sequences which surround the branch site with a portion of U2 snRNA (3–5, 6). The sequence UACUAAC is complementary to a region of U2 snRNA, with the actual site of the branch (bold) bulging out. In *Saccharomyces cerevisiae*, the UACUAAC consensus is highly conserved; in contrast, although many vertebrate branch sites conform to this consensus, significant deviations exist. Since complementarity to U2 correlates positively with splicing efficiency (4, 5, 7), pairing between U2 and the pre-mRNA is thought to be as important for splicing of vertebrate as well as yeast introns. Most vertebrate branch sites fall within 18–40 nucleotides of the 3' splice site; however, sites which lie > 40 nucleotides upstream of the 3' splice site have been characterized (8–14). Thus, sequence and positional requirements are somewhat flexible for vertebrate branch point sites.

Branch site selection can be influenced by several *trans*-acting factors. As described above, the direct interaction between U2 snRNP and the branch site consensus is well established. Splicing factor U2AF was identified as an activity which promotes the interaction between U2 and the branch site (15). The 65 kDa subunit of U2AF (U2 snRNP auxiliary factor) contacts the pre-mRNA directly, with highest affinity for pyrimidine-rich sequences (16, 17). In addition, it appears that snRNP U1, which interacts with the 5' splice site, may play an important role in promoting stable U2/pre-mRNA interactions (18, 19); the former study implicated splicing factor SC35 as well. Other activities appear to be required for efficient formation of a U2-pre-mRNA complex (20, 21).

Hundreds of eukaryotic pre-mRNAs generate multiple mRNAs by alternative splicing, and the synthesis of many splice variants is regulated in a cell or tissue-specific fashion (22–24). The sequences of many alternative splice sites resemble those of constitutive sites. However, a number of alternative exons and introns have suboptimal splicing signals (reviewed in [25]), which may serve to prevent constitutive processing. Several alternative introns contain distant and/or multiple branch point sites (8–10, 12, 14, 26, 27). The distant sites are generally associated with extensive polypyrimidine tracts that can play an important role in the the selection of the alternative exons.

*Present address: Department of Medicine, Division of Gastroenterology and Hepatology, Thomas Jefferson University, 1025 Walnut Street, Philadelphia, PA 19107, USA

One of the features involved in alternative splicing of the fibronectin gene exon EIIIB is the region of the 3' splice site/branch point site region of the intron upstream of EIIIB, IVS1 (28). A truncated version of IVS1 is functional *in vivo* (28), and this intron is processed in a conditional fashion *in vitro* by HeLa nuclear factors (8). A pre-mRNA, 7iB, containing this intron was spliced only at high (5–6 mM) Mg⁺⁺, with little processing seen at low (2–3 mM) Mg⁺⁺ concentration. However, even under permissive conditions, splicing was very inefficient. IVS1 splicing involves the use of one of three branch sites, located at –62, –70 and –76 relative to the 3' splice site; all three sites are utilized with approximately equal efficiency. A long polypyrimidine-rich tract (PPT) lies immediately downstream of the branch sites. This study investigates the roles of the branch sites and polypyrimidine tract in rendering IVS1 a poor splicing substrate, and explores the features responsible for branch site selection and conditional splicing.

MATERIALS AND METHODS

Structure and preparation of templates used for *in vitro* transcription

The construction of the 7iB template DNA was described previously (8). Mutant derivatives of 7iB were generated by inverse PCR (29). Oligodeoxyribonucleotide primers were designed in pairs such that their 5' ends abutted the region to be deleted; amplification of the entire plasmid resulted. For smaller substitutions, oligodeoxyribonucleotides incorporated the desired changes. Oligodeoxynucleotides were: 1A, CAAAAA-AAGAGGAAAAATA; 1B, CCTCCCCTTTGCTTCATAA-C; 2A, AAATAAAGCAGTGTACATCA; 2B, CTCTACTT-TTCCTTTGCCCTC; 3A, CAAAGGAAAAGTAGAGCAAAA; 3B, TAACTCAATAGAGGTGCCCC; 4, CCTCCCCTTTGC-TTCATCTCAATAGAGG; 5, CAAAAAAGAGGAAAAAT-TAGCAGTGTAC. Mutants were generated by PCR with Taq DNA polymerase (Perkin-Elmer/Cetus) using primer pairs: d1, 1A and 1B; d2, 2A and 2B; d3, 3A and 3B; d4, 2A and 1B; d5, 1A and 3B; m1, 3A and 4; m2, 2B and 5. Mutant m3 was isolated from the last reaction, and represents a mispriming event. Typically, 30 cycles of 1 minute at 94°C, 2 minutes at 50–55°C, and 5 minutes at 72°C were performed. Amplified products were treated with DNA polymerase (Klenow fragment) and T4 polynucleotide kinase and then ligated. The sequences of all mutants were confirmed by the chain termination method.

Template DNA 7iB was linearized with *Bam*HI, so as to yield a runoff transcript containing IVS1 (184 nucleotides) flanked by 5' exon III7b (154 nucleotides, including polylinker sequence) and 3' exon EIIIB (131 nucleotides). The derivatives of 7iB were treated similarly. BA.2 RNA was generated from a subclone of 7iB (pBA.2). DNA was linearized with *Hind*III to yield a transcript containing 44 nucleotides of polylinker sequence (pGEM-3, Promega Biotec), 88 nucleotides of IVS1 and 27 nucleotides of EIIIB. These transcripts are diagrammed in Fig. 1a.

In vitro transcription

Preparation of 7iB, BA.2 and mutant pre-mRNAs was as described (8). Transcription was with bacteriophage T7 RNA polymerase (Stratagene) in the presence of 50 μ Ci α -³²P-UTP and 0.4–0.5 mM of all four rNTPs, using conditions recommended by the manufacturer. Reactions were routinely performed for 30–60 minutes at 37°C. All labelled pre-mRNAs

were purified on Tris–borate buffered gels of 5% polyacrylamide and 7 M urea to remove non-incorporated radiolabel.

In vitro splicing and analysis of splicing intermediates and products

For *in vitro* splicing, RNAs were incubated with 10 μ l HeLa nuclear extract in the presence of 1.6 mM ATP, 5.0 mM creatine phosphate and the indicated amount of magnesium chloride (8). Splicing reactions (25 μ l) were incubated for 2.5–3.0 hours at 30°C, then terminated by treatment with proteinase K and RNAs recovered by precipitation with ethanol. For direct analysis, RNAs were resuspended in 80% formamide and applied to gels containing 7 M urea and 6–8% polyacrylamide, as indicated. Gels were dried and the RNAs visualized by autoradiography.

Alternatively, precipitated RNAs were resuspended in water and analyzed by RNA–PCR with enzymatic depletion; the details of this procedure have been described previously (30). Briefly, first strand cDNA was synthesized using RNA from splicing reactions as template and an oligonucleotide complementary to EIIIB (chB5: 5'-GTGATGCGGTACCCAATAATG-3') as primer. A portion of each cDNA preparation was amplified by Taq DNA polymerase (Perkin-Elmer/Cetus) for 20 cycles (94°C, 1 minute; 54°C, 1 minute and 72°C, 1 minute). Primers were chB5 and ch7b (5'-TGGTGTGGAGTACAATGTCAG-3'); the latter corresponds to sequences within exon III7b. Half of each amplified sample was digested with *Eco*NI, which cleaves the pre-mRNA once within the intron but leaves the spliced III7b–EIIIB product intact. This step substantially depletes the product derived from pre-mRNA. One tenth of this material was then subjected to an additional 20 cycles of amplification. Amplified products were separated on 2% agarose gels then transferred to nylon and hybridized to a ³²P-labelled probe containing III7b and EIIIB exon sequences. The membrane was washed and exposed to X-ray film.

UV crosslinking and analysis of labelled proteins

Splicing reactions were incubated for 20–30 minutes, then irradiated with 254 nm light (ca. 2500 mW/cm²) for 10 minutes at 0°C. RNAs were digested by addition of 1/10 vol of RNase A (10 mg/ml); incubation was at 30°C for 10 minutes. Electrophoresis sample buffer containing 2% SDS, 0.1 M DTT was added, and the samples heated to 100°C prior to loading. Electrophoresis was through 10% polyacrylamide (29:1, acrylamide:bis-acrylamide) separating gels (31). Gels were fixed, dried and exposed to X-ray film with an intensifying screen.

Isolation of RNase T1 oligonucleotides

For oligoribonucleotide isolation, transcription of 7iB was performed with one-fifth as much unlabelled UTP (0.08 mM) as other NTPs (0.4 mM). The reaction was heated to 90°C, then incubated with 4 μ g of RNase T1 for 30 minutes at 37°C. After extraction with phenol and recovery by ethanol precipitation, the T1-resistant fragments were separated by electrophoresis through a 15% polyacrylamide gel. Following autoradiography of the wet gel, bands were excised and RNAs eluted overnight in 0.3 M sodium acetate, 0.5% SDS. Yeast tRNA (1.0 μ g) was added and RNAs were precipitated with ethanol. One tenth of each recovered fragment was incubated with nuclear extract under splicing conditions at the indicated magnesium concentration for 30 minutes prior to UV crosslinking.

To distinguish the co-migrating oligos3a and 3b, each was produced by *in vitro* transcription using synthetic

oligodeoxynucleotide templates. These templates consisted of sequences complementary to oligos3a or 3b juxtaposed with the T7 RNA polymerase promoter bottom strand sequence; bottom strands were annealed to a short promoter top strand (32). Transcription of the two templates with T7 RNA polymerase resulted in RNAs corresponding to the individual T1 oligonucleotides.

RESULTS

Branch site mutation increases IVS1 splicing

Previous work indicated that the weak, conditional splicing of pre-mRNA 7iB was most likely due to sequences at the 3' end of the intron and/or within the EIIIB exon (8). Several elements within intron IVS1 are likely to influence processing. The three alternative branch sites (8) are indicated in Figure 1 and will be referred to, from 5' to 3', as BP1, BP2 and BP3. Other notable features of this intron include a long polypyrimidine tract and a purine-rich sequence immediately upstream of the 3' splice site. A series of deletion and substitution mutations were generated from the 7iB template DNA, altering one or more of these sequence features. The sequence of the relevant portion of each mutant substrate is shown in Figure 1.

Pre-mRNAs 7iB, m1, m2 and m3 were synthesized *in vitro* then incubated with HeLa nuclear extract under conditions previously established as non-permissive (2 mM Mg⁺⁺) or permissive (6 mM Mg⁺⁺) for IVS1 splicing (8). RNAs were isolated from splicing reactions and subjected to denaturing gel electrophoresis (Fig. 2). 7iB is spliced only at the permissive condition, as expected (lane 2). Deletion of two adenosine residues in mutant m1 reduced the purine content of the sequence immediately upstream of the 3' splice site and eliminated potential complementarity to U2 (UAAC). The m1 deletion had no discernable qualitative or quantitative effects on IVS1 splicing relative to wild type (compare lanes 4 and 2) and was not considered further.

Mutant m2 contained substitutions at BP3 (see Fig. 1) designed to increase similarity to the branch site consensus (UACUAAC); the improved site is referred to as BP3up. Mutant pre-mRNA m2 was spliced more efficiently than 7iB at high Mg⁺⁺,

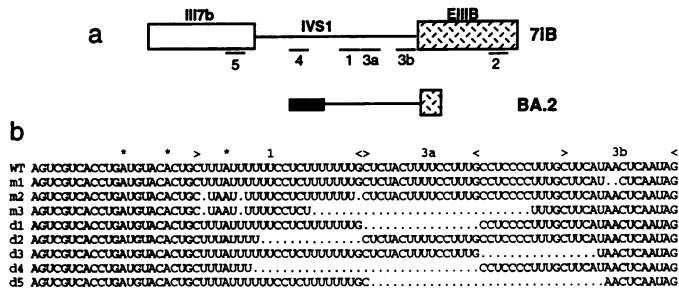


Figure 1. Structure of IVS1-containing pre-mRNAs and sub-fragments. **a.** The structure of pre-mRNA 7iB is diagrammed; line, IVS1; boxes, exons; filled box, polylinker-derived sequence. The locations of the T1-resistant fragments, oligos1–5, are indicated by the numbered bars. Below, RNA BA.2 is diagrammed, with conventions as above. **b.** At top, the sequence of the 3' portion of IVS1 from wild type 7iB is shown (8), truncated at the 3' splice site. The three branch point sites are indicated by asterisks and T1-resistant fragments oligo1, oligo3a and oligo3b are indicated by angle brackets. Below, the sequences of the various deletion and substitution mutants analyzed in this study are shown, with deleted residues indicated by dots.

(compare lanes 6 and 2), the single lariat species suggesting utilization of a single branch site. Primer extension, performed as in (8), yielded a single product 165 nucleotides in length (data not shown), consistent with exclusive use of the BP3up site. This result also identifies the 7iB lariats, from top to bottom, as BP3, BP2 and BP1. Thus, it seemed that BP3 is one determinant of weak IVS1 splicing.

Pre-mRNA m2 was not spliced at 2 mM Mg⁺⁺ (lane 5), but, in contrast to 7iB, a low level of splicing was observed at intermediate cation levels (data not shown). Related substrate m3 was spliced conditionally, but with some increased efficiency relative to 7iB, based on accumulation of spliced product (lane 8; see also Fig. 3b, lanes 7–9). Splicing of m3 also involved use of a single branch site. This site corresponds to the BP3up site, based on mobility relative to lariats derived from similarly sized plasmids (see Fig. 3b and below). Note that m2 lacks a G residue, resulting in a slightly longer PPT than wild type 7iB, whereas m3 has a shorter PPT (Fig. 1). As m2 was spliced more efficiently than m3, it is possible that IVS1 polypyrimidine tract sequences increase splicing efficiency.

Polypyrimidine tract sequences influence splicing efficiency and cation dependence

The role of the polypyrimidine tract in IVS1 was explored by testing the splicing capacity of several substrates with deletions therein. Based on the accumulation of lariat intermediates, pre-mRNAs d1 and d3 were spliced slightly more efficiently than 7iB, whereas there was little evidence for d2 splicing (Fig. 3a, lanes 1–12). (The very low level of splicing of 7iB seen in Fig. 3a is more typical than the higher level observed in Fig. 2; data

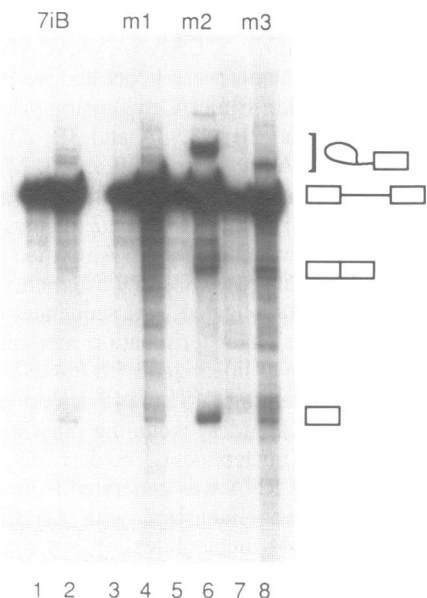


Figure 2. Effect of a branch site mutation on IVS1 splicing. Pre-mRNAs 7iB (lanes 1 and 2), m1 (lanes 3 and 4), m2 (lanes 5 and 6) and m3 (lanes 7 and 8) were incubated with nuclear extract at 2 (odd lanes) or 6 (even lanes) mM Mg⁺⁺. Reactions were terminated and RNAs isolated and analyzed by electrophoresis through 7% polyacrylamide in the presence of urea. The various intermediates and products of 7iB splicing were characterized previously (8); these are indicated by the schematic diagrams to the right. The triplet lariat intermediate triplet represents the use of the three branch sites. A minor lariat of reduced mobility was detected in lanes 6 and 8; this species has not been characterized.

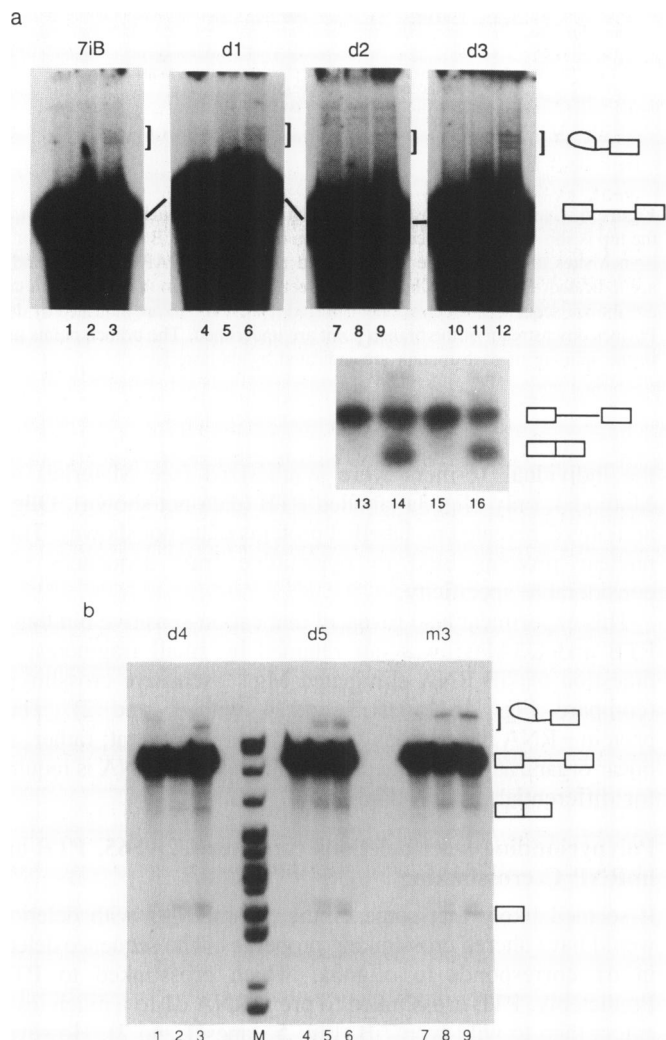


Figure 3. Deletions within the polypyrimidine tract affect splicing efficiency and branch site selection. **a.** At top, RNAs are analysed as in Figure 2, but 8 % polyacrylamide was used. Splicing of 7iB (lanes 1–3), d1 (4–6), d2 (7–9) and d3 (10–12) was performed at 2, 4 or 6 mM Mg⁺⁺ (first, middle and last lane of each set, respectively). Only the lariat intermediates and unspliced pre-mRNA are shown. Unmarked bands of low mobility do not appear to be related to splicing, as they are not cation-dependent. Below, RNA-PCR was performed on RNAs derived from splicing of d2 (lanes 13 and 14) and d3; splicing reactions were incubated with 2 (lanes 13 and 15) or 6 (lanes 14 and 16) mM Mg⁺⁺. Unspliced pre-mRNA and spliced product are indicated. **b.** Splicing of d4 (lanes 1–3), d5 (lanes 4–6) and m3 (7–9) and analyses were performed as above. RNA species are identified as in Fig. 2. Lane M, pBR322 DNA cut with *Msp*I and radiolabelled; from the top, fragments are 623/528 (unresolved), 405, 308, 243/239, 218, 202, 191, 181, 161, 148, 124 and 111.

not shown). Extensive degradation of pre-mRNAs accompanied by poor splicing efficiency precluded direct visualization of other splicing intermediates and products. However, a sensitive RNA-PCR assay for two-exon spliced product indicated that pre-mRNA d2 was spliced (Fig. 3a, lanes 13–16 and data not shown). Although these differences have not been quantitated carefully, it was concluded that the deletions in d1, d2 and d3 did not increase splicing efficiency when compared with 7iB, and did not alter cation-dependence.

The RNAs with larger deletions, d4 and d5, exhibited slightly elevated levels of splicing, but, more importantly, altered cation

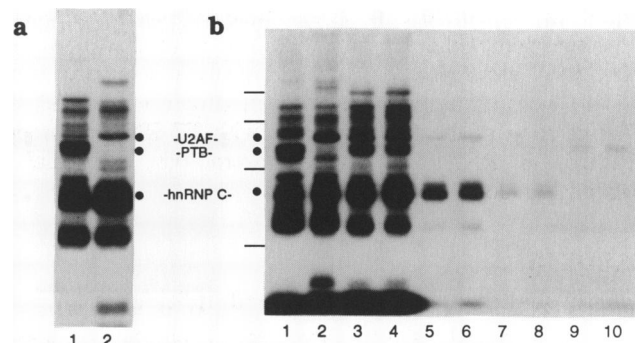


Figure 4. Localization of sites of protein crosslinking within the polypyrimidine tract. **a.** Nuclear proteins were incubated with BA.2 RNA at 2 (lane 1) or 6 (lane 2) mM Mg⁺⁺. The bound proteins were UV crosslinked to the labelled RNA. Proteins were separated by SDS-PAGE and visualized by autoradiography of the dried gel. Proteins PTB and U2AF65 (U2AF) crosslink preferentially at 2 or 6 mM Mg⁺⁺, respectively, where as hnRNP C is observed under both conditions. The identification of these species was described previously (31). **b.** Nuclear proteins were incubated with 7iB RNA (lanes 1 and 2), T1-digested 7iB (lanes 3 and 4), or T1-resistant oligonucleotides oligo1 (lanes 5 and 6), oligo2 (lanes 7 and 8) and oligo3 (lanes 9 and 10) at 2 (odd lanes) or 6 (even lanes) mM Mg⁺⁺. Dashes indicate positions of molecular size markers of apparent M_r 100, 68, 43 and 27×10^3 .

dependence; significant levels of IVS1 splicing were observed at 4 mM Mg⁺⁺ (Fig. 3b, lanes 2 and 5, respectively; note that lane 2 is somewhat underloaded). However, prolonged exposures failed to reveal any splicing of either substrate at 2 mM (Fig. 3b and data not shown). Wild type 7iB RNA undergoes little splicing below 5–6 mM Mg⁺⁺ (Fig. 3a and [8]). Mutant m3 also is active for splicing at 4 and 6, but not 2 mM Mg⁺⁺ (Fig. 3b, lanes 7–9). These results suggest that both polypyrimidine tract and branch point sequences contribute to the weak, conditional splicing of IVS1.

Polypyrimidine tract deletions alter branch site selection

Unexpectedly, some of the deletion mutants exhibited altered patterns of branch site selection. BP3 was not utilized in the splicing of RNAs d4 and d5; in the splicing of the former, BP2 was also ignored, with BP1 strongly preferred (Fig. 3b, lanes 1–3). Either sequences that lie at a distance influence branch site selection or the deletions bring BP3 too close to the 3' splice site. The latter possibility seems less likely, as the BP3up site was utilized despite a similarly sized deletion (m3, lanes 7–9). Interestingly, the region of overlap of these three deletions is precisely deleted in mutant d1, which exhibits the wild type pattern of branch site selection (Fig. 3b). Thus, these particular sequences do not specify branch site selection.

It was observed that splicing of mutant d2 involved BP3 but not BP1 or 2 (Fig 3a, lane 9). However, the signal is weak when compared to the level of splice product detected by RNA-PCR, which may indicate that other, novel sites are being used. Nevertheless, the failure to use BP1 and BP2 also suggests that alterations in sequences that lie at a distance from a branch site can influence its selection.

Nuclear proteins crosslink specifically within the IVS1 polypyrimidine tract

Using biochemical and immunological criteria, we showed that proposed splicing factors hnRNP C (33), PTB (polypyrimidine

Table 1. Similarity to branch point site consensus

Branch site	Sequence*	UNCURAC	UACUAAC	UACUAAC #
BP1	ACCUGAU	5/7	3/7	5/7
BP2	UGUACAC	4/7	3/7	5/7
BP3	GCUUUAU	3/7	2/7	4/7
BP3up	UGCUAU	6/7	5/7	7/7

* The branched residue is in bold.

G-U pairs are counted as matches.

interaction between branch site selection and polypyrimidine tract sequences. Parallel studies revealed that nuclear factors interacted specifically with sequence elements within the polypyrimidine tract. The significance of these findings is considered below.

Pyrimidine tract sequences and branch site selection

Several studies have established that branch sites with complementarity to U2 snRNA are optimal; perfect complementarity is represented by the sequence UACUAAC, with the branch A (bold) bulged out. However, a broader consensus, UNCURAC, has been established by mutagenesis (37). The match of the three IVS1 branch sites to these two consensus sequences is shown in Table 1; such comparisons predict that BP1 should be preferred, which was not observed. Thus, sequence context must be important, and it seems likely that selection of BP3 occurs as a consequence of its favorable location, despite poor complementarity to U2. Mutation of BP3 to conform more closely to consensus resulted in improved splicing of IVS1, with exclusive use of the improved site, suggests that BP3 contributes to poor IVS1 splicing. However, pre-mRNA m2 remains a fairly inefficient splicing substrate in our hands when compared to other, non-IVS1-containing pre-mRNAs (data not shown). Thus, it is likely that other features of pre-mRNA 7iB contribute to inefficient processing in addition to BP3.

A role for the pyrimidine tract in pre-mRNA splicing has been demonstrated for a number of natural and model introns (reviewed in [38]). In light of this requirement, complete deletion of the IVS1 polypyrimidine tract (IVS1 PPT) was not performed, as it seemed likely to yield non-functional pre-mRNA. However, analyses of pre-mRNAs with deletions within the PPT revealed unexpected complexity. The effects of the deletions can be due either to removal of specific sequence elements or to the juxtaposition of novel sequences at critical locations. The latter explanation best accounts for the different behaviors of d1, d4, d5 with regard to branch site selection. The loss of branch site utilization correlates with the substitution of purines at various positions (summarized in Fig. 6). Use of wild-type BP3 seems to depend on the presence of pyrimidines in the oligo3a sequence, the site of PTB crosslinking. In contrast, use of BP1 and BP2 is decreased when purines are introduced into the oligo1 sequence, the site of U2AF65 crosslinking. Interestingly, the strong BP3up branch site is not as potent in the context of pre-mRNA m3 compared with m2; note that purine substitutions occur ≥ 14 nucleotides downstream. Thus, selection of individual branch sites appears to be dependent on polypyrimidine sequences at some distance from the branch sites themselves.

A negative influence of purines within a polypyrimidine tract is not unexpected. The introduction of purines immediately downstream of the branch site has been shown to be more detrimental than similar substitutions close to the 3' splice site (13). Multiple purine substitutions tend to depress splicing and

single purine substitutions could be deleterious when approximately 14 nucleotides downstream of the branch point site (39). The data presented here extend these observations by the demonstration that sequences rather far from a branch point site (ca. 20 nucleotides) can influence its selection. Furthermore, the data point to the existence of functionally distinct regions within the IVS1 PPT (see below), which influence the selection of the individual branch sites.

Several introns with multiple branch sites have been described (9, 12, 27); their use may also be pyrimidine tract-dependent. In the splicing of the adenovirus E1A gene, various deletion and substitutions within the polypyrimidine tract influenced the selection of three branch sites (40). Evidence was presented for a stem-loop that shortened the distance between the branch sites (at -59, -55, and -51) and the 3' splice site. However, the primary sequence of this region may influence branch site selection as well. Interestingly, substitution of purines within the polypyrimidine tract in intron 1 of the adenovirus major late transcript resulted in the activation of a cryptic branch site (41); these authors also suggested that pyrimidine tract sequences influence branch site selection.

Unlike the E1A gene, the IVS1 sequence between BP3 and the 3' splice site does not favor the formation of a stem-loop structure. However, the distance between the branch sites and the influential pyrimidine tract sequences may indicate a need for the RNA to form some higher order structure, perhaps mediated by nuclear factors. The loss of cation-sensitive crosslinking with RNase T1 digestion is consistent with the need for precise spatial organization of two or more RNA elements. In summary, the data reveal a complex dependence of branch site selection, as well as IVS1 cation sensitivity, on polypyrimidine tract sequences. The findings may have implications for multiple branch sites occurring in other introns.

Pyrimidine tract-protein interactions

The existence of functional sub-regions of the IVS1 PPT crosslink is also indicated by the ability of different sequence elements to crosslink to distinct nuclear proteins. It is attractive to speculate that pyrimidine tract-dependence is due to protein-RNA interactions in this region, and indeed, several deletions within the IVS1 polypyrimidine tract significantly alter patterns of nuclear protein crosslinking. Substantial changes in PTB and hnRNP C crosslinking were observed, with more subtle alterations in U2AF65; all three are thus candidates for mediating pyrimidine tract-dependent branch site selection. However, no simple correlation can be drawn between the crosslinking changes and branch site selection. Selection of BP3 appears dependent upon sequences within the region where PTB binds; however, as PTB crosslinking was only observed under conditions non-permissive for splicing, PTB seems unlikely to be involved directly in branch site selection. Both d2 and d4 exhibited reduced hnRNP C crosslinking, with very different patterns of branch site selection. None of the deletions eliminated U2AF65 crosslinking.

PTB was initially proposed to be an essential splicing factor (34, 35), but more recent data argue against such a role (21, 42). A series of mutations involving an alternatively spliced rat β -tropomyosin intron has demonstrated that the polypyrimidine tract plays a role in the negative regulation of the intron (43). Recently, PTB has been shown to bind to this negative regulatory element (44), suggesting that the protein may have a negative effect on splicing. We proposed that PTB could act as a negative regulator

of IVS1, because the protein crosslinked under conditions non-permissive for splicing (31). In contrast, U2AF65 crosslinked under permissive conditions and was a candidate positive factor. These observations, coupled with the finding that U2AF65 and PTB crosslink near to each other, suggested a working model in which binding of PTB to IVS1 under the non-permissive condition prevents U2AF65 binding. However, eliminating PTB crosslinking (mutants d4 and d5) did not activate U2AF65 crosslinking at 2 mM Mg⁺⁺. This result is not the outcome predicted by the model (although it does suggest that PTB is non-essential). However, deletions that eliminate PTB crosslinking may affect the interaction of U2AF65 with the RNA; this is likely to be true for d4 (see below). Thus, the failure of U2AF65 to crosslink to d4 and d5 at low Mg⁺⁺ does not exclude PTB/U2AF65 interactions. It is interesting to note that reciprocal crosslinking of PTB and U2AF65 has been observed for a preprotachykinin gene-derived RNA (45). Alternatively, other proteins may be acting to prevent U2AF65 crosslinking under non-permissive condition.

The correlation between U2AF crosslinking and splicing remains strong, suggesting that U2AF is an essential factor for IVS1 splicing, but it seems likely that the protein can interact with more than one site on the pre-mRNA. Despite direct evidence for the interaction of U2AF65 with oligo 1, the deletion in d2, which removed 14 of 23 bases, reduced the labelling of this protein only slightly. The remaining nucleotides could represent the site of interaction, or binding could occur at another site. In favor of the second possibility, the d4 deletion results in a substantial decrease in U2AF65 crosslinking, suggesting involvement of sequences in addition to those removed by d2. The data presented are consistent with a model in which the precise site of U2AF binding influences branch site selection, by promoting or stabilizing U2 interactions at the individual sites. Detailed mapping of protein-pre-mRNA interactions will be needed to establish the validity of this scenario. It is also of interest to evaluate the role, if any, of these interactions in the selection of alternative exon EIIIB.

ACKNOWLEDGEMENTS

I am indebted to Renee McInnis for skilled technical assistance and to Bruce Kuo for helpful comments on the manuscript. This work was supported by the American Cancer Society (IRG-45-33 to Brown University) and by the Department of Medicine and the Division of Gastroenterology of Roger Williams Hospital.

REFERENCES

- Green, M.R., 1986, *Annu. Rev. Genet.*, **20**, p. 671–708.
- Sharp, P.A., 1987, *Science*, **235**, p. 766–771.
- Parker, R., Siliciano, P.G. and Guthrie, C., 1987, *Cell*, **49**, p. 229–239.
- Wu, J. and Manley, J.L., 1989, *Genes Dev.*, **3**, p. 1553–1561.
- Zhuang, Y., Goldstein, A.M. and Weiner, A.M., 1989, *Proceedings of the National Academy of Sciences (USA)*, **86**, p. 2752–2756.
- Black, D.L., Chabot, B. and Steitz, J.A., 1985, *Cell*, **42**, p. 737–750.
- Reed, R. and Maniatis, T., 1988, *Genes Dev.*, **2**, p. 1268–1276.
- Norton, P.A. and Hynes, R.O., 1990, *Nucl. Acids Res.*, **18**, p. 4089–4097.
- Gattoni, R., Schmitt, P. and Stevenin, J., 1988, *Nucl. Acids Res.*, **16**, p. 2389–2409.
- Goux-Pelletan, M., L., D., d'Aubenton-Carafa, Y., Fiszman, M., Brody, E. and Marie, J., 1990, *EMBO J.*, **9**, p. 241–249.
- Graham, I.R., Hamshere, M. and Eperon, I.C., 1992, *Mol. Cell. Biol.*, **12**, p. 3872–3882.
- Helfman, D.M. and Ricci, W.M., 1989, *Nucl. Acids Res.*, **17**, p. 5633–5650.
- Reed, R., 1989, *Genes Dev.*, **3**, p. 2113–2123.
- Smith, C.W.J. and Nadal-Ginard, B., 1989, *Cell*, **56**, p. 749–758.
- Ruskin, B., Zamore, P.D. and Green, M.R., 1988, *Cell*, **52**, p. 207–219.
- Valcárcel, J., Singh, R., Zamore, P.D. and Green, M.R., 1993, *Nature*, **362**, p. 171–175.
- Zamore, P.D., Patton, J.G. and Green, M.R., 1992, *Nature*, **355**, p. 609–614.
- Fu, X.-D. and Maniatis, T., 1992, *Proc. Natl. Acad. Sci. USA*, **89**, p. 1725–1729.
- Barabino, S.M.L., Blencowe, B.J., Ryder, U., Sproat, B.S. and Lamond, A.I., 1990, *Cell*, **63**, p. 293–302.
- Krämer, A. and Utans, U., 1991, *EMBO J.*, **10**, p. 1503–1509.
- Patton, J.G., Porro, E.B., Galceran, J., Tempst, P. and Nadal-Ginard, B., 1993, *Genes Dev.*, **7**, p. 393–406.
- Maniatis, T., 1991, *Science*, **251**, p. 33–34.
- McKeown, M., 1992, *Annu. Rev. Cell Biol.*, **8**, p. 133–55.
- Smith, C.W.J., Patton, J.G. and Nadal-Ginard, B., 1989, *Annu. Rev. Genet.*, **23**, p. 527–577.
- Norton, P.A., 1994, *J. Cell Sci.*, **107**, p. 1–7.
- Hartmuth, K. and Barta, A., 1988, *Mol. Cell. Biol.*, **8**, p. 2011–2020.
- Noble, J.C.S., Pan, Z.-Q., Prives, C. and Manley, J.L., 1987, *Cell*, **50**, p. 227–236.
- Huh, G.S. and Hynes, R.O., 1993, *Mol. Cell. Biol.*, **13**, p. 5301–5314.
- Ochman, H., Gerber, A.S. and Hartl, D.L., 1988, *Genetics*, **120**, p. 621–623.
- Norton, P.A. and McInnis, R., 1993, *BioTechniques*, **14**, p. 954–957.
- Norton, P.A. and Hynes, R.O., 1993, *Biochem. Biophys. Res. Comm.*, **195**, p. 215–221.
- Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C., 1987, *Nucl. Acids Res.*, **15**, p. 8783–9001.
- Choi, Y.D., Grabowski, P.J., Sharp, P.A. and Dreyfuss, G., 1986, *Science*, **231**, p. 1534–1539.
- García-Blanco, M.A., Jamison, S.F. and Sharp, P.A., 1989, *Genes Dev.*, **3**, p. 1874–1886.
- Patton, J.G., Mayer, S.A., Tempst, P. and Nadal-Ginard, B., 1991, *Genes Dev.*, **5**, p. 1237–1251.
- Zamore, P.D. and Green, M.R., 1989, *Proc. Natl. Acad. Sci. USA*, **86**, p. 9243–9247.
- Green, M.R., 1991, *Annu. Rev. Cell Biol.*, **7**, p. 559–599.
- Morris, D.R., Kakegawa, T., Kaspar, R.L. and White, M.W., 1993, *Biochem.*, **32**, p. 2931–2937.
- Roscigno, R.F., Weiner, M. and García-Blanco, M.A., 1993, *J. Biol. Chem.*, **268**, p. 11222–11229.
- Chebli, K., Gattoni, R., Schmitt, P., Hildwein, G. and Stevenin, J., 1989, *Mol. Cell. Biol.*, **9**, p. 4852–4861.
- Freyer, G.A., O'Brien, J.P. and Hurwitz, J., 1989, *J. Biol. Chem.*, **264**, p. 14631–14637.
- Jamison, S.F., Crow, A. and García-Blanco, M.A., 1992, *Mol. Cell. Biol.*, **12**, p. 4279–4287.
- Guo, W., Mulligan, G.J., Wormsley, S. and Helfman, D.M., 1991, *Genes Dev.*, **5**, p. 2096–2107.
- Mulligan, G.J., Guo, W., Wormsley, S. and Helfman, D.M., 1992, *J. Biol. Chem.*, **267**, p. 25480–25487.
- Hoffman, B.E. and Grabowski, P.J., 1992, *Genes Dev.*, **6**, p. 2554–2568.