

Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing

James G. Patton, Stephen A. Mayer, Paul Tempst,¹ and Bernardo Nadal-Ginard

Laboratory of Molecular and Cellular Cardiology, Howard Hughes Medical Institute, Department of Cardiology, Children's Hospital, Department of Molecular and Cellular Physiology, Harvard Medical School, Boston, Massachusetts 02115 USA

α -Tropomyosin exons 2 and 3 are spliced in a mutually exclusive manner. Exon 3 is included as the default exon in the mRNA of most cell types, whereas exon 2 is only included in the mRNA of smooth muscle cells. The primary determinant for the default selection of exon 3 is the branchpoint/polypyrimidine tract. This element upstream of exon 3 clearly and effectively outcompetes the corresponding element upstream of exon 2. To identify *trans*-acting factors that bind to this important *cis* element, we used UV cross-linking to identify a 57-kD protein whose binding characteristics directly correlate with 3'-splice-site selection in *cis*-competition splicing assays. This protein appears to be identical to polypyrimidine tract-binding protein. In this report we have used oligonucleotides derived from peptide sequences to isolate and sequence cDNA clones encoding this 57.2-kD protein. The primary sequence reveals a novel protein with significant homology to other RNA-binding proteins. Expression of the mRNA is detected in all tissues and cells examined, although its levels exhibit tissue-specific and developmental regulation. Using a biochemical complementation assay, we have found that this protein, along with a 100-kD protein, exists as part of a large complex that is required to rescue splicing from depleted nuclear extracts.

[Key Words: Polypyrimidine tract-binding protein; pre-mRNA splicing; α -tropomyosin]

Received March 15, 1991; revised version accepted May 1, 1991.

The splicing of pre-mRNAs occurs via a two-step mechanism and involves the assembly of multiple *trans*-acting factors with the pre-mRNA in a large complex termed the spliceosome (for reviews, see Green 1986; Padgett et al. 1986; Maniatis and Reed 1987; Sharp 1987; Steitz et al. 1988). These *trans*-acting factors recognize conserved elements within the pre-mRNA, including the 5'-splice site, the branchpoint sequence and polypyrimidine tract, and the 3'-splice site AG (Mount 1982; Ohshima and Gotoh 1987; Smith et al. 1989). The best characterized of these *trans*-acting factors is the U family of small nuclear ribonucleoproteins (snRNPs). U1 snRNP has been shown to base-pair with the 5'-splice site (Mount et al. 1983; Kramer et al. 1984; Zhuang and Weiner 1986), U2 base-pairs with the branch site (Black et al. 1985; Parker et al. 1987; Wu and Manley 1989; Zhuang and Weiner 1989), and U5 may bind to the 3'-splice site (Chabot et al. 1985; Gerke and Steitz 1986; Tazi et al. 1986). U4/U6 is also required for splicing, but this snRNP does not appear to bind the pre-mRNA directly (Berget and Robberson 1986; Black and Steitz

1986; Bindereif and Green 1987). The search for mammalian non-snRNP proteins necessary for splicing is now an intense subject of investigation. Two such factors have been definitively identified thus far: U2AF, which is necessary for U2 snRNP binding to the branchpoint sequence (Ruskin et al. 1988; Zamore and Green, 1989, 1991), and ASF/SF-2, which causes alterations in 5'-splice-site choice (Ge and Manley 1990; Krainer et al. 1990a,b). Antibodies against three additional non-snRNP proteins—heterogeneous nuclear RNP (hnRNP) C (Choi et al. 1986), SC-35 (Fu and Maniatis 1990), and an 88-kD protein (Ast et al. 1991)—have also been shown to inhibit *in vitro* splicing reactions, suggesting a necessary role for these proteins. However, complementation of immunodepleted splicing extracts has only been reported for the 88-kD protein thus far (Ast et al. 1991).

Several potential mammalian splicing factors have also been described, many of which appear to interact primarily with the polypyrimidine tract associated with the branchpoint (Gerke and Steitz 1986; Tazi et al. 1986; Swanson and Dreyfuss 1988a; Garcia-Blanco et al. 1989; Wang and Pederson 1990) or other intron sequences (Garcia-Blanco et al. 1990). In addition, multiple splicing activities have been partially characterized (Krainer and

¹Present address: Molecular Biology Program, Sloan Kettering Institute, New York, New York 10021 USA.

Patton et al.

Maniatis 1985; Kramer and Keller 1985; Mayeda et al. 1986; Kramer et al. 1987; Kramer 1988; Pruzan et al. 1990; Utans and Kramer 1990). However, none of these splicing activities has been purified to the point that specific proteins have been identified. Total protein analysis of purified spliceosomes suggests that there are many as yet unidentified spliceosomal proteins (Reed et al. 1988; Reed 1990). The role that any of these proteins play in splice site selection and/or in the splicing process itself, either catalytic or structural, remains unknown.

Previously, we have shown that the branchpoint/polypyrimidine tract is the major determinant in the default splice site selection of α -tropomyosin exon 3; and, in general, this element plays an important role in 3'-splice-site selection (Mullen et al. 1991). To identify proteins that bind to this important *cis* element, we have utilized UV cross-linking to identify and purify a 57-kD protein whose polypyrimidine tract-binding characteristics correlate closely with both splicing efficiency and 3'-splice-site selection for a large number of substrates (Mullen et al. 1991). Here, we report the cloning and sequencing of cDNAs encoding this protein. We show that the expression of the gene for this 57-kD protein is both tissue and developmentally regulated and shows strong evolutionary conservation. In addition, we show that this polypyrimidine-binding protein is a component of the spliceosome. Using an *in vitro* complementation assay, we show that this protein, in association with a 100-kD protein, restores splicing activity to depleted nuclear extracts.

Results

Identification of a 57-kD polypyrimidine-binding protein

α -Tropomyosin (α -TM) is an alternatively spliced gene containing three sets of mutually exclusive exons (Wiczorek et al. 1988). Exons 2 and 3 are one such pair, where exon 3 is the default exon found in the mRNA of most cell types while exon 2 is used only in smooth muscle cells. The default selection of exon 3 has been shown to be due to competition between the branchpoint/polypyrimidine tracts (BP/PPT) located upstream of exon 2 and exon 3. The BP/PPT upstream of exon 3 outcompetes the corresponding element upstream of exon 2 (Mullen et al. 1991). Thus, the strength of the BP/PPT is the major determinant in default splice site selection for these two mutually exclusive exons.

To search for possible factors involved in determining the competition between exons 2 and 3, we incubated transcripts containing these exons, or portions thereof (Fig. 1), in HeLa cell nuclear extracts, which splice in the default pattern, and in nuclear extracts from smooth muscle cells, which undergo the regulated splicing pattern. These reactions were then subjected to UV cross-linking, and proteins containing covalently attached labeled nucleotides were analyzed by gel electrophoresis. As shown in Figure 1A, one major protein that resolves as a doublet of \sim 57 kD was cross-linked in both cell

extracts to RNA substrates that lack a 5'-splice-site and 3'-splice-site AG and contain only a BP/PPT tract. The presence of this protein in nuclear extracts from both cell types suggests that this protein is not a cell-specific regulatory protein but is likely a constitutive RNA-binding protein.

The specificity of cross-linking was shown by effective competition with increasing amounts of self-competitor unlabeled RNA (Fig. 1A) and lack of competition with bacterial RNA (Mullen et al. 1991). In competition experiments with homoribopolymers, poly(U) is the most effective competitor for the binding of this protein followed, in decreasing order, by poly(C), poly(G), and poly(A) (Fig. 1B). These competition results strongly argue that this 57-kD protein binds to the polypyrimidine tract. To test this directly, we utilized RNA substrates containing shortened polypyrimidine tracts. As shown in Figure 1C, decreasing the number of pyrimidines from 51 nucleotides in the wild-type sequence (pGC+DX) to 17 (Δ 60-110) or 9 nucleotides (Δ 52-110) causes a dramatic decrease in cross-linking. In addition, truncation of these transcripts at the *AccI* site, which deletes the 3'-splice-site AG and 3' exon, does not abolish cross-linking (Fig. 1C). Thus, the deletion results and effective competition with homoribopolymers, especially poly(U) RNA, demonstrates that the binding of this protein is strongly dependent on the polypyrimidine tract and that neither a branchpoint sequence nor 3'-splice-site AG or 3' exon is required for binding.

A functional role for the 57-kD protein in pre-mRNA splicing is suggested by the fact that the decrease in cross-linking seen with the polypyrimidine deletion mutants correlates exactly with splicing efficiency and splice site selection. When these same constructs are spliced *in vitro*, splicing efficiency decreases dramatically (Smith et al. 1989; Mullen et al. 1991). This correlation suggests that this protein is involved in splicing and underscores the importance of the polypyrimidine tract in this process (Frendeway and Keller 1985; Ruskin and Green 1985a,b; Konarska and Sharp 1986; Bindereif and Green 1987; Smith and Nadal-Ginard 1989; Smith et al. 1989). However, it is clear that spliceosome assembly is not required for this protein to bind and cross-link to RNA. A construct (pG3X23H) that cannot form 60S spliceosomes because it lacks the spacer sequence necessary to relieve steric hindrance between exons 2 and 3 (Smith and Nadal-Ginard 1989) can be cross-linked as well as a small transcript (B3P3) containing only a branchpoint and polypyrimidine tract (used in Fig. 1A and C). For all of these transcripts, the primary determinant for cross-linking of the 57-kD protein is a polypyrimidine tract, and there is a strong correlation between splicing efficiency and cross-linking of this protein, suggesting that it plays a role in splicing.

Purification, cloning, and primary structure of polypyrimidine tract-binding protein

To further characterize the 57-kD protein, we have utilized the cross-linking assay to follow its purification to

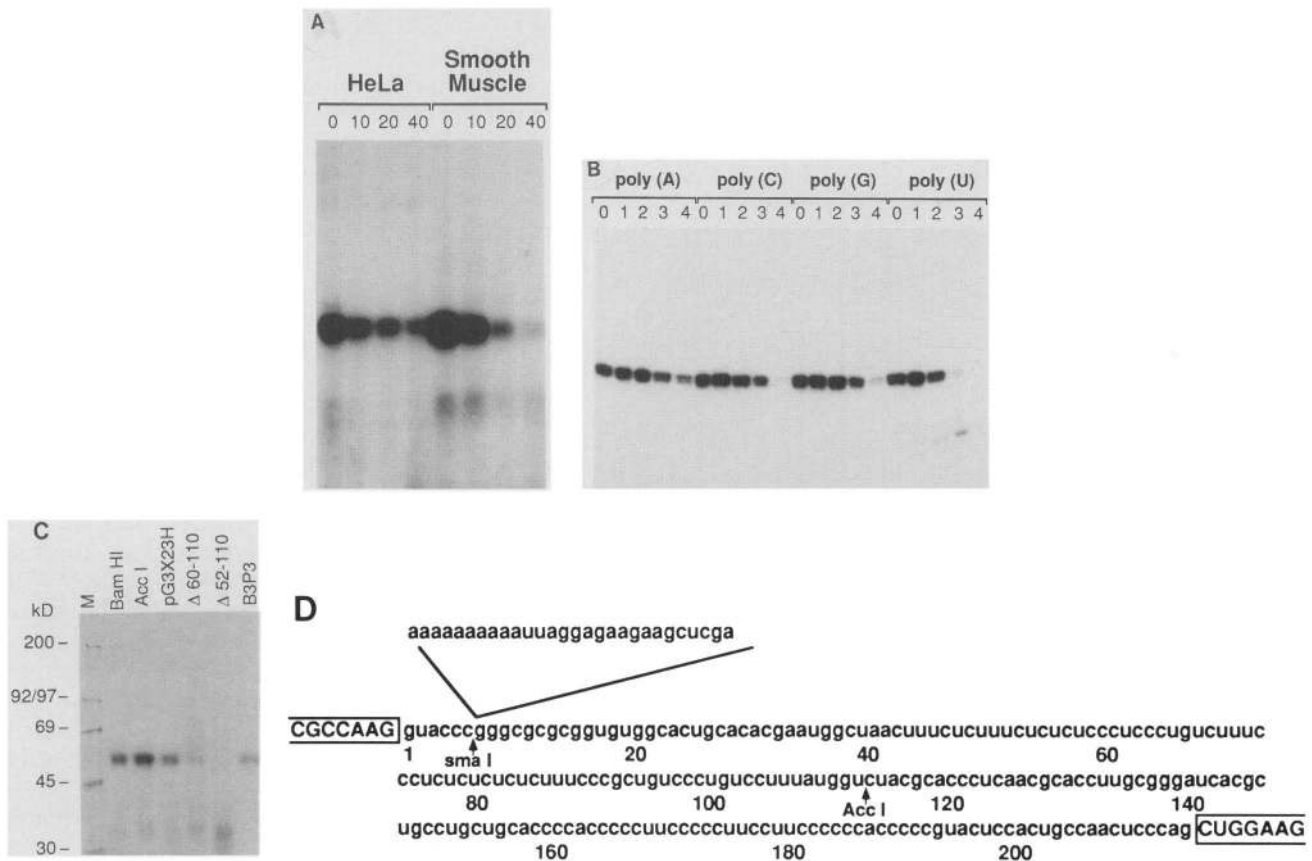


Figure 1. A 57-kD protein specifically cross-links to polypyrimidine tracts. (A) Cross-linking of the 57-kD protein is specific. Binding reactions in HeLa or smooth muscle cell nuclear extracts were prepared containing 1 ng of ^{32}P -labeled B3P3 RNA and the indicated molar excess of unlabeled B3P3 RNA. Binding was allowed to proceed for 1 hr at 30°C and cells were UV-irradiated. Samples were then processed as described in Materials and methods and analyzed on a 9% Laemmli gel. (B) Homoribopolymers compete for the binding of the 57-kD protein. Increasing amounts of homoribopolymers were added to binding reactions as in A. Lanes 1, 2, 3, and 4 refer to 2, 20, 200, and 1000 $\mu\text{g}/\text{ml}$ of homoribopolymer, respectively. (C) The binding of the 57-kD protein requires the polypyrimidine tract. The binding reactions containing equimolar amounts of the indicated RNA were incubated at 30°C for 1 hr, followed by irradiation as in A. The transcripts used have been described (Smith et al. 1989) and are as follows: *Bam*HI refers to full-length linearized template from pG3 + DX, which is a 2-intron splicing construct containing exons 2 and 3 of α -TM, along with a spacer element necessary to relieve steric hindrance (Smith and Nadal-Ginard 1989). *Acc*I refers to linearization of pG3 + DX at the *Acc*I site within the intron and produces a transcript lacking the 3' exon. pG3X23H is identical to pG3 + DX but lacks the inserted spacer element and therefore does not splice. The deletion clones refer to internal deletions between the nucleotides numbered in D and are specific deletions in the polypyrimidine tract, decreasing this tract to 17 nucleotides ($\Delta 60$ -110) or 9 nucleotides ($\Delta 52$ -110). B3P3 is a *Sma*I-*Acc*I fragment (D) from pG3 + DX subcloned into pGem. Markers were from New England Nuclear. (D) Sequence of α -TM intron 2. Boxed regions indicate exons. The spacer element necessary to relieve steric interference (Smith and Nadal-Ginard 1989) is shown above the *Sma*I site. The branchpoint is at position 41 followed by the long 51-nucleotide polypyrimidine tract. The entire splicing construct consisting of exons 2 and 3 plus the intron, as shown, was produced by linearization with *Bam*HI in the polylinker downstream of exon 3. Truncated transcripts were linearized at the *Acc*I site. The downstream polypyrimidine tract adjacent to the 3'-splice site is not utilized even when the upstream tract is deleted (Smith et al. 1989).

apparent homogeneity through several chromatographic steps (Fig. 2A). The purification scheme (see Materials and methods) was similar to that described for polypyrimidine tract-binding protein (PTB) (Garcia-Blanco et al. 1989). Previously, we have referred to this protein as PBP or PBP/PTB (Mullen et al. 1991). However, it appears that PTB and PBP are identical (Gil et al., this issue); therefore, we will refer to this protein as PTB.

Purified PTB was subjected to proteolysis, and several peptide sequences were obtained (Table 1). This protein

always migrates as a doublet and sometimes as a triplet on SDS-PAGE. Amino acid sequence of the corresponding high-performance liquid chromatography (HPLC)-fractionated tryptic peptides from each of the bands confirms that all three contain common peptides (Fig. 2B; Table 1) and, therefore, might represent post-translational modification of the same protein or alternatively spliced isoforms. Using degenerate oligonucleotides deduced from these peptide sequences, we produced by polymerase chain reaction (PCR) a 69-bp DNA fragment

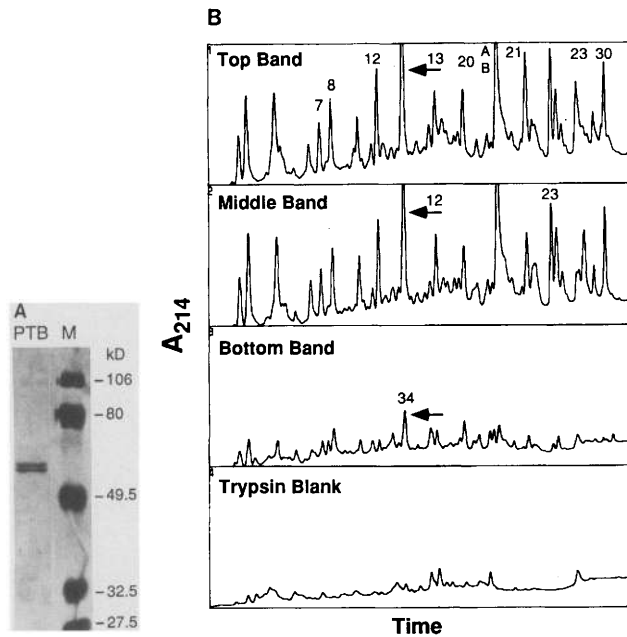


Figure 2. Tryptic digestion of purified PTB. (A) SDS-polyacrylamide gel of purified PTB. The protein was visualized by silver staining. A triplet of bands is visible on some preparations, although a doublet is mainly detected here. (B) Reverse-phase HPLC of tryptic peptides. After electroblotting of purified PTB onto nitrocellulose, peptides resulting from in situ digests of each of the bands were reduced and alkylated (4-vinylpyridine) before injection. The absorbance pattern of each of the bands, as well as a trypsin blank, is shown. (*Top* and *middle*) The two major bands visible on the silver stain in A; (*bottom*) the third band detected in some preparations. Numbered peptides were sequenced. Sequencing of the peak denoted with an arrow shows that all three bands contain a common peptide.

containing the sequence corresponding to peptide TOP T21 (Table 1) and used this cloned DNA to probe a HeLa cDNA library. Several hybridizing clones were isolated. One of these, cDNA clone 121, encodes a 57.2-kD protein that contains all of the tryptic peptides shown in Table 1 but terminates before the poly(A) tail. A separate, overlapping clone (212) contains a 5'-truncated cDNA that continues to the poly(A) tail. These DNA clones have been sequenced at least twice on both strands using multiple deletion clones to generate the sequence shown in Figure 3.

To verify that clones 121 and 212 encode PTB, the protein that cross-links to RNA, we have prepared antibodies against a bacterially expressed fusion protein containing all but the five amino-terminal amino acids fused to glutathione-S-transferase (GST; Smith and Johnson 1988). Western blots of HeLa cell nuclear proteins show that this antiserum strongly recognizes a 57-kD protein (Fig. 4A). To demonstrate the specificity of this antiserum for HeLa PTB, we have immunoprecipitated cross-linked protein from assays using pyrimidine deletion mutants as substrate RNAs. As shown (Fig. 4B), deletion of the polypyrimidine stretch again decreases cross-link-

ing, and the labeled proteins are immunoprecipitable with anti-PTB antibodies but not with preimmune sera. The recovery of the labeled material by immunoprecipitation is not complete, probably because UV cross-linking partially destroys antigenic sites recognized by the antibody. Furthermore, the same *Escherichia coli*-produced fusion protein used to generate the antiserum also cross-links efficiently to RNAs containing a BP/PPT (Fig. 4C and D). Taken together, these results demonstrate that cDNA clones 121 and 212 indeed encode PTB, the protein that cross-links to the polypyrimidine tract in a variety of nuclear extracts.

Sequence analysis of PTB

The primary structure of PTB reveals several interesting details. The longest open reading frame encodes a basic protein (pI 9.95) of 531 amino acids with a predicted size of 57,221 daltons. This sequence contains an unusual stretch of 8 alanines in a row, a common motif found in several developmentally important *Drosophila* loci that may be a linker region between functional domains (Smoller et al. 1990). The predicted secondary structure reveals a protein rich in β -sheet domains and relatively few α -helices. Notably, given the RNA-binding properties of this protein, it contains an unusual RNP-binding domain. The consensus RNP domain is an 80- to 90-amino-acid region, found in a variety of RNA-binding proteins (Bandziulis et al. 1989; Query et al. 1989), that contains a highly conserved octamer (RNP-1) sequence and a less well-conserved hexamer (RNP-2), typically separated by ~30–35 amino acids. The crystal structure of this domain has been determined recently for the U1 A protein (Nagai et al. 1990) and supports the importance of these two RNP boxes. PTB contains a region that

Table 1. Tryptic peptides from PTB

Peptide	Sequence	Position
TOP T7	ILFNK	369
TOP T8	IITFTK	212
TOP T12	XQPIYIQFSNHK	123
TOP T13	LSLDGQNIYNACCTL(R)	239
MID T12	LSLDGQNIYNACCTLR	
BOT T34	LSLDGQNIYNACCTL	
TOP T20A	NNQFQALLQYADPVSQAHA(K)	219
TOP T20B	XF(Q)NIFPPSATL(H)LNI	445
TOP T21	IAIPGLAGAGNSVLLVSNLNP	326
TOP T23	KLPLXVTE	65
MID T23	KLPIDVTEXEVISL	
TOP T230	MALIQMGSVEEAVQALIDL(HN)	493

Sequences are in single-letter code. Residues in parentheses are tentative; (X) residues could not be positively identified. Peptides are arranged according to their retention time on the HPLC chromatogram (Fig. 2B). Amino acid positions are as noted.

Polypyrimidine tract-binding protein



Figure 3. Nucleotide and amino acid sequence of PTB. Underlined regions correspond to the sequences most closely resembling the RNP-1 and RNP-2 boxes.

matches the RNP-1 box at five of eight positions (KGFKFFQK vs. consensus KGFGFVXF; see Fig. 3) and a match to the RNP-2 box at four of six positions (LLVSNL vs. consensus LFGVNL; see Fig. 3), yet the spacing between these elements is 137 amino acids, a significant difference from the consensus sequence.

Searching the GenBank and EMBL data bases for similar proteins has revealed significant homology (55% similarity; 29% identity) to the human hnRNP L protein (Piñol-Roma et al. 1989), as well as the *Drosophila melanogaster* ELAV (embryonic lethal, abnormal visual system) protein (55% similarity; 23% identity) (Robinow et

Patton et al.

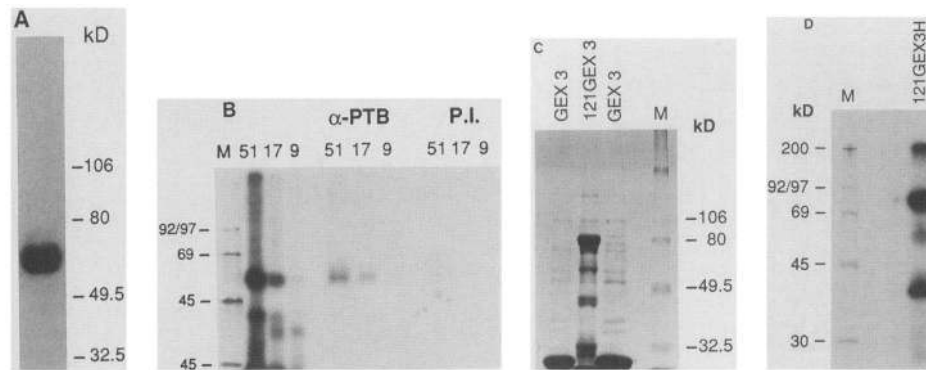


Figure 4. cDNA clones 121 and 212 encode PTB, the protein that cross-links to RNA. (A) Anti-PTB antiserum recognizes HeLa cell PTB. Antibodies were prepared against a bacterially expressed PTB fusion protein and reacted with proteins transferred to nitrocellulose. Visualization of the reacting bands was as described in Materials and methods. (B) Immunoprecipitation of cross-linked PTB. Short RNA substrates containing only a branchpoint, followed by either 51, 17, or 9 pyrimidines, were incubated with nuclear extract for 1 hr at 30°C. Binding reactions were then cross-linked and RNase-digested, followed by the addition of anti-PTB or preimmune antiserum. Immunoprecipitated material was released from protein A-Sepharose beads by boiling and was analyzed on 9% gels. The first three lanes show the cross-linked material directly. (C) Bacterially expressed PTB. 121Gex3 fusion protein was prepared as described in Materials and methods and used to immunize rabbits. Lanes show the parental Gex vector-encoded protein surrounding the PTB/Gex fusion protein expressed in HB101 cells. (D) 121Gex3 fusion protein cross-links to RNA. The fusion protein was incubated with B3P3 RNA under standard binding conditions at 4°C and cross-linked as above. Some of the breakdown products of this fusion protein are still able to cross-link.

al. 1988; Fig. 5). hnRNP L is one member of a family of proteins that associate with nascent RNA transcripts (Piñol-Roma et al. 1988). Interestingly, both hnRNP L and PTB have nonconsensus RNA-binding domains. The ELAV protein is a neuron-specific product required for neuronal maturation (Robinow et al. 1988). ELAV con-

tains three RNP consensus sequences, and the homology between PTB and ELAV is strongest in these regions. Comparison of the three proteins reveals the presence of three regions of relatively high conservation that maintain approximately equal spacing in the three proteins (Fig. 5). These conserved regions fall within the putative

Region 1.

hnRNP L	69	AS VVHIR L V EADL V	FG I S VVVM	K QALVEFE	FA VNY A
PTB	56	VPSRVIHIRK LPIDVTEGEV	ISLGLPFGKV	TNLLMLKGN	QAFIEMNTEE AANTMVNYIT
ELAV	151	L V LP MTE EI SL	G I V LIRDK	Q YID	Q AP

Region 2.

hnRNP L	193	I F KN QA VEF	SAQ RAK SLNG D IY GCCTLKI EY	K T LNV	NDQ WDY T PNL G
PTB	206	KFGTVLKIIT FTRKNQFQAL	LOYADPVSAQ HAKLSLDGQN	IYNACCTLRI DFSKLTSLNV	KYNNDKSRDY TRPDLPS
ELAV	281	FG II L N D QGF I F	A RA I LNG	I V FS	G K F POL G

Region 3.

hnRNP L	348	AD VLMV LD	KM D VF VF LYGNVE	KVKFM K G	AMVEMADG	D AITHLN N M G L V VSK Q I
PTB	334	AG NSVLLVSNLN	PERVTPOSLE ILFGVYGDVQ	RVKILFNKKE	NALVQADGN QAQLAMSHLN	GKHLHGKPIR ITLSKHQNV
ELAV	417	A EPI NL	PE LW LFG FG VQ	VKIV DRKG	GFV M N D EA MAI LN G M R L	V F N

Figure 5. Sequence comparison of PTB with hnRNP L and ELAV. The three regions of highest homology between these three proteins and their relative positions on PTB are shown. Identical amino acids are in boldface type; conserved amino acid changes are in normal type. The RNP domains in hnRNP L and ELAV are boxed in thin lines, with the conserved RNP-1 and RNP-2 sequences boxed in thick lines. The putative RNP-2 box in PTB is also enclosed in thick lines.

RNA-binding domains of hnRNP L and ELAV. Although the sequences of the two RNP boxes are somewhat variable, there is strong conservation of several hydrophobic residues in these regions. These residues are important for RNA binding (Query et al. 1989), which suggests that they may play a similar functional role in PTB.

Expression of PTB is tissue and developmentally regulated

Because of the high conservation between the human and rodent PTB sequence (see below), the pattern of expression of the gene encoding PTB has been analyzed by probing Northern blots of RNAs extracted from various rat tissues using the human cDNA clones. As shown in Figure 6, PTB was detected as an ~3-kb transcript, a size consistent with the length of the cDNA clones. A similarly sized transcript was detected in HeLa cell RNA (data not shown). Following longer exposures than those shown here, this mRNA was detected in all tissues. By normalizing the signal obtained with the PTB probe to that obtained with a probe for phosphoglycerate kinase, a constitutively expressed gene (Michelson et al. 1983), it is apparent that despite its ubiquitous presence, there is indeed developmental and tissue-specific variation in the expression of this gene. In general, the level of expression appears to be higher in fetal than in adult tissues. The expression in brain is significantly lower than that seen in most tissues, which is intriguing given the homology to the neuron-specific ELAV protein. These developmental and tissue-specific differences in the lev-

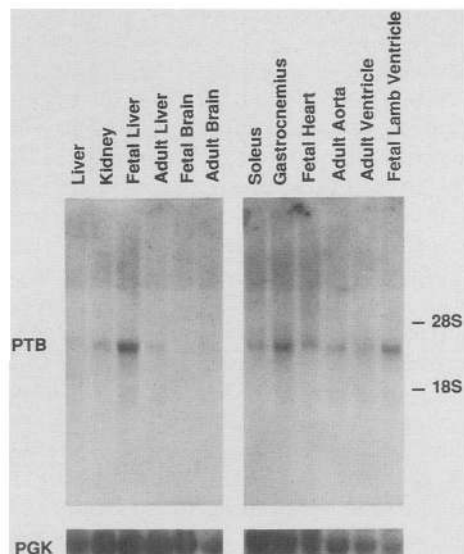


Figure 6. Expression of PTB shows developmental and tissue-specific regulation. Northern blots of RNA extracted from various rat tissues were probed with PTB cDNA clones. The same blots were then reprobed with a cDNA probe encoding phosphoglycerate kinase, a ubiquitously expressed transcript. Only the ~2-kb band is shown for this hybridization. The positions of 18S and 28S RNAs are shown.

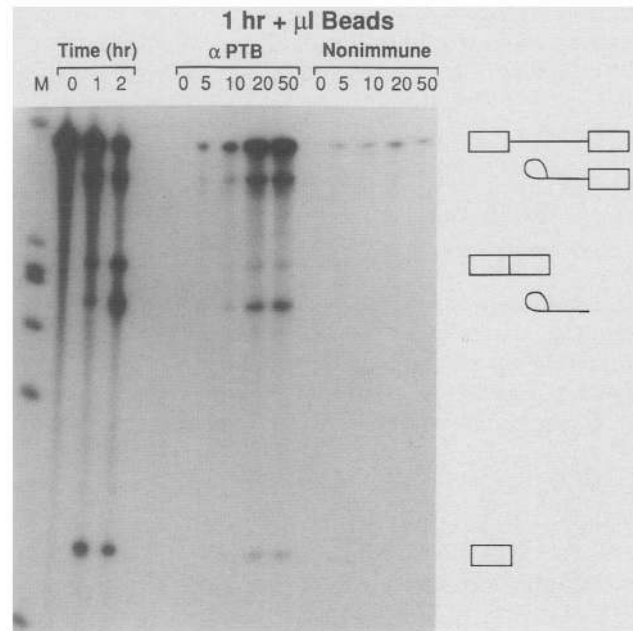


Figure 7. PTB is a spliceosome protein. Splicing reactions using a single intron construct from α -TM (pG3+DX) were incubated at 30°C for 1 hr, followed by the addition of increasing amounts of protein A-Sepharose beads prebound with either anti-PTB antisera or preimmune sera. The beads were washed, RNA-eluted, and applied to a 6% polyacrylamide gel. The various intermediates and products are as noted. (Left) A control splicing reaction at zero time, 1 hr, and 2 hr.

els of expression are unexpected if PTB is a constitutive splicing factor. However, the concentration of specific splicing factors may vary from tissue to tissue, permitting certain alternative splicing events to occur (Ge and Manley 1990; Krainer et al. 1990a,b). Despite almost undetectable levels of PTB mRNA in brain, we have isolated and sequenced cDNA clones from human fetal brain libraries confirming that the PTB gene is expressed in this tissue.

Is PTB an essential splicing factor?

The results showing specific binding of PTB to the polypyrimidine tract, as well as the direct correlation between binding and splice site selection (Mullen et al. 1991), suggest that PTB plays a role in splicing. Moreover, it appears that PTB is a component of the spliceosome as shown both by native gel electrophoresis (Garcia-Blanco et al. 1989) and by its sedimentation through glycerol gradients (data not shown). To determine this association more definitively, *in vitro* splicing reactions were immunoprecipitated with anti-PTB antibodies raised against bacterially produced PTB. These antibodies specifically immunoprecipitate spliceosomes, as indicated by the presence of splicing intermediates (Fig. 7). Importantly, relatively less of the spliced product is immunoprecipitated compared with the amount of lariat product. This is expected because the mRNA is released

from the spliceosome and does not have the intron-containing polypyrimidine tract sequences necessary for PTB binding. This result clearly demonstrates that PTB is a spliceosome protein.

To test directly the role of PTB in splicing, we have used a complementation assay for *in vitro* splicing similar to the one described previously by Zamore and Green (1991). As shown in Figure 8, a three-component system was derived by fractionating nuclear extract (adjusted to 500 mM KCl) on poly(U)-Sephacel columns and substituting purified material for the appropriate fractions. Nuclear extracts were divided into three fractions (Fig. 8): the flowthrough fraction (500 mM F/T) containing essential snRNPs and other necessary factors; a PTB-containing fraction (1 M step); and a third fraction

consisting of the essential splicing factor U2AF (Zamore and Green 1991). Significant depletion of the flowthrough material of both PTB and U2AF was verified by both Western analyses and UV cross-linking (data not shown). Use of the depleted flowthrough material, as well as the complementing fractions in a splicing assay, shows that all three fractions are required for efficient splicing (Fig. 8, middle lane). None of the fractions by themselves (data not shown), or in two-way combinations, support splicing. Importantly, addition of a PTB-containing fraction is required for splicing, and substitution of purified PTB (see Fig. 2A) for the 1 M step fraction also restores splicing activity (Fig. 8, right lane). The complementation activity of both the crude PTB fraction (1 M step) and column-purified protein is concentration dependent (data not shown). Taken at face value, the results of the complementation assay indicate that PTB is an essential pre-mRNA splicing factor.

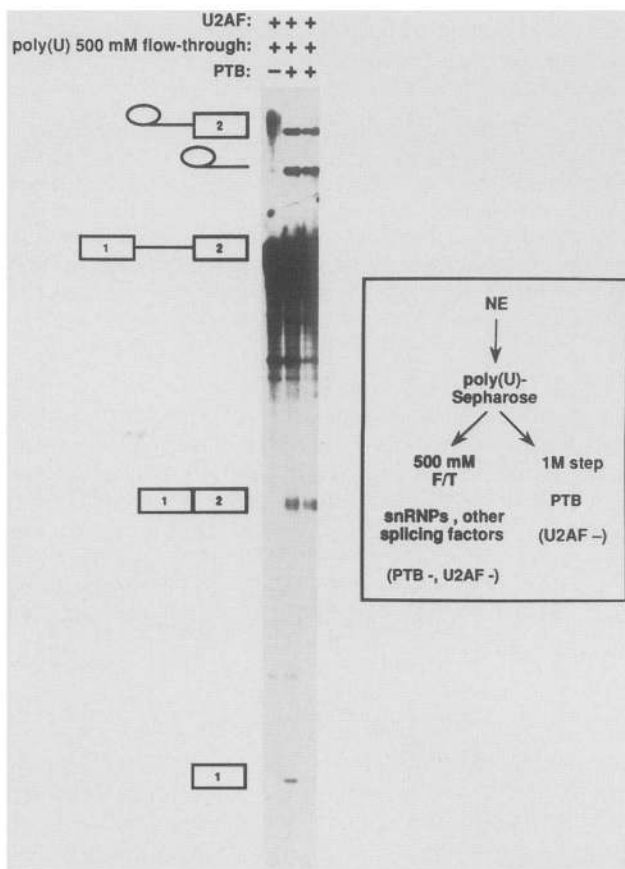


Figure 8. PTB-enriched fractions are necessary for *in vitro* splicing. Nuclear extracts were adjusted to 500 mM KCl and applied to poly(U)-Sephacel columns. Flowthrough material was collected, washed extensively at 500 mM KCl, and step-eluted at 1.0 M KCl. The 1 M step fraction contains PTB. (*Left* lane) Lack of splicing in the absence of PTB-containing material; (*center* lane) restoration of splicing activity with the 1 M step elution; (*right* lane) restoration of activity with purified PTB. The splicing substrate for the reactions shown was an adenovirus-derived pre-mRNA (MINX; see Materials and methods). Identical results were obtained using the α -TM construct pG3+DX.

PTB is part of the large protein complex required for splicing

The cross-linking of PTB to different RNA substrates is quite efficient, reflecting either the kinetics of PTB binding to RNA and/or the possibility that more than one PTB molecule is cross-linked per pyrimidine tract. This behavior has raised the question of whether PTB is present in nuclear extracts as a monomer or multimer. To address this issue, purified PTB was subjected to glycerol gradient sedimentation and found to sediment as part of a large complex. To determine more precisely the size of this complex, PTB, prepared as shown in Figure 2A, was analyzed by gel filtration chromatography. No PTB eluted with fractions corresponding to monomer size. Rather, the peak of PTB activity eluted with fractions corresponding to ~325 kD (Fig. 9A), suggesting that PTB does not exist as a free 57-kD protein. When stained with Coomassie blue, the gel filtration fractions revealed the presence of two major proteins of 100 and 57 kD (PTB), as well as a group of four smaller proteins of which a 33-kD protein was the most abundant. Detection of these additional proteins stained with Coomassie blue prompted us to re-examine the previous preparations of PTB that had been visualized by silver staining. As shown in Figure 9B, these preparations, judged previously to be homogeneous, all contained the three major proteins detected in the gel filtration fractions. These additional proteins are not visualized by silver staining. (Note that the two bands between the 100-kD protein and PTB are breakdown products of the 100-kD protein, as determined by their kinetics of appearance in the purified preparations, as well as their staining properties). This differential staining property has been noted for other proteins such as Ca^{2+} -binding proteins (Mooibroek et al. 1987). These three proteins, 100 kD, PTB, and 33 kD, copurify through DEAE ion-exchange columns and two affinity chromatography steps despite the presence of detergent, a highly reducing environment, and KCl concentrations as high as 500 mM. Although

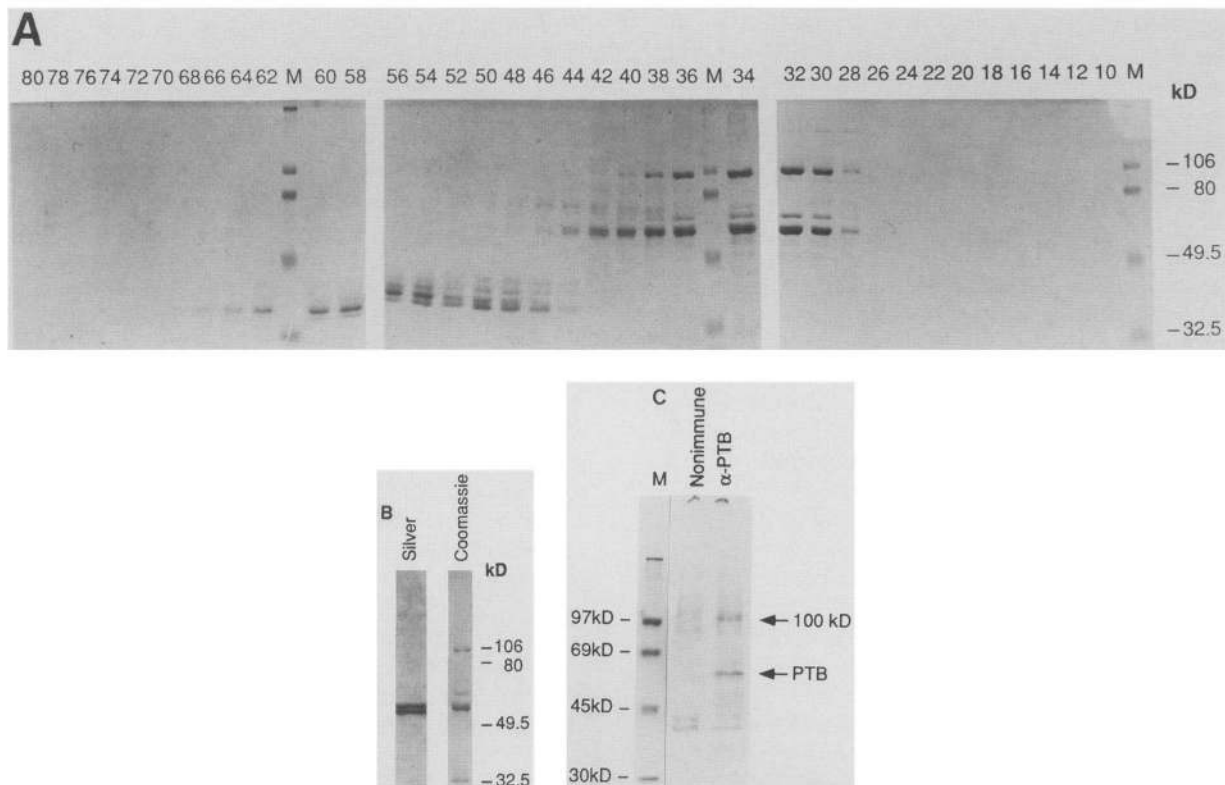


Figure 9. PTB exists as part of a 325-kD complex with a 100-kD protein. (A) Partially purified PTB was applied to a Sephacryl S300HR column. Fractions were collected, and an aliquot was precipitated and applied to a 9% Laemmli gel. Visualization was by Coomassie staining. Numbering is from right to left, beginning with the tenth fraction collected off the column. Protein standards from Pharmacia were used to calibrate the column. Thyroglobulin (669 kD) elution peaked at fraction 25, ferritin (440 kD) peaked at fraction 31, and catalase (232 kD) peaked at fraction 37. (B) PTB fractions contain peptides that do not silver-stain. Equivalent amounts of purified PTB protein (the same preparations as shown in Fig. 2A) were loaded onto gels and stained using either a Bio-Rad silver staining kit or Coomassie blue. (C) Immunoprecipitation of the 100-kD protein with anti-PTB antibodies. ^{35}S -Labeled HeLa mininuclear cell extracts were prepared as described (Schreiber et al. 1989). Protein A-Sepharose prebound with either preimmune or anti-PTB antisera was then added, followed by washing and elution by boiling in loading buffer. The positions of PTB and the 100-kD protein are as indicated.

other interpretations are possible, the results obtained are most consistent with the hypothesis that the 100-kD protein and PTB are present in a large complex. Alternatively, self-aggregation could also account for the gel filtration results shown above. We feel that this is unlikely because these fractions retain splicing activity (see below), and aggregated material would likely elute in the void volume of the gel filtration column. Nevertheless, to control for this possibility, anti-PTB antibodies were used to immunoprecipitate proteins from ^{35}S -labeled HeLa nuclear extracts. As shown in Figure 9C, two major proteins were detected of ~100 and 60 kD. Western blot analysis has shown that the 60-kD protein is indeed PTB. Given the results shown in Figure 9A, it is likely that the protein migrating at 100 kD is the protein associated with PTB. As expected, additional proteins are also immunoprecipitated because these antibodies immunoprecipitate spliceosomes (Fig. 7). Further evidence that this complex is genuine was obtained by incubating partially purified PTB (containing the 100-kD protein) with PTB fusion protein, followed by glutathione-agarose chromatography. This results in copurification of the 100-kD protein with the fusion protein (data not shown) and

demonstrates the specific interaction between these two proteins.

In light of the data in Figure 9, the complementation assay shown in Figure 8 must be strictly interpreted to indicate the presence of a splicing activity in the PTB preparations that may be associated with one of the individual polypeptides or a combination thereof. To determine which of these three proteins is able to restore splicing to depleted extracts, three pools were prepared from the gel filtration fractions shown in Figure 9A. Pool I (fractions 30–36) and pool II (fractions 37–42) contain PTB and the 100-kD protein but lack the 33-kD component. Pool III (fractions 43–54) contains a group of proteins, including the 33-kD protein. When these pooled fractions were used in the same complementation assay shown in Figure 8, only pools I and II were able to rescue splicing (data not shown). No activity was detected in pool III. However, it should be noted that these pools have a low protein concentration and failure to rescue splicing could be due to the addition of insufficient material. These results indicate that the splicing activity missing from depleted extracts is associated with PTB, the 100-kD protein, or a combination of the two.

Patton et al.

To further define this activity, a new scheme for the purification of PTB was devised that yields homogeneously pure PTB, as determined by a variety of staining procedures, including Coomassie blue. However, these preparations did not complement depleted splicing extracts as did preparations containing both PTB and the 100-kD protein. This result indicates that either PTB is not the splicing factor missing from the depleted extracts or that these extracts are depleted of more than one factor. Because both proteins, together with the 33-kD component, elute from poly(U)-Sephacryl at similar salt concentrations, the latter possibility appears most likely. At this time we are only able to monitor the depletion of PTB from the extract but not the depletion state of either the 100- or 33-kD protein. For this reason, we attempted to remove PTB exclusively from splicing extracts by immunodepletion using specific polyclonal antibodies (Fig. 4). These extracts are efficiently depleted of PTB, as detected by Western blots, and are unable to splice pre-mRNA (data not shown). However, complementation of these extracts with crude or enriched preparations of PTB fails to restore splicing activity. This result is probable, because the immunoprecipitation removes multiple proteins, including the 100-kD protein (Fig. 9C). Attempts to find conditions that only immunoprecipitate PTB from extracts have failed so far. These results are consistent with the complementation data shown above and further suggest the existence of complexes that might include more than one splicing factor.

The presence of preformed splicing complexes in nuclear extracts has been documented previously (Konarska and Sharp 1988; Ast et al. 1991). Their existence makes immunodepletion and complementation experiments of factors present in such complexes, as appears to be the case for PTB, a more difficult task than was first envisioned. The nature of these preformed complexes and the partition of a given splicing factor between the free and complexed fractions remains to be determined. Thus, determination of the specific role of PTB, the 100- and 33-kD proteins, will require the separation of the individual proteins in a functional form, as well as a suitable complementation assay in which the depletion of each protein can be monitored. Despite these difficulties, it seems fair to conclude that PTB, together with the 100-kD protein, is required for splicing.

Discussion

This report describes the identification, characterization, molecular cloning, and sequence of a 57-kD polypyrimidine-binding protein. This protein is identical to a previously identified 62-kD protein termed PTB (Garcia-Blanco et al. 1989; see Gil et al., this issue). PTB is a spliceosome protein that binds the BP/PPT of mammalian introns in a manner that correlates directly with 3'-splice-site selection as determined in *cis*-competition assays (Mullen et al. 1991), and it exists as part of a large complex consisting of PTB and a 100-kD protein. As demonstrated by depletion-complementation assays, this complex is required for splicing.

PTB is an evolutionarily conserved member of a group of RNA-binding proteins

The PTB cDNA sequence predicts a protein of 531 amino acids with a size of 57.2 kD, which matches well with the electrophoretically determined size of the corresponding protein from HeLa cell nuclear extracts. PTB contains an unusual RNP-binding domain (Bandzulis et al. 1989; Query et al. 1989). It does not contain sequence motifs similar to other RNA-binding proteins, such as Tra (Boggs et al. 1987), suppressor of white apricot (Chou et al. 1987), the family of RNA-dependent ATPases (see Wasserman and Steitz 1991), or a serine/arginine-rich domain seen in still other RNA-binding proteins (Lazinski et al. 1989; Amrein et al. 1988). A search of a data base containing a collection of both yeast and mammalian splicing factors did not detect any significant homology of any of these proteins to PTB (J. Arenas, pers. comm.). Sequences similar to the RNP-1 and RNP-2 boxes are present in PTB, although the distance between these elements is >100 amino acids greater than that found in consensus RNP domains (Bandzulis et al. 1989; Query et al. 1989). Nevertheless, several hydrophobic residues found within this domain are conserved. A search of the GenBank and EMBL data bases for similar proteins revealed that a previously sequenced but unidentified cDNA clone was indeed the mouse homolog for PTB (A.L.M. Bothwell, pers. comm.). The homology at the amino acid level is 98%. The high degree of conservation agrees with genomic DNA analysis using the PTB cDNA as a hybridization probe, indicating that this sequence can be detected in a wide variety of species. Hybridizing bands are detected in species as diverse as yeast, *Drosophila*, *Caenorhabditis elegans*, frogs, fish, rat, and human (data not shown). If these cross-hybridizing bands represent true PTB gene homologs, it would reveal a remarkable conservation through evolution and further suggest an important functional role for PTB. Structure-function relationships would be greatly facilitated if organisms favorable to genetic manipulation contain a PTB homolog.

Additional genes were also found to be similar, the most striking of which was the homology to the human hnRNP L protein (Piñol-Roma 1989) and to the *Drosophila* ELAV protein (Robinow et al. 1988). Comparison between these proteins reveals three highly conserved regions (Fig. 5). The two downstream regions show remarkable identity, especially between hnRNP L and PTB, and maintain similar spacing in the two proteins. These regions coincide with the predicted RNA-binding domains present in both hnRNP L and ELAV. In addition, the third region of homology contains a sequence found in the consensus RNP domain of the U1 snRNP 70K protein (Theissen et al. 1986; Query and Keene 1987; Spritz et al. 1987). The closest match to the RNP 2 box in PTB (LLVSNL) is located near the 5' end of this conserved region, suggesting that despite the abnormal spacing of the two RNP boxes in PTB, this region may function in RNA binding. It will be necessary to accumulate more related proteins to determine the impor-

tance of these conserved regions, as well as direct determination of the RNA-binding site in PTB using mutational analysis and cross-linking assays.

Is PTB an hnRNP protein?

The similarity of PTB to hnRNP L suggests that PTB may be a member of the family of hnRNP proteins (Piñol-Roma et al. 1988). By strict size comparison, PTB is most similar to the hnRNP M protein (Swanson and Dreyfuss 1988b). However, this protein binds to poly(U)-Sephacel in the presence of 2 M KCl, whereas PTB elutes from this resin at 750 mM KCl, suggesting that these proteins are distinct. Moreover, hnRNP proteins are generally thought to bind RNA nonspecifically (Conway et al. 1988), much the way histones bind DNA. If this is the case, PTB does not fit the description of an hnRNP protein. However, several hnRNP proteins have been shown to exhibit some sequence specificity (Swanson and Dreyfuss 1988a,b; Wilusz and Shenk 1990). Because the hnRNP classification is based on the binding to nascent RNA transcripts, it is likely that the family of proteins now referred to as hnRNPs consists of proteins that exhibit general RNA binding, sequence-specific RNA binding, spliceosome constituents, and perhaps essential splicing factors. As the specific nature and functional role of individual proteins emerges, a more precise classification will be necessary to appropriately categorize the multiplicity of pre-mRNA-binding proteins.

A large complex containing PTB is required for splicing

In contrast to the large number of pre-mRNA splicing mutants in yeast (Ruby and Abelson 1991), the lack of mammalian splicing mutants has made the biochemical approach the method of choice for the identification and characterization of splicing factors. To determine biochemically whether a specific protein is an essential splicing factor, a complementation assay is required in which addition of the factor restores splicing activity to depleted extracts. In the case of U2AF, depleted splicing extracts were obtained due to high-affinity binding of this factor to poly(U)-Sephacel (Zamore and Green 1991). For ASF/SF2, the precise nuclear retention of this factor during cell fractionation allowed preparation of cytoplasmic extracts deficient in splicing only because they lack this protein. For splicing proteins contained in larger complexes, as appears to be the case for PTB, the complementation of depleted extracts presents more difficulties due to the fact that all of the components of the complex are removed during depletion. In these cases, complementation of the depleted extract requires all of the components of the complex. The lack of complementation by an individual component is therefore not unexpected.

The functional characterization of PTB has to be interpreted in light of the issues raised above. In the work presented here, nuclear extracts were divided into three fractions necessary for splicing. Complementation anal-

ysis carried out with these fractions has confirmed the requirement for U2AF in splicing and has shown that PTB-containing fractions are also necessary for splicing. PTB appears to exist with the 100-kD protein as part of a large complex. Our results suggest that this complex is required for splicing because homogeneously purified PTB does not rescue splicing while fractions containing both PTB and the 100-kD protein are able to restore splicing to depleted extracts. It could be argued from this observation that PTB is not a splicing factor. However, the specific spliceosome association of PTB, together with the complementation data, suggest that PTB is required as part of a large complex, including the 100-kD protein. Although a third protein of 33 kD is also present in these fractions, it is not required in the complementation assay shown here. Whether this protein is an essential splicing factor remains to be determined, because we cannot determine the depletion of either the 100- or 33-kD proteins from splicing extracts in our complementation assay. Because purification of three similarly sized proteins is observed when biotinylated B3P3 RNA is incubated with nuclear extract and associated proteins are determined (R. Reed, pers. comm.; see Reed 1990) it is possible that all three bind the polypyrimidine tract as a complex and all three might be required for splicing.

The relationship of the 100- and 33-kD proteins to other splicing factors is not certain. Partial peptide sequence of the 100-kD protein has revealed that this protein contains a consensus RNP domain but does not match existing sequences in the data banks (J.G. Patton and E.B. Porro, unpubl.). A 100-kD factor, termed intron-binding protein (IBP), has been identified that binds to the 3'-splice site and polypyrimidine tract (Gerke and Steltz 1986; Tazi et al. 1986). However, IBP is precipitable with anti-Sm sera, whereas Western blots of our PTB-enriched material do not show any reactivity with Sm antibodies. The size of the 33-kD protein is similar to ASF/SF2 (Ge and Manley 1990; Krainer et al. 1990a,b), SC-35 (Fu and Maniatis 1990), and the small subunit of U2AF (Zamore and Green 1989). Western blots have shown that the 33-kD protein is not the small subunit of U2AF (P. Zamore, pers. comm.). Both SC-35 and ASF/SF-2 can be silver-stained, suggesting that these proteins are not related to the PTB-associated 33-kD protein. More work will be required to determine such relationships, if any.

The data presented here indicate that PTB is one subunit of a large complex necessary for splicing. Whether PTB is individually required for splicing and/or whether the 100-kD protein, when dissociated, will be able to rescue splicing remains to be determined. However, several lines of indirect evidence strongly suggest that PTB is required for splicing. First, it binds specifically to the functionally important polypyrimidine tract and remains associated with spliceosomes (Fig. 7). Second, the binding of PTB correlates exactly with both splicing efficiency in single intron constructs (Smith et al. 1989) and with splice site selection in two intron *cis*-competition splicing assays both *in vitro* and *in vivo* (Mullen et al. 1991). Third, immunodepletion of PTB from nuclear

Patton et al.

extracts abolishes splicing. Therefore, the indirect evidence, together with the direct complementation results and spliceosome association, strongly indicates that PTB is one component of an essential splicing factor.

Materials and methods

Constructs

Constructs pG3 + DX, pG3X23H, Δ 60-110, and Δ 52-110 are all as described previously (Smith and Nadal-Ginard 1989; Smith et al. 1989). B3P3 was prepared by subcloning the *SmaI*-*AccI* fragment from α -TM intron 2 (see Fig. 1D) into the *SmaI* site of pGem-4 after fill-in of the *AccI* site. RNA substrates containing only a branchpoint (B3) and either 17 or 9 pyrimidines were prepared by subcloning the corresponding fragments, amplified by PCR, from the full-length splicing substrates, Δ 60-110 and Δ 52-110 (for sequences, see Mullen et al. 1991). All of the above clones were verified by sequencing. 121Gex3 was prepared by subcloning an *EcoRI* fragment from λ clone 121 into the Gex3 vector (Smith and Johnson 1988). Expression and purification of bacterial protein was as described (Smith and Corcoran 1990). MINX is an adenovirus-derived pre-mRNA substrate, as described previously (Zillman et al. 1988).

Nuclear extracts, RNA transcriptions, and splicing reactions

Nuclear extracts were prepared from HeLa cell suspension cultures and from a rat smooth muscle cell line (kindly provided by A. Rothman et al. 1991) by the procedure of Abmayr et al. (1988). Transcription reactions for the preparation of RNA templates were all performed using SP6 or T7 polymerase as described (Smith and Nadal-Ginard 1989). Splicing reactions contained 0.5 mM ATP, 2 mM MgCl₂, 20 mM creatine phosphate, 0.5 mM DTT, 20 units of RNasin, 60% nuclear extract, and 5–10 ng of RNA. Binding reactions were carried out under identical conditions except for the omission of ATP and creatine phosphate. Reactions were maintained at 4°C or 30°C for the times indicated in the figure legends.

UV cross-linking

Binding reactions were set up in a 96-well microtiter dish (Falcon MicroTest III) and incubated for various times at 30°C. The dish was placed in a box lined with aluminum foil, and a 254-nm UV light transilluminator was placed over the box at a distance of 7 cm and irradiated for 15 min. In control experiments, irradiation over time showed an almost linear response for the first 10 min where the signal reached a plateau for the cross-linking of PTB. Following irradiation, RNase A was added to a final concentration of 25 μ g/ml, micrococcal nuclease to 50 U/ μ l, and CaCl₂ to 1 mM. Digestion was for 1 hr at 37°C. The reactions were adjusted to 40 mM Tris (pH 6.8), 7% glycerol, 1.5% SDS, and 400 mM 2-mercaptoethanol and electrophoresed on 9% Laemmli gels (Laemmli 1970).

Purification of PTB

Nuclear extracts were prepared as described above and applied to DEAE-Sephacel CL-6B columns equilibrated in buffer A plus 150 mM KCl. Buffer A is 20 mM Tris (pH 8.0), 3 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, and 0.05% NP-40. The flowthrough was collected and applied to a heparin-agarose column equilibrated in buffer A plus 150 mM KCl. The column was washed extensively with buffer A plus 100 mM KCl, and PTB was eluted with buffer A containing 300 mM KCl. The eluted

material was adjusted to 150 mM KCl and applied to a poly(U)-Sephacel 4B column equilibrated in buffer A plus 150 mM KCl. The column was washed extensively with buffer A plus 500 mM KCl and step-eluted at 750 mM KCl. Purified material was dialyzed against buffer A plus 150 mM KCl containing 20% glycerol. This protocol yielded a preparation that, when silver stained, showed a doublet or triplet of proteins of apparent molecular mass of 57 kD. However, when examined by Coomassie blue staining, copurification of PTB with a 100- and 33-kD protein is apparent with this protocol.

Peptide sequencing

Purified PTB was transferred to nitrocellulose and digested as in Ghosh et al. (1990). Separation of peptides and sequencing was as in Tempst and Riviere (1989) and Tempst et al. (1990).

Gel filtration analysis of PTB

PTB-containing fractions in dialysis buffer were applied to a Sephacryl S300HR column (0.8 \times 47 cm) at a flow rate of 3.0 ml/hr in buffer A containing 150 mM KCl. Aliquots of each of the fractions were precipitated with 25% TCA and analyzed on 9% Laemmli gels.

Cloning and sequencing of PTB

A 69-bp cloned fragment derived from PCR amplification using degenerate oligonucleotides corresponding to the 5' and 3' ends of peptide 2T21 was 3'-end-labeled and used to screen a HeLa cell cDNA library (Clontech) or a human fetal brain cell library primed with oligonucleotide d(T). Of several clones obtained, two clones were examined in detail, a 2.7-kb clone (212) and an overlapping and 5'-extended 1.7-kb clone (121). Plating, lifting, probing, and phage purification were all as described (Sambrook et al. 1989).

Sequence analysis

Sequence analysis was carried out on exonuclease III deletion subclones (Henikoff 1984). Sequence was obtained both manually, using dideoxy terminators, and on an Applied Biosystems model 370A automated DNA sequencer, using *Taq* or Sequenase polymerase on double-stranded templates. Sequences were analyzed using the University of Wisconsin (Madison) GCG software, as well as the BLAST program (Altschul et al. 1990).

Northern blot analysis

Poly(A)⁺ RNA (10 μ g) was isolated from various rat tissues, separated on formaldehyde-agarose gels, and transferred to GeneScreen Plus (New England Nuclear) following the manufacturer's protocols. Hybridization was in the presence of 50% formamide using random-primed PTB cDNA probes.

Immunoprecipitations and Western blots

Antibodies were prepared by injection of rabbits with glutathione-agarose-purified 121 Gex3 fusion protein following standard protocols (Harlow and Lane 1988). For immunoprecipitation, sera were incubated with protein A-Sephacel for 1 hr at 4°C, followed by addition of cross-linked material for 1 hr at 4°C. Beads were then washed with at least 10 column volumes of a solution containing 150 mM NaCl, 5 mM EDTA, 50 mM Tris (pH 7.5), and 0.05% NP-40. Bound material was eluted by boiling in Laemmli loading buffer (above). For Western blots, pro-

teins were transferred to nitrocellulose in 20% methanol, 0.4% SDS, 200 mM glycine, and 20 mM Tris. Following transfer, blots were blocked in 5% nonfat dry milk with 10 mM Tris (pH 7.4), 0.1% Tween 20, and 0.9% NaCl. Antibody dilutions were in this same buffer. Antibody incubations were at room temperature for 1 hr, followed by incubation with donkey anti-rabbit IgG coupled to horseradish peroxidase for 1 hr, at room temperature. Visualization was with the ECL reagents from Amersham following their protocol.

Preparation of PTB-depleted nuclear extracts and splicing complementation

Nuclear extract (1.5 ml) was adjusted to 500 mM KCl in buffer A and applied to a 3.0-ml poly(U)-Sepharose column. Fractions of 0.5 ml were collected, and the peak of protein was determined using a Bradford assay. Step fractions were eluted with 1.0 M KCl or 2 M guanidine, and the resultant peaks were dialyzed in buffer A with 100 mM KCl and 20% glycerol. Typically, the complementation reactions utilized purified U2AF (Zamore and Green 1989) rather than the guanidine-eluted U2AF. The effective depletion of PTB was monitored by both cross-linking and Western blot assays. Complementation reactions were as described (Zamore and Green 1991) except that in a typical 20- μ l splicing reaction, the flowthrough material was 3 μ l, PTB-containing fractions were 5 μ l, and purified U2AF was 2 μ l. Splicing reactions were as described (Zamore and Green 1991) using the MINX substrate or, as above, using the α -TM construct pG3+DX.

Acknowledgments

We especially thank Phillip D. Zamore and Michael Green for their invaluable and generous help in setting up the complementation assays and providing purified U2AF. Their comments and input in the generation and evaluation of the functional data are greatly appreciated. The human fetal brain cDNA library was kindly provided by Lou Kunkel. PGK DNA was kindly provided by Stuart Orkin. We thank Al Bothwell for communicating results prior to publication. Thanks also go to Emily Fernandes, Ok Hwang, and Lise Riviere for their expert technical assistance, Tom O'Keefe for help in sequencing the cDNA clones, Emily Flynn McIntosh for preparing the figures, and Chris Smith for advice and comments throughout the project. J.G.P. acknowledges support from National Institutes of Health postdoctoral fellowship 1F32GM12748.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

The isoform sequenced in this manuscript will be referred to as PTB-1. Alternatively spliced forms will be labeled as in Gil et al. (this issue). To date, there are four isoforms of PTB, PTB-1 through PTB-4.

References

Abmayr, S.M., J.L. Workman, and R.G. Roeder. 1988. The pseudorabies immediate early protein stimulates in vitro transcription by facilitating TFIID: promoter interactions. *Genes & Dev.* **2**: 542–553.

- Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- Amrein, H., M. Gorman, and R. Nothiger. 1988. The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell* **55**: 1025–1035.
- Ast, G., D. Goldblatt, D. Offen, J. Sperling, and R. Sperling. 1991. A novel splicing factor is an integral component of the 200S large nuclear ribonucleoprotein (lnRNP) particles. *EMBO J.* **10**: 425–432.
- Bandziulis, R.J., M.S. Swanson, and G. Dreyfuss. 1989. RNA-binding proteins as developmental regulators. *Genes & Dev.* **3**: 431–437.
- Berget, S.M. and B.L. Roberson. 1986. U1, U2, and U4/U6 small nuclear ribonucleoproteins are required for in vitro splicing but not polyadenylation. *Cell* **46**: 691–696.
- Bindereif, A. and M.R. Green. 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. *EMBO J.* **6**: 2415–2424.
- Black, D.L. and J.A. Steitz. 1986. Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein. *Cell* **46**: 697–704.
- Black, D.L., B. Chabot, and J.A. Steitz. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* **42**: 737–750.
- Boggs, R.T., P. Gregor, S. Idriss, J.M. Belote, and M. McKeown. 1987. Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene. *Cell* **50**: 739–747.
- Chabot, B., D.L. Black, D.M. LeMaster, and J.A. Steitz. 1985. The 3' splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science* **230**: 1344–1349.
- Choi, Y.D., P.J. Grabowski, P.A. Sharp, and G. Dreyfuss. 1986. Heterogeneous nuclear ribonucleoproteins: Role in RNA splicing. *Science* **231**: 1534–1539.
- Chou, T-B., Z. Zachar, and P.M. Bingham. 1987. Developmental expression of a regulatory gene is programmed at the level of splicing. *EMBO J.* **6**: 4095–4104.
- Conway, G., J. Wooley, T. Bibring, and W. LeSturgeon. 1988. Ribonucleoproteins package 700 nucleotides of pre-mRNA into a repeating array of regular particles. *Mol. Cell. Biol.* **8**: 2884–2895.
- Friendway, D. and W. Keller. 1985. The stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell* **42**: 355–367.
- Fu, X-D. and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**: 437–441.
- Garcia-Blanco, M.A., S.F. Jamison, and P.A. Sharp. 1989. Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. *Genes & Dev.* **3**: 1874–1886.
- Garcia-Blanco, M., G. Anderson, J. Beggs, and P.A. Sharp. 1990. A mammalian protein of 220 kDa binds pre-mRNAs in the spliceosome: A potential homologue of the yeast PRP8 protein. *Proc. Natl. Acad. Sci.* **87**: 3082–3086.
- Ge, H. and J.L. Manley. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**: 25–34.
- Gerke, V. and J.A. Steitz. 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3' splice site of premessenger RNA. *Cell* **47**: 973–984.
- Ghosh, S., A.M. Gifford, L.R. Riviere, P. Tempst, G. Nolan, and D. Baltimore. 1990. Cloning of the p50 DNA binding subunit of NF- κ B: Homology to rel and dorsal. *Cell* **62**: 1019–1029.
- Green, M.R. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.*

Patton et al.

- 20: 671–708.
- Harlow, E. and D. Lane. 1988. *Antibodies*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**: 351.
- Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46**: 845–855.
- . 1988. Association of U2, U4, U5, and U6 small nuclear ribonucleoproteins in a spliceosome-type complex in absence of precursor RNA. *Proc. Natl. Acad. Sci.* **85**: 5459–5462.
- Krainer, A.R. and T. Maniatis. 1985. Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. *Cell* **42**: 725–736.
- Krainer, A.R., G.C. Conway, and D. Kozak. 1990a. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & Dev.* **4**: 1158–1171.
- . 1990b. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**: 35–42.
- Kramer, A. 1988. Presplicing complex formation requires two proteins and U2 snRNP. *Genes & Dev.* **2**: 1155–1167.
- Kramer, A. and W. Keller. 1985. Purification of a protein required for the splicing of pre-mRNA and its separation from the lariat debranching enzyme. *EMBO J.* **4**: 3571–3581.
- Kramer, A., W. Keller, B. Appel, and R. Luhrmann. 1984. The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell* **38**: 299–307.
- Kramer, A., M. Frick, and W. Keller. 1987. Separation of multiple components of HeLa cell nuclear extracts required for pre-messenger RNA splicing. *J. Biol. Chem.* **262**: 17630–17640.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.
- Lazinski, D., E. Grzadzińska, and A. Das. 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59**: 207–218.
- Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325**: 673–678.
- Mayeda, A., K. Tatei, H. Kitayama, K. Takemura, and Y. Ohshima. 1986. Three distinct activities possibly involved in mRNA splicing are found in a nuclear fraction lacking U1 and U2 RNA. *Nucleic Acids Res.* **14**: 3045–3057.
- Michelson, A.M., A.F. Markham, and S.H. Orkin. 1983. Isolation and DNA sequence of a full-length cDNA clone for human X chromosome-encoded phosphoglycerate kinase. *Proc. Natl. Acad. Sci.* **80**: 472–476.
- Mooibroek, M.J., D.F. Michiel, and J.H. Wang. 1987. Clathrin light chains are calcium-binding proteins. *J. Biol. Chem.* **262**: 25–28.
- Mount, S.M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**: 459–472.
- Mount, S.M., I. Pettersson, M. Hinterberger, A. Karmas, and J.A. Steitz. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* **33**: 509–518.
- Mullen, M.P., C.W.J. Smith, J.G. Patton, and B. Nadal-Ginard. 1991. α -Tropomyosin mutually exclusive exon selection: Competition between branchpoint/polypyrimidine tracts determines exon choice. *Genes & Dev.* **5**: 642–655.
- Nagai, K., C. Oubridge, T.H. Jessen, J. Li, and P.R. Evans. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**: 515–520.
- Ohshima, Y. and Y. Gotoh. 1987. Signals for the selection of a splice site in pre-mRNA. *J. Mol. Biol.* **195**: 247–259.
- Padgett, R.A., P.J. Grabowski, M.M. Konarska, S. Seiler, and P.A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55**: 1119–1150.
- Parker, R., P.G. Siliciano, and C. Guthrie. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**: 229–239.
- Piñol-Roma, S., Y.D. Choi, M. Matunis, and G. Dreyfuss. 1988. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes & Dev.* **2**: 215–227.
- Piñol-Roma, S., M.S. Swanson, J.G. Gall, and G. Dreyfuss. 1989. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J. Cell Biol.* **109**: 2575–2587.
- Pruzan, R., H. Furneaux, P. Lassota, G.Y. Hong, and J. Hurwitz. 1990. Assemblage of the prespliceosome complex with separated fractions isolated from HeLa cells. *J. Biol. Chem.* **265**: 2804–2813.
- Query, C.C. and J.D. Keene. 1987. A human autoimmune protein associated with U1 RNA contains a region of homology that is cross-reactive with a retroviral p30^{gag} antigen. *Cell* **51**: 211–220.
- Query, C.C., R.C. Bentley, and J.D. Keene. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* **57**: 89–101.
- Reed, R. 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proc. Natl. Acad. Sci.* **87**: 8031–8035.
- Reed, R., J. Griffith, and T. Maniatis. 1988. Purification and visualization of native spliceosomes. *Cell* **53**: 949–961.
- Robinow, S., A. Campos, K.-M. Yao, and K. White. 1988. The elav gene product of *Drosophila*, required in neurons, has three RNP consensus motifs. *Science* **242**: 1570–1572.
- Ruby, S.W. and J. Abelson, J. 1991. Pre-mRNA splicing in yeast. *Trends Genet.* **7**: 79–85.
- Ruskin, B. and M.R. Green. 1985a. Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. *Cell* **43**: 131–142.
- . 1985b. Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. *Nature* **317**: 732–734.
- Ruskin, B., P.D. Zamore, and M.R. Green. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52**: 207–219.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Schreiber, E., P. Matthias, M. Müller, and W. Schaffner. 1989. Rapid detection of octamer binding proteins with mini-extracts prepared from a small number of cells. *Nucleic Acids Res.* **17**: 6419.
- Sharp, P. 1987. Splicing of messenger RNA precursors. *Science* **235**: 766–771.
- Smith, D.B. and L.M. Corcoran. 1990. Expression and purification of glutathione-S-transferase fusion proteins. In *Current protocols in molecular biology* (ed. F. Ausubel et al.), pp. 167.1–167.8. John Wiley/Greene, New York.
- Smith, D.B. and K.S. Johnson. 1988. Single step purification of polypeptides expressed in *E. coli* as fusions with glutathione S-transferase. *Gene* **67**: 31–40.
- Smith, C.W.J. and B. Nadal-Ginard. 1989. Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual

- lariat branchpoint location; implications for constitutive splicing. *Cell* **56**: 749–758.
- Smith, C.W.J., E.B. Porro, J.G. Patton, and B. Nadal-Ginard. 1989. Scanning from an independently specified branchpoint defines the 3' splice site of mammalian introns. *Nature* **342**: 243–247.
- Smoller, D., C. Friedel, A. Schmid, D. Bettler, L. Lam, and B. Yedvobnick. 1990. The *Drosophila* neurogenic locus *mastermind* encodes a nuclear protein unusually rich in amino acid homopolymers. *Genes & Dev.* **4**: 1688–1700.
- Spritz, R.A., K. Strunk, C.S. Surowy, S.O. Hoch, D.E. Barton, and U. Francke. 1987. The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing, and RNA binding. *Nucleic Acids Res.* **15**: 10272–10391.
- Steitz, J.A., D.L. Black, V. Gerke, K.A. Parker, A. Kramer, D. Frendeway, and W. Keller. 1988. Functions of the abundant U-snRNPs. In *Small nuclear ribonucleoprotein particles* (ed. M.L. Birnstiel), pp. 115–154. Springer-Verlag, Berlin.
- Swanson, M.S. and G. Dreyfuss. 1988a. RNA binding specificity of hnRNP proteins: A subset bind to the 3' end of introns. *EMBO J.* **7**: 3519–3529.
- . 1988b. Classification and purification of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. *Mol. Cell. Biol.* **8**: 2237–2241.
- Tazi, J., C. Alibert, J. Temsamani, I. Reveillaud, G. Cathala, C. Brunel, and P. Jeanteur. 1986. A protein that specifically recognizes the 3' splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. *Cell* **47**: 755–766.
- Tempst, P., A.J. Link, L.R. Riviere, M. Fleming, and C. Elicone. 1990. Internal sequence analysis of proteins separated on polyacrylamide gels at the sub-microgram level: Improved methods, applications and gene cloning strategies. *Electrophoresis* **11**: 537–553.
- Tempst, P. and L.R. Riviere. 1989. Examination of automated polypeptide sequencing using standard phenyl isothiocyanate reagent and subpicomole high performance liquid chromatographic analysis. *Anal. Biochem.* **183**: 290–300.
- Theissen, H., M. Etzerodt, R. Reuter, C. Schneider, F. Lottspeich, P. Argos, R. Lührmann, and L. Phillipson. 1986. Cloning of the human cDNA for the U1 associated 70K protein. *EMBO J.* **5**: 3209–3217.
- Utans, U. and A. Kramer. 1990. Splicing factor SF4 is dispensable for the assembly of a functional splicing complex and participates in the subsequent steps of the splicing reaction. *EMBO J.* **9**: 4119–4126.
- Wang, J. and T. Pederson. 1990. A 62,000 molecular weight spliceosome protein crosslinks to the intron polypyrimidine tract. *Nucleic Acids Res.* **18**: 5995–6001.
- Wassarman, D.A. and J.S. Steitz. 1991. Alive with DEAD proteins. *Nature* **349**: 463–464.
- Wieczorek, D., C.W.J. Smith, and B. Nadal-Ginard. 1988. The rat α -tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth and non-muscle isoforms by alternative splicing. *Mol. Cell. Biol.* **8**: 679–694.
- Wilusz, J. and T. Shenk. 1990. A uridylate tract mediates efficient heterogenous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. *Mol. Cell. Biol.* **10**: 6397–6407.
- Wu, J. and J.L. Manley. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes & Dev.* **3**: 1553–1561.
- Zamore, P.D. and M.R. Green. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl. Acad. Sci.* **86**: 9243–9247.
- . 1991. Biochemical characterization of U2 snRNP auxiliary factor: An essential pre-mRNA splicing factor with a novel intranuclear distribution. *EMBO J.* **10**: 207–214.
- Zhuang, Y. and A.M. Weiner. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**: 827–835.
- . 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes & Dev.* **3**: 1545–1552.
- Zillman, M., M.L. Zapp, and S.M. Berget. 1988. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **8**: 814–821.



Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing.

J G Patton, S A Mayer, P Tempst, et al.

Genes Dev. 1991, **5**:

Access the most recent version at doi:[10.1101/gad.5.7.1237](https://doi.org/10.1101/gad.5.7.1237)

References

This article cites 81 articles, 28 of which can be accessed free at:
<http://genesdev.cshlp.org/content/5/7/1237.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

A banner advertisement for Dharmacon Reagents and Horizon. On the left, it says 'Dharmacon Reagents' with the tagline 'Custom synthesis, RNAi, and CRISPR solutions'. In the center, the text 'Infinite Reliability' is displayed in large white font. To the right, the 'horizon' logo is shown in white, with 'a PerkinElmer company' underneath. A 'More' button is visible in the bottom right corner of the banner. The background features a colorful, abstract image of what appears to be a DNA double helix or a similar molecular structure.