

Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns

Mariano A. García-Blanco, Sharon F. Jamison, and Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

A protein of molecular size 62,000 daltons (p62) was detected in HeLa cell nuclear extracts by UV cross-linking to mRNA precursors. p62 binds specifically to the polypyrimidine tract of the 3' splice site region of introns. p62 purified to homogeneity binds the polypyrimidine tract of pre-mRNAs. This binding does not require the AG dinucleotide at the 3' splice site. Alterations in the polypyrimidine tract that reduce the binding of p62 yield a corresponding reduction in the efficiency of formation of a U2 snRNP/pre-mRNA complex and splicing. The p62 protein is retained in the spliceosome, where it remains bound to the pre-mRNA. This polypyrimidine tract binding protein (pPTB) is proposed to be a critical component in recognition of the 3' splice site during splicing.

[*Key Words:* Polypyrimidine tract; intron; 3' splice site; RNA-binding protein; UV cross-linking]

Received August 24, 1989; revised version accepted September 27, 1989.

Several RNA-binding proteins have been implicated in the splicing of nuclear mRNA precursors (pre-mRNAs). Most of the newly synthesized nuclear pre-mRNAs in mammalian cells are associated with heterogeneous ribonucleoproteins (hnRNPs) (Dreyfuss 1986). Both immunodepletion and antibody-inhibition experiments suggest that hnRNP C protein is important for splicing (Choi et al. 1986). This protein binds preferentially to sequences typical of the polypyrimidine tract upstream of the 3' splice site (Kumar et al. 1987; Swanson and Dreyfuss 1988). It also has significant affinity, however, for any single-stranded RNA. Similarly, the hnRNP A1 and D proteins bind preferentially to sequences typical of 3' splice sites. In fact, mutation of the highly conserved AG sequence at the 3' splice site reduces the binding of these proteins (Swanson and Dreyfuss 1988). Two apparently related proteins of 68,000 and 110,000 daltons, which bind specifically to the 3' splice site region, have been detected (Gerke and Steitz 1986; Tazi et al. 1986). These intron-binding proteins (IBPs) are recognized by Sm lupus patient sera and monoclonal antibodies and are thought to be associated with U5 small nuclear ribonucleoproteins (snRNP) (Gerke and Steitz 1986; Tazi et al. 1986). Interestingly, the binding of these proteins is also reduced if the AG sequence of the 3' splice site is altered.

Complementation assays have been used to identify factors important for either splicing in vitro or formation of specific complexes of substrate RNA and snRNPs (Krainer and Maniatis 1985; Krämer and Keller 1985;

Kramer 1988). For example, a soluble factor U2AF (U2 snRNP auxiliary factor) was partially purified on the basis of its requirement for efficient binding of U2 snRNP to pre-mRNA (Ruskin et al. 1988).

Until recently, it was thought that all of the conserved sequences at the 5' splice site are recognized by sequence complementarity to the U1 snRNA (Padgett et al. 1986). Recent results suggest that these base-pairing interactions are not sufficient to determine the precise 5'-cleavage site (Seraphin et al. 1988). This suggests that these sequences are also recognized by other components essential for splicing, perhaps sequence-specific binding proteins (Mayeda et al. 1986).

It has long been recognized that the formation of specific complexes containing U1 and U2 snRNPs, respectively, bound to the 5' splice site, and the branch site is an early step in the splicing of pre-mRNAs (Mount et al. 1983; Black et al. 1985; Konarska and Sharp 1986; Ruby and Abelson 1988). Depletion of nuclear extracts of U1 snRNP abolishes the activity of the 5' splice site in promoting the U2 snRNP/pre-mRNA association (Ruby and Abelson 1988; Seraphin et al. 1988). Recognition of both sites is important in the formation of each of these snRNA/pre-mRNA complexes. For example, mutational changes of sequences at the 5' splice site reduce the rate of formation of the U2 snRNP complex at the branch site (Lamond et al. 1987). Similarly, mutational changes at the branch site region reduce the rate of association of U1 snRNP at the 5' splice site (LeGrain et al. 1988; Ruby and Abelson 1988). Depletion of nuclear extracts of U2

snRNP, however, does not affect the activity of the branch site in promoting the U1 snRNP association with the 5' splice site (LeGrain et al. 1988; Ruby and Abelson 1988). Thus, it is likely that RNA sequence-specific binding proteins are important in recognition of the branch site at early stages in splicing and in promoting formation of complexes containing U1 and U2 snRNPs (Gerke and Steitz 1986; Mayeda et al. 1986; Steitz 1986; Tazi et al. 1986; Ruskin et al. 1988).

The total number of different proteins required for splicing of pre-mRNA is difficult to estimate. A large number of yeast mutants, temperature sensitive for growth, have been shown to be defective for splicing at the nonpermissive temperature (Rosbash et al. 1981; Lustig et al. 1986). In fact, 10 of the first 11 temperature-sensitive mutants isolated on the basis of deficiency in synthesis of mRNAs for ribosomal protein, thus deficient in synthesis of rRNA, have a defect in splicing as assayed in vitro (Hartwell et al. 1970; Rosbash et al. 1981; Lustig et al. 1986). Thus, the total number of gene products necessary for splicing activity may be very large. To date, the nature of the proteins encoded by several of these temperature-sensitive mutants has been characterized. Using immunological tools, Last and Woolford (1986) have shown that the protein products of the RNA 2 and RNA 3 genes are nuclear proteins of ~100,000 and 55,000 daltons, respectively. The RNA 8 gene specifies a 260,000-dalton protein that is complexed to U5 snRNA and, after incubation in the presence of ATP, also to U4 and U6 snRNAs (Lossky et al. 1987).

In this study we have identified a 62,000-dalton protein (p62) on the basis of a UV cross-linking assay, which binds specifically to the polypyrimidine tract of pre-mRNAs. Experiments indicate that the binding of this protein is required for the formation of the U2 snRNP complex. This protein has been purified to homogeneity.

Results

Detection of proteins binding pre-mRNAs in nuclear extracts

Proteins in HeLa cell nuclear extracts that bound to the adenovirus pre-mRNA (Ad10 RNA) were detected by a UV cross-linking assay (Shoemaker and Schimmel 1974). Uniformly labeled RNAs were incubated with nuclear extracts under splicing conditions, and proteins were cross-linked to RNA by irradiation at 254 nm. After digestion of RNA by nuclease, individual RNA-protein adducts were resolved by SDS-PAGE (Fig. 1). Many different proteins were cross-linked to the Ad10 pre-mRNA (lanes 2–4), as well as to a length-matched control RNA (pBS) derived from the plasmid pBS⁺ (lane 5). There was, however, one protein of 62,000 daltons (p62) that cross-linked specifically to the pre-mRNA. Longer exposures of autoradiographs showed that a second protein of apparent molecular mass >200,000 daltons, was also specifically cross-linked to the pre-mRNA (lane 3; >200 K). The hnRNP C proteins, identi-

fied by reactivity with a specific monoclonal antibody bound both the pre-mRNA and pBS RNA (Fig. 1 and data not shown).

Generation of distinct bands depended on addition of HeLa cell nuclear extract (Fig. 1, lane 1) and UV irradiation (data not shown). Cross-linked products were shown to represent protein-RNA adducts by protease and nuclease digestion studies (data not shown).

The p62 protein bound the pre-mRNA very rapidly under reaction conditions optimal for splicing; the extent of binding was maximal by 1 min and remained constant through 3 hr of incubation (data not shown). In contrast to the splicing reaction, p62 binding did not require the addition of exogenous ATP (Fig. 1). The lower signal observed in the presence of ATP is due to lower stability of the RNA. Both p62 and hnRNP C proteins bound pre-mRNA after incubation at 4°C (data not shown). The products designated as hnRNP C protein migrate as a doublet and probably correspond to the previously described C1 and C2 polypeptides (Dreyfuss 1986). The p62 product was also resolved into a doublet and, in some rare occasions, a triplet of bands. The biochemical nature of this variation in p62 is unclear, although it is not an artifact of the cross-linking assay, as will be shown below.

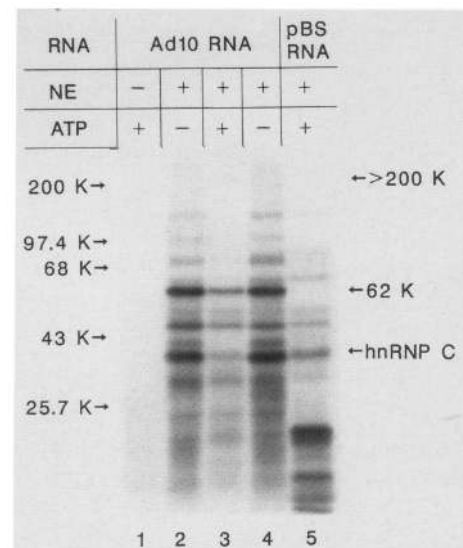


Figure 1. UV cross-linking of HeLa cell nuclear extract proteins to an adenovirus pre-mRNA. A uniformly labeled adenovirus pre-mRNA (Ad10 RNA) or a size-matched RNA (pBS RNA) was incubated with HeLa cell nuclear extracts under splicing conditions. (Top) Incubated RNA. (NE) The addition of nuclear extract; (ATP) the addition of 1 mM ATP and 5 mM creatine phosphate. The incubated reactions were exposed to 254-nm UV light and treated with RNase A (1 mg/ml), as described in Methods. The UV cross-linked RNA-protein products were resolved by SDS-PAGE. Molecular weight markers are myosin (200,000 daltons) phosphorylase b (97,400 daltons) bovine serum albumin (BSA) (68,000 daltons) ovalbumin (43,000 daltons) and α -chymotrypsinogen (25,700 daltons). Three cross-linked proteins (M_r >200,000, M_r 62,000 and hnRNP C) are indicated.

Several RNA-binding proteins have been well characterized, and monoclonal antibodies specific for these proteins have been isolated. The potential relationship of these proteins to p62 was tested by immunoprecipitation. Anti-hnRNP C (41–43 kD), anti-hnRNP M (68 kD), anti-hnRNP L (68 kD), and anti-Sm monoclonal antibodies all immunoprecipitated unique UV cross-linked products, but none immunoprecipitated the p62 (see Methods; data not shown).

p62 specifically binds pre-mRNAs

The specificity of binding of p62 to the Ad10 splicing precursor was investigated. First, other ³²P-labeled RNAs were tested as potential substrates for cross-linking to the protein. Second, unlabeled RNAs were tested for competition for p62 binding to the Ad10 substrate.

Two efficiently spliced pre-mRNAs, Ad10 and a rabbit β -globin pre-mRNA (Lamond et al. 1987), were cross-linked efficiently to p62 (data not shown). As mentioned previously, the pBS RNA could not be cross-linked to this protein. A substrate RNA for polyadenylation in vitro (Skolnik-David et al. 1987) was not cross-linked to p62 (data not shown). The polyadenylation substrate was efficiently cross-linked to the hnRNP C protein and to a protein of 68 kD, which is probably that described by Wilusz and Shenk (1988) and Moore et al. (1988) (data not shown).

As a test for specificity, which did not require activity

in photo cross-linking, labeled Ad10 mRNA was incubated with nuclear extract in the presence of unlabeled competitor RNAs. Ad10 RNA (self-competition and β -globin pre-mRNA competed with the cross-linking of labeled Ad10 RNA to p62 (Fig. 2A,B). The pBS RNA, even at 3000-fold molar excess, did not compete (Fig. 2A). Similarly, the polyadenylation substrate did not compete (data not shown). Finally, tRNA and total cytoplasmic RNA from HeLa cells, when added at >1000-fold molar excess, did not compete with the cross-linking of p62 to the Ad10 substrate (data not shown). The result with the fibronectin pre-mRNA shown in Figure 2B will be discussed below.

p62 recognizes sequences in the vicinity of the 3' splice site

Deletions of the Ad10 RNA were used to map the binding site of p62. A series of deletions in the 3' end of the Ad10 RNA distinguish an element critical for cross-linking to the p62 protein (RNAs 1–4 in Fig. 3A). This critical element was located between 61 nucleotides upstream of the 3' splice site and 4 nucleotides in the flanking exon. Both an RNA lacking the first exon and the 5' half of the intron (RNA 5 in Fig. 3A) and a shorter RNA containing the 3' half of the intron and four nucleotides of the second exon (RNA 6 in Fig. 3A) bound p62.

Fortunately, there are two unique RNase T1-resistant oligonucleotides in the 3' half of the intron: a 14-nu-

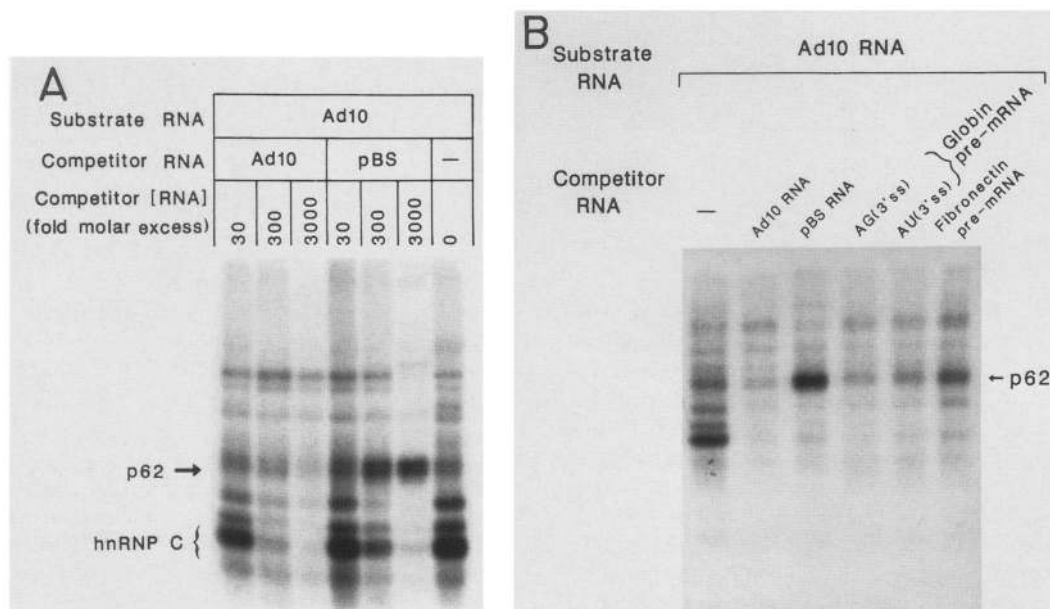


Figure 2. The binding of p62 to pre-mRNAs is specific. (A) Labeled Ad10 RNA was incubated with HeLa cell nuclear extracts, as described in Fig. 1, except that a 1 : 20 dilution of the nuclear extract was used in the presence of no competitor RNA or 30-, 300-, and 3000-fold molar excess of Ad10 RNA or pBS RNA, respectively. p62 and hnRNP C protein are indicated. (B) Uniformly labeled Ad10 RNA was incubated with HeLa cell nuclear extract, as described in A. The lanes represent reactions that also contained a 300-fold molar excess of competitor unlabeled RNAs. The unlabeled RNAs were Ad10 and pBS RNAs (as described in Fig. 1), rabbit β -globin pre-mRNAs [a wild-type and a 3' splice site mutant (AG \rightarrow AU)] and a fibronectin pre-mRNA (see Methods). p62 and hnRNP C protein are indicated.

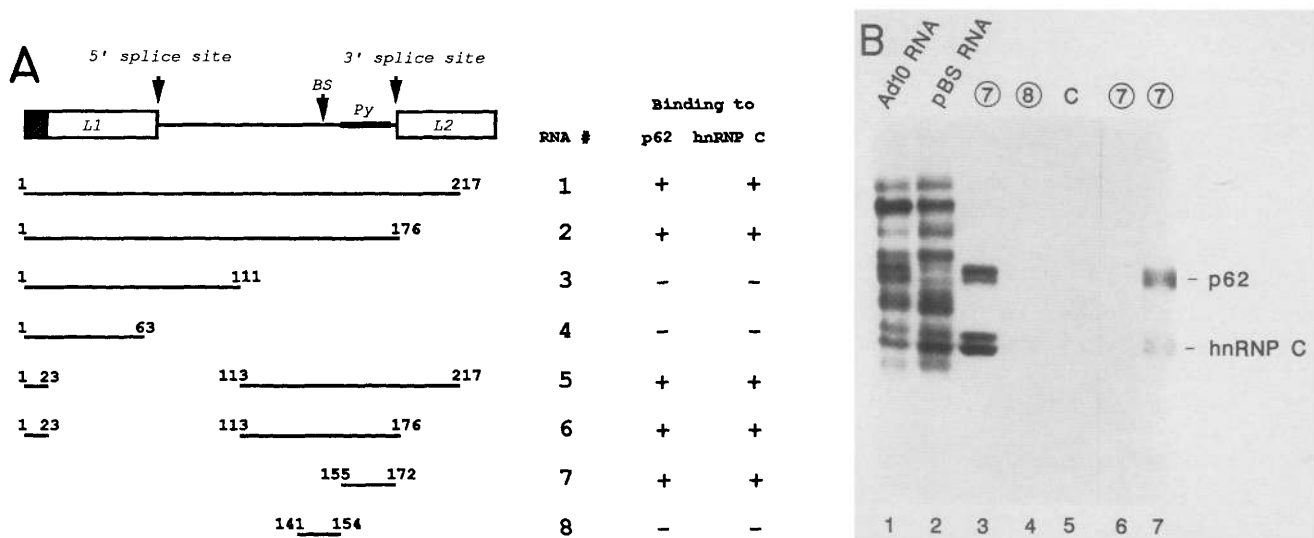


Figure 3. Mapping of the binding site of p62 on the Ad10 RNA. (A) Ad10 RNA and shorter RNAs spanning portions of Ad10 RNA were incubated with HeLa cell nuclear extracts and processed as described in Fig. 1. The efficiency of UV cross-linking to p62 and hnRNP C protein is indicated. RNAs 1–4 are derived from vector pBSAd10 by truncating the template with *Sau3A*, *FnuD2*, *HhaI*, and *HaeIII*, respectively. RNAs 5 and 6 are derived from pBSAd13 by truncating the template with *Sau3A* and *FnuD2*, respectively. RNAs 7 and 8 are RNase T1-resistant oligonucleotides derived from Ad10 RNA (RNA 1). (Top left) A schematic representation of the splicing unit of Ad10 RNA. The boxes represent exon sequences; the shaded area indicates polylinker sequences in the pBS⁻ vector. (L1 and L2) Leader 1 and 2 sequences of the adenovirus 2 major late transcription unit; (BS) branch site; (Py) polypyrimidine tract. The numbers bordering the transcript icons indicate the equivalent nucleotide number in the Ad10 RNA. (B) Uniformly labeled Ad10 RNA (RNA 1), pBS RNA, or RNase T1 oligonucleotides (RNAs 7, 8, and C) were incubated with HeLa cell nuclear extract in the absence (lanes 1–5) or presence of competitor unlabeled Ad10 RNA (lane 6) or pBS RNA (lane 7). A 1 : 20 dilution of the nuclear extract was used in lanes 6 and 7. Circled numbers 7 and 8 refer to RNA numbers in A; (7) An 18-mer (5'-UCCCUUUUUUUC-CACAG-3'); (8) a 14-mer (5'-UCAUACUUAUCCUG-3'); (C) a 12-mer (5'-ACAAACUCUUCG-3').

cleotide fragment (5'-UCAUACUUAUCCUG-3') encompassing the branch site (RNA 8 in Fig. 3A) and an 18-nucleotide fragment (5'-UCCCUUUUUUUC-CACAG-3') encompassing the polypyrimidine tract and the AG dinucleotide at the 3' splice site (RNA 7 in Fig. 3A). Three RNase T1-resistant oligonucleotides were purified from the Ad10 substrate and tested for cross-linking to the p62 protein. The 18-nucleotide fragment containing the polypyrimidine tract and the 3' splice site, but not the 14-nucleotide fragment, cross-linked to p62 (Fig. 3A,B, lanes 3 and 4). A 12-nucleotide (5'-ACAAACUCUUCG-3') RNase T1-resistant fragment from the Ad10 RNA (C in Fig. 3B) did not cross-link to p62. As a test of the specificity of this oligonucleotide assay, p62 cross-linking to the 18-nucleotide fragment was competed by unlabeled Ad10 RNA but not by unlabeled pBS RNA (Fig. 3B, lanes 6 and 7).

Sequences important for efficient cross-linking to the hnRNP C protein also mapped to the same 18-nucleotide fragment at the 3' splice site. This is in agreement with the work of others (Kumar et al. 1987; Swanson and Dreyfuss 1988). As expected, hnRNP C cross-linking to this 18-nucleotide fragment was competed by both Ad10 and pBS RNAs (Fig. 3B, lanes 6 and 7).

We constructed plasmids (pBPY-1 through pBPY-3) in which either the Ad10 RNA 18-mer oligonucleotide or two mutant oligonucleotides were inserted in the polylinker of the pBS⁻ vector (see Methods). This permits

the synthesis of RNAs containing the Ad10 RNA polypyrimidine tract and 3' splice site or mutants of these sequences. The rest of the sequence of the BPY RNAs is totally unrelated to Ad10 RNA. The 35-nucleotide BPY-1 RNA, containing the Ad10 RNA 18-mer sequence (5'-UCCCUUUUUUUC-CACAG-3') cross-linked to p62, thus demonstrating again that this 18-mer was sufficient for binding (data not shown). Unlabeled BPY-1 RNA and a 35-nucleotide BPY-3 RNA, which contained a C-rich polypyrimidine tract (5'-UCCCUCCCUCC-CUCCACAG-3') both competed efficiently with Ad10 RNA for binding of p62 (data not shown).

Poly(I), poly(U), and poly(AUC) were tested as competitors of p62 binding. These RNAs did not compete with Ad10 RNA for p62 binding even at 100 µg/ml. This represents a 2000-fold mass excess. In contrast, poly(U) at a concentration of 100 µg/ml competed with Ad10 RNA for hnRNP C protein binding (data not shown).

The importance of the AG dinucleotide at the 3' splice site for p62 binding was tested with a previously characterized β-globin mutant (the AG at 3' splice site was converted to AU; Lamond et al. 1987). Substrate RNA from this mutant template was efficiently cross-linked (data not shown) and competed for cross-linking of p62 to the Ad10 RNA, as well as wild-type β-globin premRNA (Fig. 2B). A 35-nucleotide RNA containing the Ad10 RNA polypyrimidine tract and a mutated 3' splice site (5'-UCCCUUUUUUUC-CACCC-3') was tran-

scribed from plasmid pBPY2 (see Methods). This RNA cross-linked to p62 and competed efficiently with Ad10 RNA for cross-linking to p62 (data not shown). Therefore, the AG dinucleotide at the 3' splice site is not an important determinant for p62 binding.

p62 is bound to pre-mRNA in splicing complexes

Pre-mRNAs have been shown to interact with snRNPs and proteins in nuclear extracts to form splicing complexes (Brody and Abelson 1985; Grabowski et al. 1985) that can be separated by gel electrophoresis (Konarska and Sharp 1986, 1987; Pikielny et al. 1986; Lamond et al. 1988). Three complexes (A,B,C) are discernible using Tris-glycine native gels (Fig. 4A). A two-dimensional gel system was used to determine whether or not the p62 protein was in contact with the substrate RNA in these complexes. The complexes were separated on a native polyacrylamide gel (first dimension), UV cross-linked, digested with nuclease in situ, and subjected to electrophoresis in a polyacrylamide gel containing SDS (second dimension; see Methods). The cross-linked p62 product was generated by both the heterogeneous complex (H) and the splicing complexes (A–C) (Fig. 4B). When the reaction was incubated in the absence of ATP, p62 was associated with the H complex, as was expected (data not shown).

Surprisingly, only a very low level of hnRNP C protein was cross-linked to the pre-mRNA using the two-dimensional protocol. In contrast, efficient cross-linking was observed with the unfractionated aliquot of the same reaction loaded on the native gel (Fig. 4B, lane 2). It is likely that the Tris-glycine gel condition perturbs the interactions between hnRNP C protein and the RNA. This perhaps accounts for the enhanced resolution of this gel system.

A protein of ~50,000 daltons (50 kD) was cross-linked to pre-mRNA only in the splicing complexes (Fig. 4B). A protein of this molecular weight had not been detected as a specific RNA-binding product in analysis of total solution reactions. The low level of signal for this protein could have been obscured in analysis of the total reactions (+ ATP lane; Fig. 4). We have observed, however, a UV cross-linked protein of 50 kD that can be detected by immunoprecipitation with the Y12 Sm-type monoclonal antibody (data not shown).

p62 binding and U2 snRNP complex formation

A pre-mRNA derived from a fibronectin intron, which is alternatively spliced in vivo, was used in the above assays (Schwarzbauer et al. 1987). The fibronectin pre-mRNA was inactive for splicing in vitro both in regard to the formation of any of the splicing complexes, A–C,

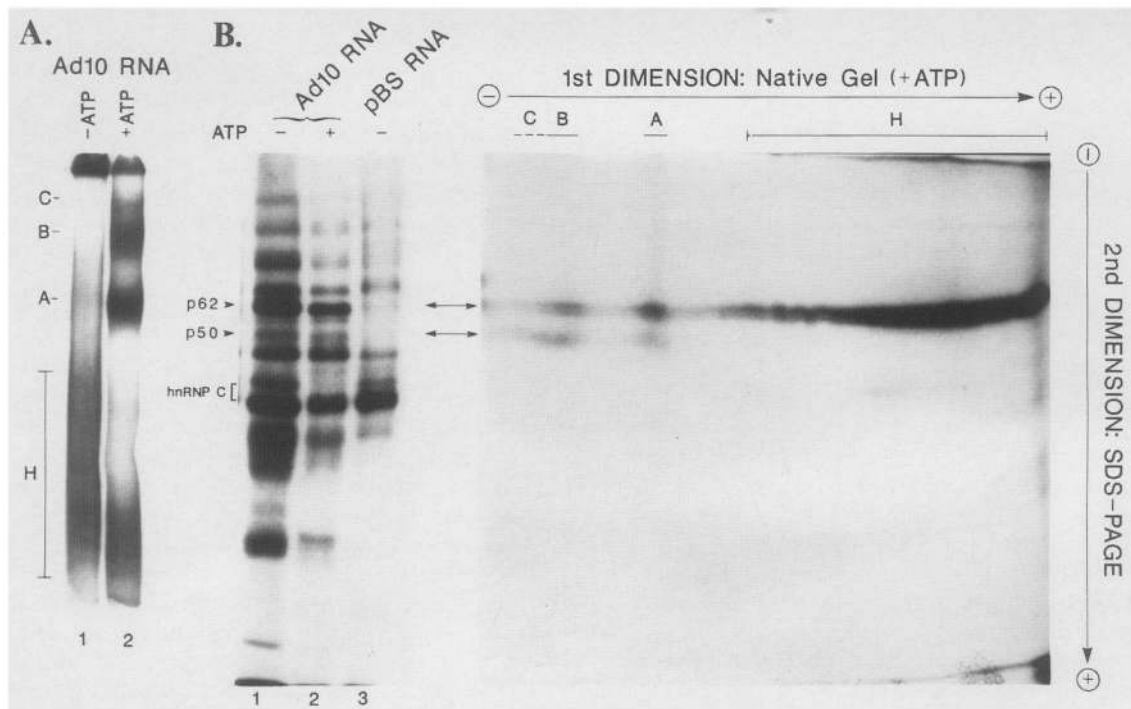


Figure 4. p62 is bound to pre-mRNA in splicing complexes. Labeled Ad10 RNA was incubated with HeLa cell nuclear extract under splicing conditions. A preparative splicing reaction was loaded on a native Tris-glycine gel. The gel was UV-irradiated to cross-link RNA and protein in situ, digested with RNase A, equilibrated with SDS-PAGE sample buffer, and embedded in the stacking gel of an SDS-PAGE gel (see Methods). (A) Native gel showing the pattern of splicing complexes (lane adjacent to that used in the second dimension). Nonspecific complex (H) and splicing-specific complexes (A–C) are indicated. (B) RNA–protein cross-linked products separated by one-dimensional SDS–PAGE and by the two-dimensional procedure. Lanes 1–3 and the two-dimensional panel were from the same SDS–PAGE gel. p62, p50, and hnRNP C proteins are indicated; comigrating spots in the second dimension are indicated with bidirectional arrows.

and the generation of the lariat intron and spliced exons (data not shown). This pre-mRNA did not cross-link directly to p62 (data not shown) and did not compete with the Ad10 substrate for cross-linking to p62 (Fig. 2B). Moreover, a human immunodeficiency virus (HIV) pre-mRNA (see Fig. 6), which formed splicing complexes very inefficiently, if at all, did not cross-link to p62 (D. Chang and M. García-Blanco, unpubl.) These observations suggested that p62 binding may be important in the formation of the U2 snRNP complex (complex A).

To test the correlation between p62 binding and U2 snRNP complex formation, a pre-mRNA-encoding plasmid was constructed, pPIP3: precursor in pieces. The sequence of pPIP3 was derived from that of pBSAd10 and is shown in Figure 5. pre-mRNA (PIP3 RNA) derived from this plasmid was spliced *in vitro* and formed the anticipated splicing complexes (see Fig. 7A, below). pPIP3 has unique restriction endonuclease sites embedded within the region encoding the PIP3 RNA, allowing the easy substitution of mutant sequences within RNA substrates (Fig. 5). The PIP3 RNA was shown to bind p62 as was expected, given that it shares all but the first uridine of the 18-nucleotide-long fragment of Ad10 RNA. Several mutants of pPIP3 were constructed that encode pre-mRNAs with mutated polypyrimidine tract/3' splice site regions (Fig. 6).

Inspection of the polypyrimidine tract/3' splice site sequences of the fibronectin and HIV pre-mRNAs (FNE and HIV in Fig. 6) revealed several elements that differed from the corresponding regions of Ad10 and β -globin RNAs (wt Ad10 and wt β -gl in Fig. 6). A mutant (pPIP3py4) was constructed where the polypyrimidine tract of pPIP3 was substituted by that present in the fibronectin pre-mRNA (Fig. 6). PIP3py4 did not form a U2 snRNP complex, and it did not cross-link efficiently to

p62 (Fig. 7). This suggests that the polypyrimidine tract of the fibronectin pre-mRNA is not efficiently recognized by p62, and this may be responsible for the inefficient formation of splicing complexes.

A second mutant of pPIP3 was constructed by deleting 6 nucleotides in the polypyrimidine tract/3' splice site segment of pPIP3 (see pPIP3py2 in Fig. 6). This mutant reduces the length of the polypyrimidine tract from 13 nucleotides in PIP3 RNA to 9 in PIP3py2 RNA. Furthermore, this mutant changes the sequence immediately upstream of the invariant 3' splice site AG dinucleotide. PIP3py2 RNA did not form a complex with U2 snRNP and did not cross-link to p62 (Fig. 7).

A third mutant, pPIP, was constructed by replacing three uridine residues with adenosines in the polypyrimidine tract of pPIP3 (see U2 pPIP3py6 in Fig. 6). The sequence around the 3' splice site was not changed from that in pPIP3. The pPIP3py6 RNA formed U2 snRNP complex very inefficiently and did not bind p62 (Fig. 7). The binding to hnRNP C protein was also diminished (Fig. 7).

A mutant pre-mRNA was constructed that modified the sequences immediately upstream of the invariant AG dinucleotide but did not change the sequence of polypyrimidine tract (pPIP3py5 in Fig. 6). pPIP3py5 RNA did not form a complex with U2 snRNP; however, this RNA was recognized, albeit less efficiently than pPIP3 RNA, by p62 (Fig. 7). These results with pPIP3py5 RNA are similar to those obtained with mutants in the 3' splice site AG dinucleotide in β -globin RNA, which bind p62 but do not form U2 snRNP complex efficiently (Fig. 7; Lamond et al. 1987). Taken together, these data strongly suggest that binding to p62 is required but is not sufficient for efficient formation of the U2 snRNP complex.

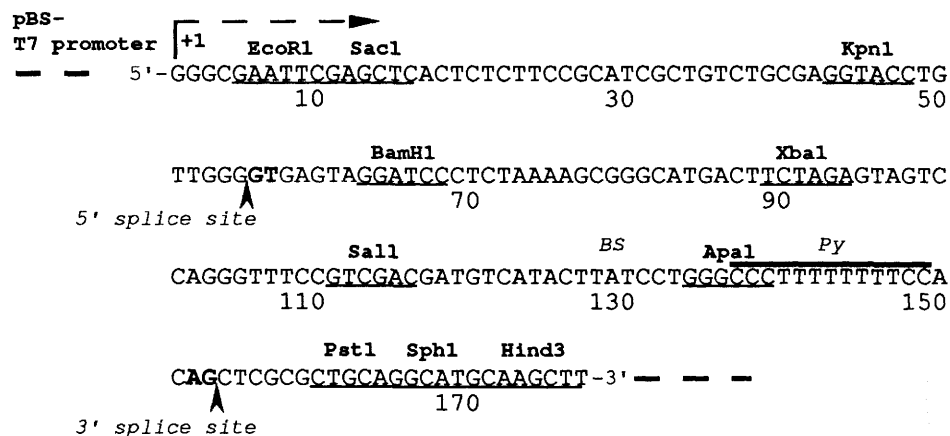


Figure 5. Construction of pPIP3. The plasmid pPIP3 was constructed by the sequential addition of pairs of adjacent double-stranded DNA oligonucleotides. These oligonucleotides were cloned between the *SacI* and *PstI* sites of pBS⁻ and were designed so that the ends would reconstruct enzyme-restricted ends. The six double-stranded DNA oligonucleotides contained the following pPIP3 sequences: *SacI*–*KpnI*, *KpnI*–*BamHI*, *BamHI*–*XbaI*, *XbaI*–*SalI*, *SalI*–*ApaI*, *ApaI*–*PstI*, respectively. The sequence shown is equivalent to the pre-mRNA synthesized from the T7 promoter in *HindIII*-linearized pPIP3 (5' end is at +1). Unique restriction sites are underlined and indicated in boldface type. Landmarks in the pre-mRNA encoded by pPIP3 are indicated: 5' splice site, 3' splice site, branch site (BS), and polypyrimidine tract (Py). The invariant dinucleotides GT and AG, at the 5' and 3' splice sites, respectively, are indicated in boldface type.

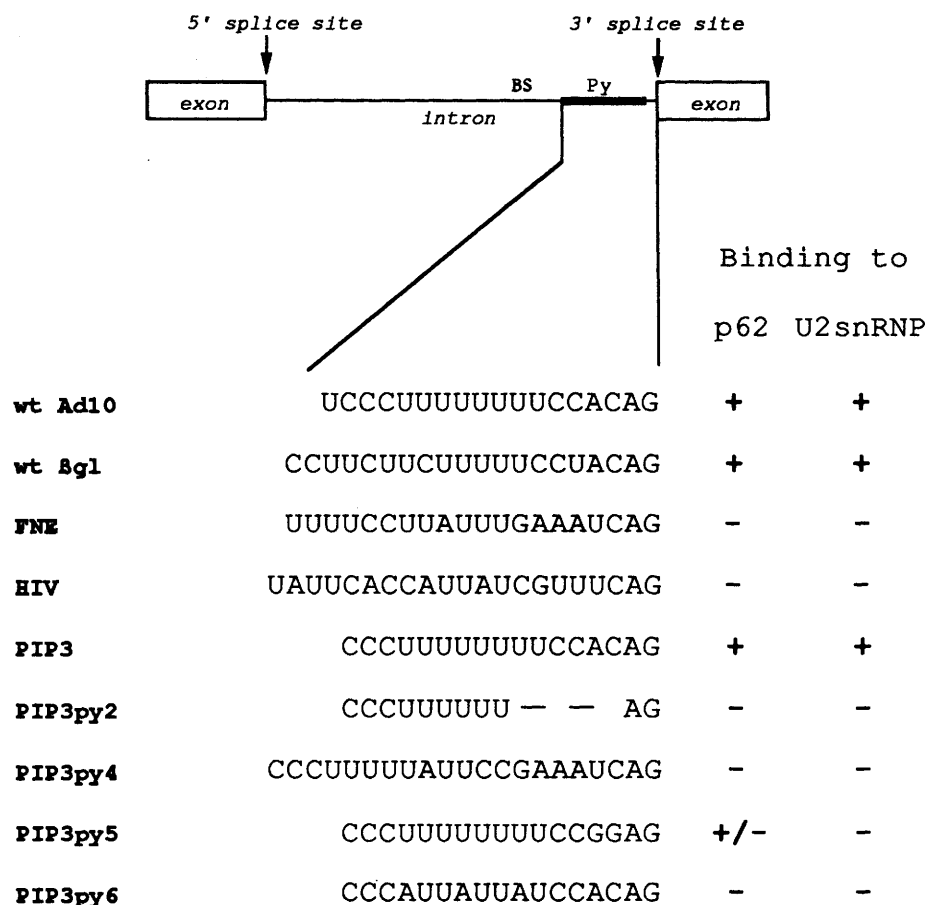


Figure 6. p62 binding and formation of complex containing U2 snRNP with wild-type and mutant pre-mRNAs. The polypyrimidine tract and sequences up to and including the invariant AG dinucleotide at the 3' splice site are shown for various pre-mRNAs. The DNA sequence for pBSAd10 and all pPIP-derived plasmids were determined by the Sanger dideoxy method (Sanger et al. 1977), using a primer complementary to the T3 promoter in pPIP3 and pBSAd10. The sequence of pBSAL4 [template for the wild-type β -globin RNA (β gl)] was published by Lamond et al. (1987). The sequence of pBi8.3 (template for the FNE RNA) was communicated by P. Norton (pers. comm.). The sequence of pHXB3 (template for the HIV RNA; D. Chang, unpubl.) is that of the parental plasmid pgtat (Malim et al. 1988). The HIV RNA spans the *tat*-coding exons and the large intron that separates them in the HIV-1 genome. A schematic representation of a general splicing unit is shown above. Cross-linking to p62 and the formation of the U2 snRNP complex by these pre-mRNAs are indicated (see Fig. 7).

Purification of p62

Using UV photo-cross-linking as the assay, p62 was purified to homogeneity by column chromatography using sequentially DEAE-Sephacel, heparin-agarose, and poly(U)-agarose matrices (for technical details, see Methods and Table 1). Several preliminary experiments indicated that >90% of the detectable p62 eluted in the flowthrough and washes (100 mM KCl) of DEAE-Sephacel (not shown). In the purification scheme presented in Table 1, this fraction (FT/W) was immediately loaded on an equilibrated heparin-agarose matrix. After extensive washing, the matrix was developed with a linear salt gradient. p62 eluted, together with many other RNA-binding proteins, at a KCl concentration of ~300 mM (Fig. 8A). The first two steps of this scheme gave a 140-fold purification and a yield of 411%. The yield value >100% is probably due to purification of p62 away from nucleic acids in the nuclear extract. The p62-con-

taining fractions were diluted with buffer to reduce the salt concentration to 100 mM KCl, and chromatographed on poly(U)-agarose matrix. After extensive washes, the matrix was developed with a linear salt gradient, and p62 eluted at 600 mM KCl as a homogeneous protein (Figs. 8B and 9A).

The purified protein resolved into a doublet of bands following electrophoresis in SDS-polyacrylamide gels and silver staining (Fig. 9A). Most of the tryptic peptides obtained from these two proteins coeluted from a reverse-phase high performance liquid chromatography (HPLC) matrix, indicating that the two proteins share most of their sequence (data not shown). This doublet corresponds to the doublet observed in the initial cross-linking pattern of p62 in the crude nuclear extract used as the starting material for the purification (see Fig. 2A). In one experiment, homogeneous p62 was cross-linked to labeled Ad10 RNA, and the products were resolved on an SDS-PAGE gel that was both silver-stained and ex-

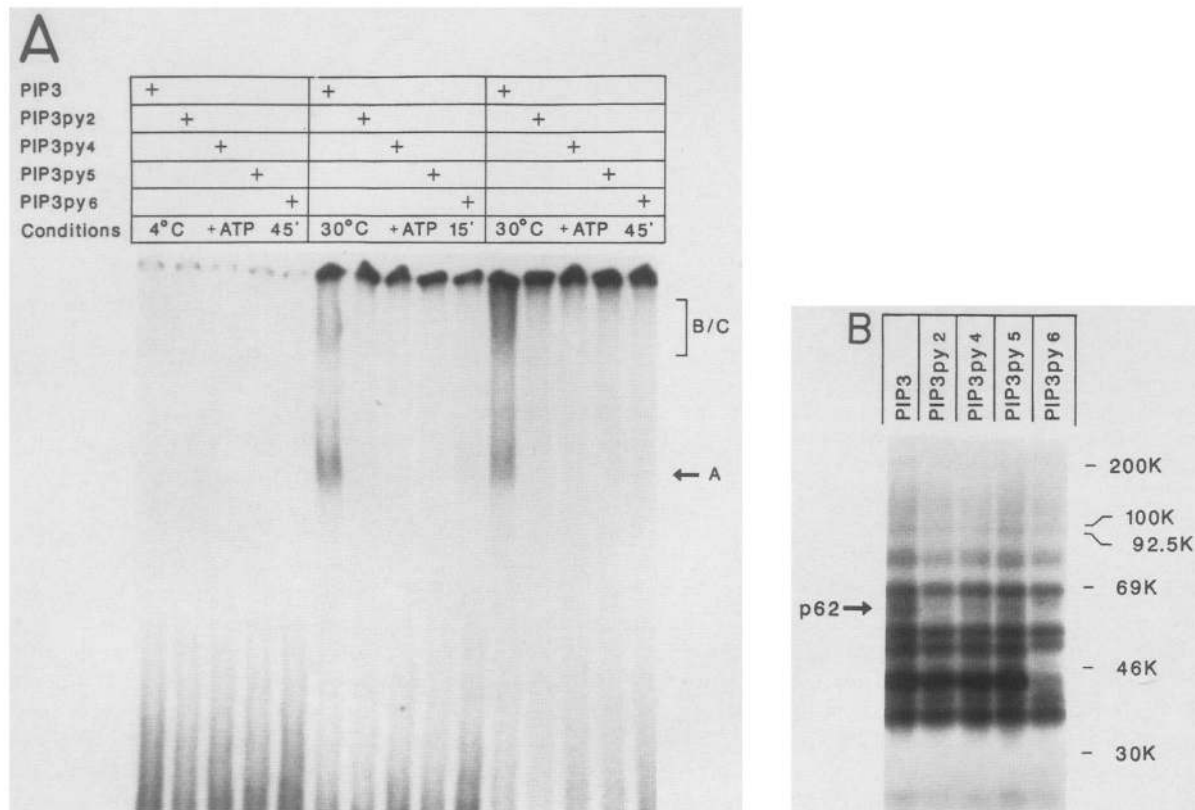


Figure 7. p62 binding and U2 snRNP complex formation. (A) Labeled PIP3 RNA and PIP3 mutant RNAs were incubated with HeLa nuclear extract in buffer optimal for splicing either at 4°C for 45 min or at 30°C for 15 or 45 min (*top*). Splicing complexes were resolved by native gel electrophoresis. The U2 snRNP complex (A) and larger splicing complexes (B/C) are indicated. (+) Incubated RNA. (B) A fraction of the splicing reactions incubated at 30°C for 45 min (described in A) was irradiated with 254-nm UV light and processed as described in Methods. The RNA incubated is indicated above the corresponding lane. The autoradiogram of the SDS-polyacrylamide gel is shown, and p62 is indicated. Molecular weight markers are myosin (200,000 daltons), phosphorylase b (100,000 and 92,500 daltons), BSA (69,000 daltons), ovalbumin (46,000 daltons), and carbonic anhydrase (30,000 daltons).

posed to film. The silver-stained bands in the doublet comigrated exactly with the radiolabeled bands (data not shown), demonstrating that both polypeptides have RNA binding activity.

The purified p62 protein had RNA binding properties similar to that of the p62 product in the initial extracts. Unlabeled Ad10 RNA competed efficiently with the cross-linking of Ad10 RNA to homogeneous p62, whereas pBS RNA, even at 1000-fold molar excess, did not significantly compete with the binding of homogeneous p62 (Fig. 9B). Moreover, purified p62 cross-linked to the labeled RNase T1-resistant 18-nucleotide RNA generated from sequences spanning the polypyrimidine tract and the 3' splice site of Ad10 RNA (data not shown). Importantly, the purified p62 was shown to bind the polypyrimidine tract in the absence of other RNA-binding proteins, particularly hnRNP C protein.

Discussion

A 62,000-dalton protein (p62) that binds to the polypyrimidine tract of the 3' splice site region of introns has been identified and purified. Recognition of the pre-

cursor RNA by this protein may partially account for the critical role of the polypyrimidine tract in specifying the splicing of RNA in mammalian cells.

Binding of p62 to the polypyrimidine tract was detected by UV cross-linking of labeled RNA to cellular proteins. The specificity of binding was also determined by comparing the efficiency of competition of various RNAs for formation of the cross-linked product. Polypyrimidine tracts from two substrate RNAs that are efficiently spliced *in vitro*—the first intron of adenovirus major late transcription unit (5'-UCCCUUUUUUUC-CACAG-3') and the second intron of the rabbit β -globin (5'-CCUUCUUCUUUUUCCUACAG-3') are readily bound by p62. Similarly, a C-rich tract (5'-UCCCUCC-CUCCCUCCACAG-3') is also efficiently recognized. The specificity of p62 recognition, however, is not simply determined by a polypyrimidine tract, as poly(U) does not efficiently compete for p62 binding. At the moment, it is difficult to definitively assign the sequence specificity of p62 binding. The protein does not bind efficiently to RNAs with polypyrimidine tracts shorter than 10 nucleotides, e.g., 5'-CCCUUUUUUAG-3', nor to tracts interrupted with purines, e.g., 5'-

Table 1. Purification of p62

Step	Total protein ^a (mg)	Specific activity ^b (cross-linked p62/mg of total protein)	Yield ^b (%)
Nuclear extract	136.0	1.0	100
DEAE–Sephacel	19.0	ND	ND
Heparin–agarose	4.0	140.0	411.0
Poly(U)–agarose	0.025 ^c	2000.0 ^c	35.5

^aTotal protein was determined using the Bio-Rad assay, following the recommendations of the manufacturer.

^bThe level of UV cross-linking was determined by scanning autoradiograms with an LKB soft laser scanner. Quantification of cross-linking of p62 to Ad10 RNA was obtained by cross-linking serial dilutions of the samples. This resulted in arbitrary units of cross-linking to p62 for a given volume which, divided by the milligrams of total protein in this volume, gave the specific activity. The level of cross-linked p62 for the total volume in a pooled fraction was the yield. The cross-linking assays for starting material and purified fractions were all repeated in the same experiment to accurately compare levels of cross-linking.

^cAmino acid analysis of the purified protein suggested that the amount of protein shown may have been a 3- to 5-fold overestimate; thus, the purification could have been as high as 10,000-fold.

CCCAUUAUUAUCCACAG-3' or 5'-CCCUUUUUAU-UCCGAAAUCCAG-3'. The binding specificity of p62 does not apparently extend to the AG of the 3' splice site, which can be altered without a reduction in binding affinity. Alteration of the next 2 nucleotides in the 3' splice site, ACAG → GGAG, partially reduces the binding affinity of p62. Thus, it is possible that p62 recognizes a long polypyrimidine tract in conjunction with the AC sequence at the 3' splice site. This specificity of binding of p62 is interesting in the context of efficient splicing of substrate RNA in vitro.

The polypyrimidine tract and highly conserved PyAG have been shown to be the major determinants of the 3' splice site in mammalian introns (Padgett et al. 1986). In contrast, the conserved sequences at the branch site and the AG dinucleotide at the 3' splice site in yeast appear to be the major determinants (Padgett et al. 1986). Con-

sistent with this, there is little evidence of an enriched polypyrimidine tract in the 3' region of most yeast introns. One example of the importance of the polypyrimidine tract in mammalian introns is that the efficiency of splicing in vivo of an intron of rabbit β-globin decreased >13-fold when a deletion shortened the tract from 16 to 8 nucleotides (Wieringa et al. 1984). Consistent results have been reported with other mutations in the polypyrimidine tract (Ruskin and Green 1985). The best evidence relating this dependence of splicing on the polypyrimidine tract to binding of p62 emerged from analysis in vitro of substrate RNAs from mutant templates. Shortening of the polypyrimidine tract or substitutions of adenosine residues for pyrimidine residues coordinately reduced p62 binding, formation of U2 snRNP complex, and splicing in vitro. Similarly, substrate RNAs from introns of fibronectin and HIV were ineffi-

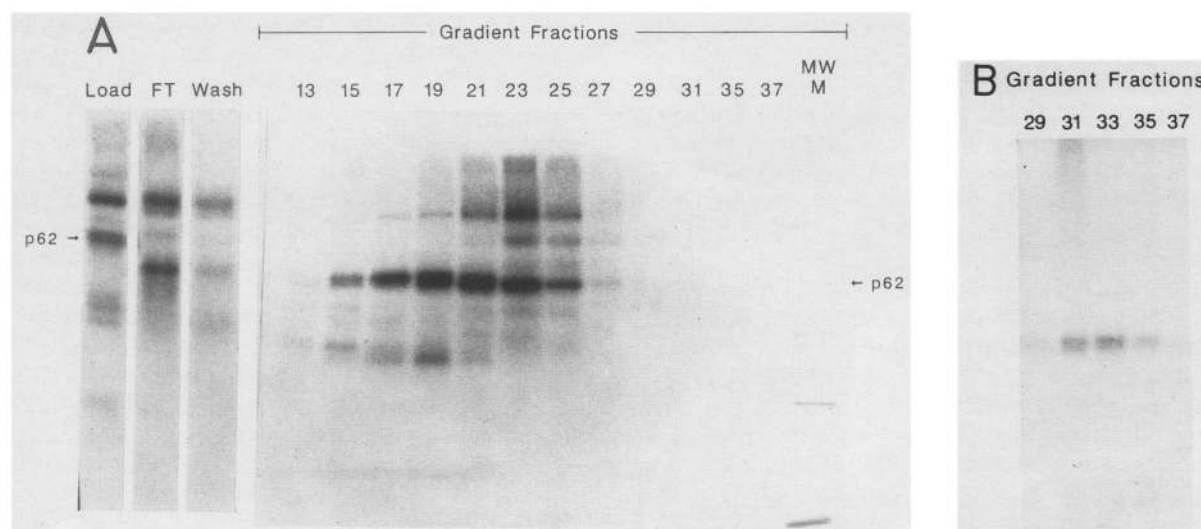


Figure 8. Purification of p62. p62 was purified through DEAE–Sephacel, heparin–agarose, and poly(U)–agarose matrices, and activity was followed by UV cross-linking of the eluted fractions to Ad10 RNA (see Table 1). The cross-linked products obtained with fractions 13–37 eluted from heparin–agarose (A) and fractions 29–37 eluted from poly(U)–agarose (B) were resolved by SDS–PAGE and are shown. (MWM) Molecular weight markers (described in Fig. 7B).

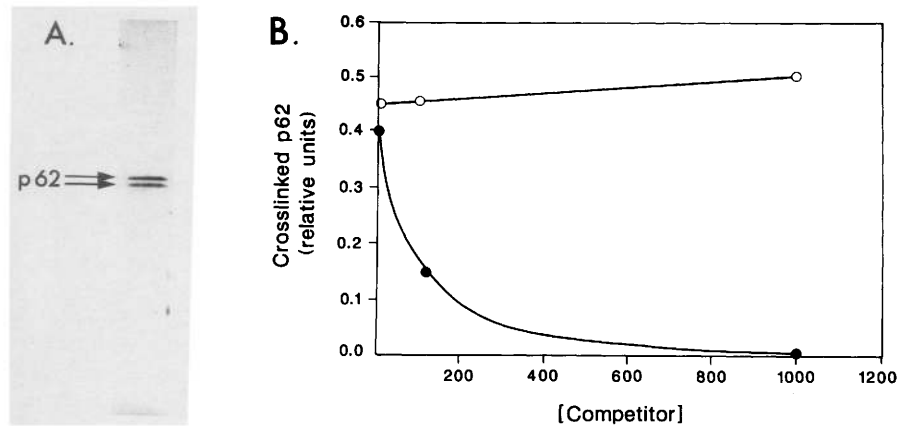


Figure 9. Properties of purified p62. (A) When p62 fraction 33 of the elution gradient from poly(U)-agarose (see Fig. 8B) was resolved by SDS-PAGE and detected by silver staining (Blum et al. 1987), a homogeneous p62 doublet was observed. (B) Fraction 33 from poly(U)-agarose was incubated with labeled Ad10 RNA in the presence of varying amounts of unlabeled Ad10 RNA (●) or pBS⁻ RNA (○). The level of cross-linked p62 was plotted vs. the concentration of competitor RNA, given in fold molar excess over labeled Ad10 RNA.

ciently spliced *in vitro* and were inefficiently recognized by p62. Thus, in general, all substrate RNAs with an intact CAG at the 3' splice site that were bound by p62 also efficiently formed the U2 snRNP complex and were spliced *in vitro*.

Recognition of the 3' splice site region during splicing involves more than the binding of p62. A stable complex of U2 snRNP and the branch site–3' splice site is formed as a result of this recognition. Mutation of the CAG at the 3' splice site reduces formation of this U2 snRNP complex and splicing (Aebi et al. 1986; Lamond et al. 1987). Because changes in the AG dinucleotide do not affect p62 binding, another factor must be involved in mediating the effect of AG dinucleotide on U2 snRNP complex formation. Candidates for this factor would be the IBPs, mentioned previously, or hnRNP proteins A1 or D, which have been shown to have specificity for the AG sequence (Gerke and Steitz 1986; Tazi et al. 1986; Swanson and Dreyfuss 1988). It is possible that the p62 protein could interact with one or more of these proteins in binding to the substrate RNA. Such an interaction, however, is not essential for p62 binding, because the purified p62 protein specifically binds RNA. Experiments indicating a direct role for either IBP or hnRNP A1 or D in splicing are not available. The hnRNP C protein preferentially recognizes the polypyrimidine tract of the 3' splice site region. Whether or not hnRNP C protein and p62 could simultaneously bind a single RNA is unclear. Depletion of nuclear extracts of hnRNP C protein with a monoclonal antibody inactivates the splicing activity, suggesting this protein is important (Choi et al. 1986).

Formation of the U2 snRNA complex also involves recognition of sequences at the branch site. Part of this recognition involves sequence complementarity between U2 snRNA and the branch site (Parker et al. 1987). It is also possible that a protein factor could also bind to sequences at the branch site and promote splicing. Again it is unclear how p62 might cooperate

with factors recognizing the branch site. A factor U2AF was partially purified on the basis of an assay for components important in U2 snRNP binding to the branch site. The most purified fraction contained components that recognized sequences in the polypyrimidine tract and the AG of the 3' splice site. If the U2AF factor is homogeneous, then it can be differentiated from p62 on the basis of its chromatographic behavior on DEAE matrix and its dependence on the AG sequence. It is also possible that U2AF is a mixture of more than one component (e.g., see Krämer 1988), one of which might be p62.

Thus, we conclude that p62 is a novel polypyrimidine tract binding protein that is probably important for RNA splicing, and we propose the name polypyrimidine tract binding protein (pPTB) based on its biochemical activity.

The UV cross-linking method provides a simple and general way of identifying proteins that are in contact with the pre-mRNA during splicing. In addition to p62, we have identified several other proteins that interact with splicing precursors. A protein >200,000 daltons binds the pre-mRNAs only during splicing reactions and is probably the human homologue of the yeast prp8 splicing factor (M.A. García-Blanco, G. Anderson, J. Beggs, and P.A. Sharp, in prep.). A 50,000-dalton protein was seen to cross-link to the pre-mRNA in the splicing complexes, is Sm-precipitable, and binds at or near the 5' splice site (M.A. García-Blanco, S.F. Jamison, and P.A. Sharp, unpubl.). X. Fu and, independently, R.A. Padgett, have also identified what is probably this 50-kD protein (both pers. comm.). As demonstrated in this paper, the cross-linking method can be used as an assay for purification of these proteins. Other proteins that bind RNA in a sequence-specific fashion have been identified (Bandziulis et al. 1989; Hodgkin 1989), and more will undoubtedly be identified in the near future. These proteins will be responsible for specifying the sequences to be excised as introns during splicing and for regulation of this process. Characterization of this family of pro-

teins will be of primary importance in understanding these processes.

Methods

Nuclear extracts and RNA substrates

HeLa cell nuclear extracts were obtained using the procedure of Dignam et al. (1983). RNAs were synthesized from DNA templates using T7 or T3 RNA polymerases (Stratagene) and purified by gel electrophoresis, as described previously (Grabowski et al. 1984; Melton et al. 1984). RNAs were labeled to high specific activity with [³²P]UTP (Melton et al. 1984). Unlabeled RNAs were synthesized in the absence of labeled NTPs or with labeled NTPs at 2500-fold lower specific activity. The plasmids used as templates for RNA synthesis were derivatives of pBS⁻ (Stratagene). pBSAd10 was constructed by ligation of a 90-bp *Fnu*D2–*Hha*I fragment to a 168-bp *Hha*I–*Bam*HI fragment of pBSAd1 (Konarska and Sharp 1987) and cloning of this complex into a *Sma*I–*Bam*HI-restricted pBS⁻ vector. pBSAd13 was constructed by ligation of the 168-bp *Hha*I–*Bam*HI fragment of pBSAd1 into a *Sma*I–*Bam*HI-restricted pBS⁻ vector. T7 RNA polymerase-directed transcription of *Sau*3A-restricted pBSAd10 yielded a 217-nucleotide Ad10 pre-mRNA that contained an intron with a 136-nucleotide internal deletion when compared to the previously described pBSAd1 pre-mRNA (Konarska and Sharp 1987). β-Globin (second intron) pre-RNAs were synthesized from *Eco*RI-linearized pBSAL4 and pBSAL8 (Lamond et al. 1987). PIP3 and mutant PIP3 RNAs were synthesized using T7 RNA polymerase from *Hind*III-linearized pPIP3 and mutant constructs. BPY-1 through BPY-3 RNAs (35-nucleotide long) were synthesized using T3 RNA polymerase from *Kpn*I-linearized pBPY-1 through pBPY-3. pBPY-1 through pBPY-3, plasmids were constructed by cloning oligonucleotides 5'-GTCCCTTTTTTTTCCACAG-3', 5'-GTCCCTCCCTCCCTC-ACAG-3', or 5'-GTCCCTTTTTTTTCCACCC-3', respectively, into a *Hind*III–*Sma*I-digested and Klenow filled-in pBS⁻ vector. A polyadenylation precursor was synthesized from pT3L3 (Skolnik-David et al. 1987), and a fibronectin pre-mRNA was synthesized from pBi8.3 (a gift of P. Norton and R. Hynes). This construct contains the 5' splice site and adjacent sequences of the III8a exon of the rat fibronectin gene (Schwarzbauer et al. 1987; P. Norton and R. Hynes, pers. comm.). A 229-nucleotide-long RNA (pBS RNA) was synthesized from *Hha*I-restricted pBS⁻. RNase T1 oligonucleotides were obtained by RNase T1 digestion of the pBSAd10 transcript. RNA (1 × 10⁸ cpm) was incubated with 10 μg of tRNA (Bethesda Research Laboratories) and 1 unit of RNase T1 (Pharmacia) in 10 mM Tris-HCl (pH 7.5) and 1 mM EDTA for 2 hr at 37°C. The reactions were terminated by lyophilization. The dried RNAs were resuspended in 90% formamide and loaded on a 20% polyacrylamide–7 M urea gel, from which the RNA oligonucleotides were eluted as described (Grabowski et al. 1984).

Monoclonal antibodies that recognize hnRNP C (Choi and Dreyfuss 1984), M, and L proteins were a gift of M. Swanson and G. Dreyfuss, and the anti-Sm (Y12) monoclonal antibody (Lerner et al. 1981) was a gift of J. Steitz.

In vitro assays

Labeled RNA (1 × 10⁵ cpm) was usually added to an in vitro binding reaction. The conditions employed were optimal for splicing in vitro: 20 mM HEPES–KOH (pH 7.9), 62 mM KCl, 2 mM MgCl₂, 1 mM ATP, 5 mM creatine phosphate, 150 μM di-

thiothreitol, and 6% (vol/vol) glycerol. Nuclear extract was spun at 14,000g for 2 min at room temperature, and the supernatant fraction was added to 33% (vol/vol) of the final reaction volume (24 μl). Incubation was at 30°C for 20 min, unless stated otherwise. At this point, the sample was processed by UV cross-linking (see below) or direct loading on a native polyacrylamide gel (Konarska and Sharp 1987), or the RNA was extracted from protein and analyzed on denaturing polyacrylamide gels (Grabowski et al. 1984). In many cases, all three steps were done for the same reaction.

UV cross-linking

The reaction mixtures were transferred to ice and irradiated with a UV light (maximum intensity at 254 nm of 10 mW/min on the surface of the lamp, model UVG-54 from UVP, Inc.) at 4.5 cm from the surface of the lamp for 10 min (Shoemaker and Schimmel 1974; Wilusz and Shenk 1988). The intensity of the lamp was measured using a UV light meter (UVP, Inc.). RNase A was added to 1 mg/ml, and the samples were incubated at 30°C for 15 min. SDS–PAGE sample buffer was added, and the samples were boiled for 3 min and loaded on discontinuous 15% SDS–polyacrylamide gels (Laemmli 1970). The gels were run at a constant 35 mA. The bromophenol blue (BPB) dye front was electrophoresed off the end of the gel to ensure that short RNA oligonucleotides would not be present in the gel. The gel was fixed by shaking in 50% (vol/vol) ethanol and 10% (vol/vol) glacial acetic acid for >1 hr, soaked in 10% (vol/vol) glycerol for 5 min, and dried at 80°C under vacuum. The proteins now labeled by the covalent addition of RNA oligonucleotides were detected by autoradiography on XAR-5 film using one intensifying screen.

Two-dimensional electrophoresis

A 240-μl splicing reaction was loaded on a native 4% acrylamide/bis-acrylamide gel (80 : 1) buffered with 50 mM Tris-glycine (Konarska and Sharp 1987), which was the first dimension. The gel was run at a constant 200 V (13 V/cm) for 5.5 hr; the gel was then placed on developed XAR film for support and UV-irradiated at 4°C for 20 min with the 254-nm lamp described above. The distance from the surface of the gel to the surface of the lamp was 0.5 cm, giving ~8 mW/min of irradiation on the surface of the gel. The wet gel was exposed on film overnight at 4°C to ensure that the expected complexes had been separated on the native gel. The lane containing the complexes was excised and introduced in a sealable plastic bag where it was incubated with 1 mg/ml of RNase A by immersion in a 37°C water bath for 2 hr. The RNase A solution was replaced by 2 × SDS–PAGE buffer, and the gel lane was incubated further at 37°C for 1 hr and at 65°C for 15 min. The gel lane was embedded into the stacking gel portion of a discontinuous SDS–polyacrylamide gel, and this second-dimension gel was run at a constant 25 mA until the BPB reached the bottom of the gel. At this point, the second-dimension gel was processed as described above.

Column chromatography

Nuclear extract [8.5 ml in buffer D, 20 mM HEPES–KOH (pH 7.9), 200 μM EDTA, 0.5 mM dithiothreitol, 0.4 mM PMSF, and 20% (vol/vol) glycerol with 100 mM KCl] was loaded on a DEAE–Sephacel (Pharmacia) matrix column (cross-sectional area = 4.9 cm², height = 4 cm, V = 19.6 cm³) at a flow rate of 5 ml/h-cm² (see Table 1). The matrix was washed with 60 ml of buffer D with 100 mM KCl at the same flow rate. The

flowthrough and the wash were collected and assayed for RNA-binding proteins using the UV-cross-linking assay, for total protein using the Bio-Rad assay, and for conductivity. The DEAE-Sephacel wash (59 ml) was loaded on a heparin-agarose (Sigma) matrix column (CSa = 0.78 cm², h = 6.4 cm, V = 5 cm³), at a flow rate of 11 ml/h-cm². The matrix was washed with 120 ml of buffer D and 100 mM KCl and developed with a 100-ml gradient of KCl concentration from 100 mM to 600 mM KCl in buffer D at a flow rate of 21 ml/h-cm². The flowthrough, wash, and gradient fractions (2.5 ml) were collected and assayed as above. The gradient fractions containing the p62 activity were pooled to ~17 ml; of this, 2.8 ml was diluted with 9 ml of buffer D without KCl and loaded on a AG(U)-agarose [poly(U)-agarose] (Pharmacia) matrix column (CSa = 0.78 cm², h = 3.2 cm, V = 2.5 ml) at a flow rate of 11 ml/h-cm². The flowthrough was loaded again and collected after the second pass through the column. The matrix was washed once with 12 ml of buffer D with 100 µg/ml heparin sulfate. The matrix was developed with a KCl concentration gradient from 0.1 to 1 M KCl in buffer D, and 1-ml fractions were collected.

Acknowledgments

We thank M.M. Konarska for helpful advice and for providing us with valuable materials. We also thank R. Clerc, M. Brown, R. Marciniak, A. Virtanen, D. Chang, and C. Moore for valuable discussions and M. Swanson, G. Dreyfuss, J. Steitz, P. Norton, and R. Hynes for providing us with important materials. We thank M. Siafaca for her help in the preparation of this manuscript. M.A.G.-B. is supported by a U.S. Public Health Service postdoctoral fellowship (F32-CA08342) from the National Institutes of Health. This work was supported by U.S. Public Health Service grants from the National Institutes of Health (R01-GM32467 and R01-GM34277) and partially from the National Cancer Institute Cancer Center core grant (P30-CA14051) to P.A.S.

References

- Aebi, M., H. Herrig, R.A. Padgett, J. Reiser, and C. Weissmann. 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell* **47**: 555–565.
- Bandziulis, R.J., M.S. Swanson, and G. Dreyfuss. 1989. RNA binding proteins as developmental regulators. *Genes Dev.* **3**: 431–437.
- Black, D.L., B. Chabot, and J.A. Steitz. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in pre-mRNA splicing. *Cell* **42**: 737–750.
- Blum, H., H. Beier, and H.J. Gross. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8**: 93–99.
- Brody, J. and J. Abelson. 1985. The 'spliceosome': Yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**: 1344–1349.
- Choi, Y.D. and G. Dreyfuss. 1984. Monoclonal antibody characterization of the C protein of heterogeneous nuclear ribonucleoprotein complexes in vertebrate cells. *J. Cell. Biol.* **99**: 1997–2004.
- Choi, Y.D., P.J. Grabowski, P.A. Sharp, and G. Dreyfuss. 1986. Heterogeneous nuclear ribonucleoproteins: Role in RNA splicing. *Science* **231**: 1534–1539.
- Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**: 1475–1489.
- Dreyfuss, G. 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. *Annu. Rev. Cell Biol.* **2**: 459–498.
- Gerke, V. and J.A. Steitz. 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3' splice site of pre-messenger RNA. *Cell* **47**: 973–984.
- Grabowski, P.J., R.A. Padgett, and P.A. Sharp. 1984. Messenger RNA splicing in vitro: An excised intervening sequence and a potential intermediate. *Cell* **37**: 415–427.
- Grabowski, P.J., S.R. Seiler, and P.A. Sharp. 1985. A multicomponent complex is involved in the splicing of precursors to mRNAs. *Cell* **42**: 345–353.
- Hartwell, L.H., C.S. McLaughlin, and J.R. Warner. 1970. Identification of ten genes that control ribosome formation in yeast. *Mol. Gen. Genet.* **109**: 42–56.
- Hodgkin, J. 1989. Drosophila sex determination: A cascade of regulated splicing. *Cell* **56**: 905–906.
- Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNA. *Cell* **46**: 845–855.
- . 1987. Interaction between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* **49**: 763–774.
- 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.
- Krämer, A. 1988. Pre-splicing complex formation requires two proteins and U2 snRNP. *Genes Dev.* **2**: 1155–1167.
- Krämer, 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.
- Kumar, A., H. Sierikowska, and W. Szer. 1987. Purification and RNA binding properties of a C-type hnRNP protein from HeLa cells. *J. Biol. Chem.* **262**: 17126–17137.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.
- Lamond, A.I., M.M. Konarska, and P.A. Sharp. 1987. A mutational analysis of spliceosome assembly: Evidence for splice site collaboration during spliceosome formation. *Genes Dev.* **1**: 532–543.
- Lamond, A.I., M.M. Konarska, P.J. Grabowski, and P.A. Sharp. 1988. Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoproteins. *Proc. Natl. Acad. Sci.* **85**: 411–415.
- Last, R.L. and J.L. Woolford, Jr. 1986. Identification and nuclear localization of yeast pre-messenger RNA processing components: RNA 2 and RNA 3 proteins. *J. Cell. Biol.* **103**: 2103–2112.
- LeGrain, P., B. Seraphin, and M. Rosbash. 1988. Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol. Cell. Biol.* **8**: 3755–3760.
- Lerner, E.A., M.R. Lerner, C.A. Janeway, and J.A. Steitz. 1981. Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. *Proc. Natl. Acad. Sci.* **78**: 2737–2741.
- Lossky, M., G.J. Anderson, S.D. Jackson, and J. Beggs. 1987. Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions. *Cell* **51**: 1019–1026.
- Lustig, A.J., R.-J. Lin, and J. Abelson. 1986. The yeast RNA gene products are essential for mRNA splicing in vitro. *Cell* **47**: 953–963.
- Malim, M.H., J. Hauber, R. Fenrick, and B.R. Cullen. 1988. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335**: 181–183.
- Maniatis, T. and R. Reed. 1987. The role of small nuclear ribon-

- ucleoprotein 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.
- Melton, D.A., P.A. Krief, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**: 7035–7056.
- Moore, C.L., J. Chen, and J. Whorisky. 1988. Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. *EMBO J.* **7**: 3159–3169.
- 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.
- Padgett, R.A., P.J. Grabowski, M.M. Konarska, S.R. 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.
- Pikielny, C.W., B.C. Rymond, and M. Rosbash. 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature* **324**: 341–345.
- Rosbash, M., P.K.W. Harris, J.L. Woolford, Jr., and J.L. Teem. 1981. The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. *Cell* **24**: 679–686.
- Ruby, S.W. and J. Abelson. 1988. An early hierarchic role of U1 small nuclear ribonucleoproteins in spliceosome assembly. *Science* **242**: 1028–1035.
- Ruskin, B. and M.R. Green. 1985. Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. *Cell* **43**: 131–142.
- 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.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**: 5463–5467.
- Shoemaker, H.J.P. and P.R. Schimmel. 1974. Photoinduced joining of a transfer RNA with its conjugate aminoacyl-transfer RNA synthetase. *J. Mol. Biol.* **84**: 503–513.
- Schwarzbauer, J.E., R.S. Patel, D. Fonda, and R.O. Hynes. 1987. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J.* **6**: 2573–2580.
- Seraphin, B., L. Kretzner, and M. Rosbash. 1988. A U1 snRNA: Pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J.* **7**: 2533–2538.
- Skolnik-David, H., C.L. Moore, and P.A. Sharp. 1987. Electrophoretic separation of polyadenylation-specific complexes. *Genes Dev.* **1**: 672–682.
- Swanson, M.S. and G. Dreyfuss. 1988. RNA binding specificity of hnRNP proteins: a subset binds to the 3' end of introns. *EMBO J.* **7**: 3519–3529.
- Tazi, J., C. Alibert, J. Temsamani, I. Reveilland, 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.
- Wieringa, B., E. Hofer, and C. Weissmann. 1984. A minimal intron length but no specific internal sequence is required for splicing the large rabbit β -globin intron. *Cell* **37**: 915–925.
- Wilusz, J.T. and Shenk. 1988. A 64 kD nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. *Cell* **52**: 221–228.



Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns.

M A García-Blanco, S F Jamison and P A Sharp

Genes Dev. 1989, **3**:

Access the most recent version at doi:[10.1101/gad.3.12a.1874](https://doi.org/10.1101/gad.3.12a.1874)

References

This article cites 50 articles, 13 of which can be accessed free at:
<http://genesdev.cshlp.org/content/3/12a/1874.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](#).

