*RNA* (2001), *7*:1239–1253. Cambridge University Press. Printed in the USA.

DOI: 10.1017.S1355838201000607

# An in vitro-selected RNA-binding site for the KH domain protein PSI acts as a splicing inhibitor element

**ASOKA K. AMARASINGHE,[1] ROBIN MACDIARMID,[1,2] MELISSA D. ADAMS,[1,3] and DONALD C. RIO[1]**

Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720, USA

## ABSTRACT

**P element somatic inhibitor (PSI) is a 97-kDa RNA-binding protein with four KH motifs that is involved in the inhibition of splicing of the *Drosophila* P element third intron (IVS3) in somatic cells. PSI interacts with a negative regulatory element in the IVS3 5′ exon. This element contains two pseudo-5′ splice sites, termed F1 and F2. To identify high affinity binding sites for the PSI protein, in vitro selection (SELEX) was performed using a random RNA oligonucleotide pool. Alignment of high affinity PSI-binding RNAs revealed a degenerate consensus sequence consisting of a short core motif of CUU flanked by alternative purines and pyrimidines. Interestingly, this sequence resembles the F2 pseudo-5′ splice site in the P element negative regulatory element. Additionally, a negative in vitro selection of PCR-mutagenized P element 5′ exon regulatory element RNAs identified two U residues in the F1 and F2 pseudo-5′ splice sites as important nucleotides for PSI binding and the U residue in the F2 region is a nearly invariant nucleotide in the consensus SELEX motif. The high affinity PSI SELEX sequence acted as a splicing inhibitor when placed in the context of a P element splicing pre-mRNA in vitro. Data from in vitro splicing assays, UV crosslinking and RNA-binding competition experiments indicates a strong correlation between the binding affinities of PSI for the SELEX sequences and their ability to modulate splicing of P element IVS3 in vitro.**

**Keywords: KH domains; pre-mRNA splicing; RNA-binding proteins; SELEX**

## INTRODUCTION

The alternative splicing of pre-mRNA is one of the major mechanisms involved in regulating gene expression in eukaryotes. It allows a single pre-mRNA to be processed to encode many different protein isoforms and can restrict expression to a given cell or tissue type and/or a specific developmental stage. The regulation of alternative splicing involves both *cis*-acting pre-mRNA sequence elements and *trans*-acting protein factors. These exonic or intronic regulatory RNA sequences can act as splicing enhancers or inhibitors and their interaction with serine-arginine (SR) proteins, heterogeneous nuclear ribonucleoproteins (hnRNPs), and other *trans*-acting RNA-binding proteins modulate alternative splice site selection (for reviews, see Adams et al., 1996; Lopez, 1998; Graveley, 2000; Smith & Valcarcel, 2000).

In *Drosophila*, tissue-specific P element transposition is regulated by alternative splicing of the P element transposase pre-mRNA. *Drosophila* P element pre-mRNA contains four exons separated by three introns and the third intron (IVS3) is spliced only in germ line cells. The fully spliced mRNA encodes the transposase protein that is required for transposition of P elements in germ line cells (Laski et al., 1986). In somatic cells, IVS3 is not spliced and resulting mRNA encodes a truncated protein that represses transposition (Misra & Rio, 1990). A negative regulatory element has been identified that lies upstream of the accurate 5′ splice site of IVS3 and contains two pseudo-5′ splice sites, termed F1 and F2 (Laski & Rubin, 1989; Siebel & Rio, 1990; Chain et al., 1991). This regulatory element interacts with U1 snRNP, the hnRNPA1-like protein, hrp48, and the KH domain protein, P element Somatic Inhibitor (PSI). These interactions result in inhibition of U1

Reprint request to: Donald C. Rio, Department of Molecular and Cell Biology, University of California, Berkeley, 401 Barker Hall, Berkeley, California 94720, USA; e-mail: don_rio@uclink4.berkeley.edu.

[2]Present address: Department of Plant and Microbial Biology, University of California, Berkeley, 111 Koshland Hall, Berkeley, California 94720, USA.

[3]Present address: Department of Biology, University of North Carolina–Chapel Hill, 303 Fordham Hall, Chapel Hill, North Carolina 27599, USA.

snRNP binding to the accurate IVS3 5′ splice site, an early step in splice site selection and spliceosome assembly (Siebel et al., 1992). The hnRNP protein hrp48 specifically interacts with the F2 pseudo-5′ splice site (Siebel et al., 1994). It has been shown that the F1 pseudo-5′ splice site binds U1 snRNP, but this site is not used for splicing (Siebel et al., 1992).

The PSI protein was identified as a component that interacts with the IVS3 5′ exon and is required for the inhibition of IVS3 splicing in vitro. PSI is highly expressed in the soma and is thought to be the determinant of the tissue-specific expression of the transposase protein (Siebel et al., 1995). Ectopic expression of PSI in the female germ line is sufficient to repress splicing of an IVS3 reporter transgene (Adams et al., 1997). PSI contains a N-terminal domain consisting of four hnRNP K homology (KH) RNA-binding motifs and a C-terminal region with two glutamine-rich repeats (A and B; Siebel et al., 1995). An N-terminal fragment of PSI containing only the four KH motifs (KH0-3) can specifically bind to the IVS3 5′ exon RNA (R. MacDiarmid and D.C. Rio, unpubl.; see below). PSI is known to interact directly with U1 snRNP through the A and B motifs (Labourier et al., 2001).

KH domains have been identified in over 100 RNA-binding proteins with diverse cellular functions from a wide variety of organisms (Burd & Dreyfuss, 1994). In *Drosophila*, it is predicted that the genome sequence encodes 27 KH domain proteins (Lasko, 2000). The KH motif is about 70 amino acids in length and can be present in single or multiple copies within a given protein. The KH domain was first identified as a region in hnRNP K (the K-homology region) that binds to single-stranded RNA nonspecifically (Siomi et al., 1993). The hnRNP K protein has three KH motifs and binds poly r(C) (Matunis et al., 1992). The RNA-binding properties of several KH domain proteins have been studied, including Nova-1, Nova-2, and vigilin (Buckanovich & Darnell, 1997; Kanamori et al., 1998). These proteins have been shown to bind to specific target RNA sequences. The Nova-1 protein recognizes RNA stem loops containing the $(UCAUY)_3$ (Jensen et al., 2000), and the Nova-2 protein binds to a high affinity consensus site, GAGUCAU, in RNA stem loops (Yang et al., 1998). These studies show that these KH domain proteins can recognize short pyrimidine-rich RNA sequences. A recent crystal structure of the Nova-2 protein KH domain bound to RNA shows that a single KH motif interacts with as few as 4–5 nt of single-stranded RNA (Lewis et al., 2000). These studies also showed that other regions of the RNA-bound KH domain are accessible to carry out protein–protein interactions. Studies on vigilin, a multi-KH domain protein containing 14 KH motifs, showed that it recognizes sequences with conserved $(A)_nCU$ and $UC(A)_n$ motifs (Kanamori et al., 1998). The Sam68 protein, a specific target of the src tyrosine kinase in mitosis, is another KH domain-containing protein that recognizes specific target RNAs. Sam68 was shown to recognize RNA sequences with a UAAA motif (Lin et al., 1997). Thus, KH domains appear to recognize short motifs in RNA.

Although the PSI protein is known to bind the P element IVS3 5′ exon RNA, its exact RNA-binding site has not been identified. Using an in vitro genetic selection approach (SELEX), we have identified high affinity RNA-binding sequences for PSI. The deduced consensus sequence suggests that the PSI protein may interact with U residues in the upstream regulatory element containing the F1 and F2 pseudo-5′ splice sites and this idea is supported by data obtained using mutations in the P element 5′ exon RNA. In addition, we show that a high affinity SELEX sequence can act as a splicing inhibitor element to block P element pre-mRNA IVS3 splicing in vitro. Finally, we show a correlation between the binding of specific *Drosophila* somatic Kc nuclear extract proteins, including PSI, to these in vitro selected RNAs and their activities in in vitro splicing assays.

## RESULTS

### Selection of high affinity RNA-binding sequences for PSI protein

Although PSI protein binds to the IVS3 5′ exon (Siebel et al., 1992), it has not been possible to identify the RNA-binding sites using conventional footprinting methods (M.D. Adams, A.K. Amarasinghe, unpubl.). To identify RNA sequences with high affinity binding sites for PSI protein, in vitro genetic selection (SELEX) was performed (Tuerk & Gold, 1990). Full-length PSI protein was purified to homogeneity from *Drosophila melanogaster*, human HeLa cells infected with a vaccinia virus vector, and *Escherichia coli* (Fig. 1A) and each protein was confirmed to be active for binding to P element IVS3 5′ exon RNA by native gel electrophoresis (data not shown). Although the PSI protein expressed in eukaryotic cells appears to be phosphorylated (Siebel et al., 1995; R. MacDiarmid, unpubl.), both eukaryotic proteins and the nonphosphorylated *E. coli* protein exhibit similar RNA-binding activities (Siebel et al., 1995). A pool of radiolabeled RNA ligands containing 30 random nucleotides between unique flanking regions was generated by T7 transcription from a degenerate DNA oligonucleotide pool. In vitro genetic selection was carried out through 10 cycles of reverse transcription, PCR amplification, and nitrocellulose filter binding using PSI protein purified either from *D. melanogaster* or from human HeLa cells (Fig. 1B). Increasing amounts of bound RNA was observed with increasing rounds of selection (Fig. 1C). Purified *Drosophila* PSI protein gave a slightly higher percentage of bound RNA than the purified PSI protein from HeLa cells. A total of 33 individual RNA sequences selected from both recombinant

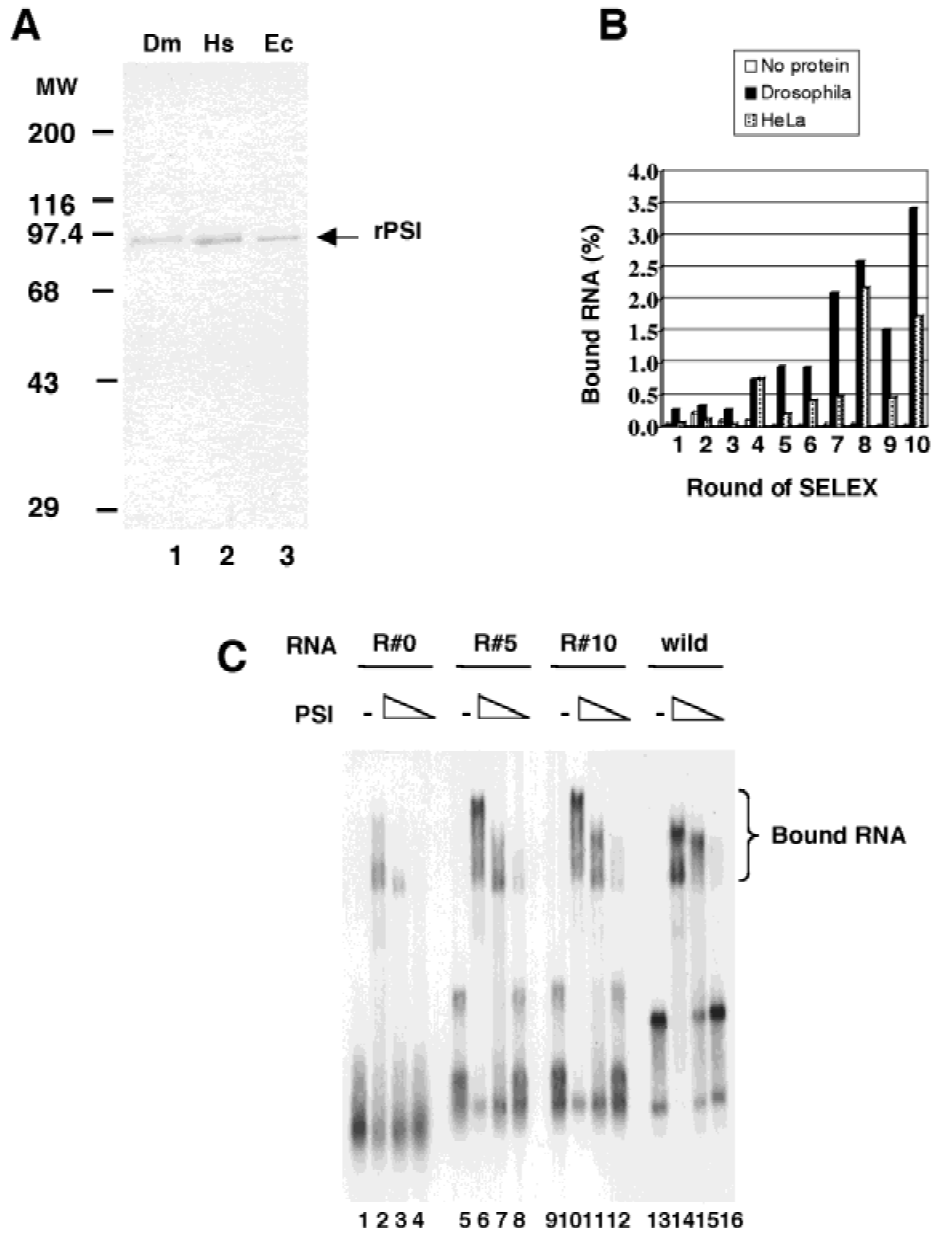



**FIGURE 1.** In vitro selection of PSI RNA-binding sequences. **A:** Purified PSI proteins. Recombinant PSI (12 pmol each) isolated to near homogeneity from three different cell types, separated by 5% SDS PAGE and visualized by coomassie staining. Lane 1: *Drosophila* cells; lane 2: HeLa cells; lane 3: *E. coli.* Note the difference in migration due to the posttranslational modifications of PSI in *Drosophila* and HeLa cells. Molecular weight markers are given to the left. **B:** Selection of RNA ligands results in an increase in affinity to PSI. The increase in affinity of RNA ligands to PSI protein purified from *Drosophil*a (solid) or HeLa cells (hatched) was determined after each round of selection. The SELEX RNA bound to PSI retained on the nitrocellulose filter in each round as a percentage of the total pool RNA was plotted against number of rounds of selection. **C:** Native gel analysis of SELEX RNA pools. Native gel analysis was used to assay a fraction of the above SELEX RNAs affinity to PSI protein. Fivefold serial dilutions of PSI isolated from *Drosophila* cells (highest concentration is 12 pmol) were used to bind RNA after 0 (lanes 1–4), 5 (lanes 5–8) or 10 (lanes 9–12) rounds of selection compared to wild-type 5′ exon RNA (lanes 13–16). Note that the amount of higher molecular weight RNA–protein complex formation has increased in round 10 (lanes 10 and 11) compared to the round 5 (lanes 6 and 7).

proteins were cloned and sequenced for comparative analysis and for comparison to the one known natural RNA target of PSI, the P element IVS3 5′ exon RNA. Computer-generated alignments of these sequences (using the MAST algorithm, see Fig. 2A, legend) identified a relatively degenerate consensus sequence consisting of a core motif of CUU flanked by alternating purines and pyrimidines (RCYYCUURYRC; Fig. 2A). This short core motif was present in more than 60% of the cloned SELEX sequences and the individual sequences generally showed several mismatches to the above extended consensus sequence. Most sequences

**A**

| | F1 | F2 | | KH0–3 Binding |
|---|---|---|---|---|
| Wild 5′exon | UUUAAGU | AUAGG**UU**AAGA | AAAUAUAUAAU | ++++ |
| Up#2 | GC | UCUAU**UU**GUCC | CUCUCUUCUUCGUUCUU | + |
| Up#3 | CAGCAUUGGAACUGCU | GCUA**CUU**UAGC | CUU | +++ |
| Up#4 | CUGUC | CUAC**CUU**ACGC | GCAGUGUCUGCUCU | +++ |
| Up#5 | A | ACUC**CUU**GCUC | UCCUAAAUUUGCAGCUCU | ++++ |
| Up#6 | C | GAGA**CUU**UUGC | UUUUCUUGCGUCUUUU | +++++ |
| Up#7 | AUCCUCUGAA | CUGU**CCU**ACUG | UGUAUUCCG | +++++ |
| Up#8 | | AAGAGC**U**ACGC | CUGUCCACUCUUGCGCUUC | ++++ |
| Up#9 | UCUUG | CCGA**CUU**ACAC | UGUGUUUACCUUU | ++++ |
| Up#10 | UCU | GCAGG**U**AAGUC | UAGGUUUUGGUUCUGC | +++ |
| Up#11 | | GCAG**CUU**GCGC | AUGUGGUCAGGUUUGCGG | ++++ |
| Up#12 | CU | AUAC**CUU**ACGA | CACACUAUGUUUUGGCG | +++ |
| Up#13 | G | AAACGG**U**ACGC | CUACCAAGGUGUCCUCCU | +++ |
| S#1 | UACCAAUCGCU | ACUCU**UU**ACGC | UCUCUGUG | |
| S#2 | CCGA | ACCG**CUU**UCAU | GACUCACCUUACUAC | |
| S#7 | UCCUGUUAUUUAU | CGGC**CUU**AUAA | CUCUUU | |
| S#20 | UUUUUGCAAC | UAGU**CUU**AGUG | UACCCCCAC | |
| S#25 | C | UGCUG**UU**GACU | CAUGCCACCUAUGACUA | |
| S#31 | UGCUUCC | UCUC**CUU**UGAC | UAGACGUUAGU | |
| S#34 | GUUGCCCAUAUGUACGC | GUAA**C**C**U**ACUC | UU | |
| S#40 | UCUGUAUAUUAAA | CUGU**C**C**U**CUCA | CCCUGC | |
| S#43 | GUGU UCUGGCCGCCUC | CCUU**CUU**AGCU | UGU | |
| S#46 | | UCAU**CUU**GUUA | UGUGCAUCUAAUCGUCACU | |
| S#6 | ACCGACUAUC | UCUAU**UU**UAUC | UCUCGNCGC | |
| S#12 | GGUCUCCUGC | GUGGA**UU**GUGU | GUGGUGCUC | |
| S#18 | CGCCAAAAC | AUAG**U**GU**U**GUCG | UAAGGUAUAG | |
| S#24 | ACCUCC | UGCC**CUU**ACUA | CUAAUACGCUAUC | |
| S#27 | UCUGC | AUUA**CUU**CUGA | CCUUACCUUUGUUC | |
| S#29 | UCG | AAUG**CUU**UCAC | CUCGUCUCUCCUCUGU | |
| S#32 | UCGCG | UCUG**CUU**CUGA | UUUGUGUCCCUGC | |
| S#35 | CCUCACUUCUCUCUUUU | ACUAGC**U**UACC | CUUG | |
| S#36 | CUUUA | CCGA**CUU**UCAC | UGUGUUACCGCUCU | |
| S#42 | AUACGCCUG | CUAA**C**C**U**CGUC | UACUACUCGU | |
| S#45 | UAA | ACCU**CUU**UUCU | AGCUCCUCUCGCCUUA | |
| | Consensus motif | RCYY**CUU**RYRC | | |

**B**

| | F1 F2 |
|---|---|
| Wild | UUUAAGUAUAGGUUAAGAAAAUAUAUAAU |
| D#48 | ------C------C---------- |
| D#424 | -----C-------C---------- |
| D#520 | ----------C---------- |
| D#519 | ----------C----------A |
| D#44 | ----------U---------- |
| D#513 | ----------C---------- |
| CONSENSUS | RCYYCUURYRC |

**C**

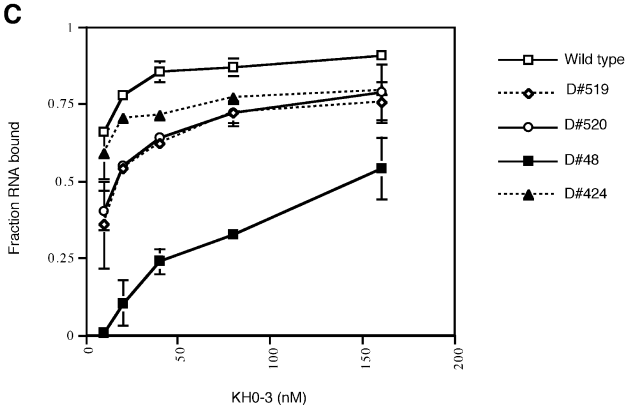



**FIGURE 2.** Alignment of the high affinity RNA sequences obtained from the in vitro selection and their KH0–3 binding activities. **A:** Sequences of individual SELEX clones selected by PSI are shown. The sequences were analyzed by MAST algorithm (WWW.sdsc.edu/MEME) to search for a common motif, and aligned on the basis of the best fit to the consensus in each sequence. The consensus sequence for the PSI high-affinity binding site is given below (Y: pyrimidines, R: purines). The CUU core motif is shown in bold. The F1 and F2 pseudo-5′ splice sites are underlined. The SELEX sequences used for gel shift assays with KH0–3 protein are named Up#2 to Up#13 and their relative binding efficiencies compared to the wild-type 5′ exon is given to the right. The SELEX Sequences Up#2 and Up#6 were used for further characterization. **B:** Sequences of the wild-type IVS3 5′ exon regulatory element and "down SELEX" PCR mutants. Pseudo-5′ splice sites F1 and F2 in the 5′ exon regulatory element are shown. Nucleotide changes in the 5′ exon PCR mutant are given below the wild-type sequence. Only an aligned 29-nt-long RNA from each substrate is shown. The common flanking regions in the "down SELEX" substrates are not shown for clarity. **C:** RNA-binding curves for down SELEX RNAs. Increasing amounts of purified KH0–3 protein was incubated with radiolabeled down SELEX RNAs and complexes were run on a native gel. The intensity of shifted and free RNA bands was quantitated, and bound RNA fraction was plotted against the protein concentration (data obtained from at least three separate experiments).

showed an alternate pattern of purines and pyrimidines on either side of this short core motif. Equivalent numbers of the cloned SELEX sequences analyzed came from either the *D. melanogaster* or the human HeLa cell purified PSI proteins and gave similar results. Additionally, 18% of the sequences analyzed carry a core motif of either CCU or UUU indicating a preference for the pyrimidine-rich core motif. Interestingly, the consensus sequence identified by in vitro selection (RCYYCUURYRC) resembles the F2 pseudo-5′ splice site present in the P element third intron (IVS3) 5′ exon negative regulatory element (UAGGUUAAG; Fig. 2B). Two uridine residues found in the F2 pseudo-5′ splice site are nearly invariant in the above selected se-

quences. This observation suggests that the PSI binding site in the IVS3 5′ exon may be present in the region of the F2 pseudo-5′ splice site.

### Binding of full-length PSI and the PSI four KH domains to SELEX sequences

The PSI protein contains four N-terminal KH RNA-binding domains and two 100 amino acid C-terminal repeat motifs termed A and B that interact with U1 snRNP-70K protein, but that do not appear to bind RNA (Labourier et al., 2001). Previous RNA-binding studies of the full-length recombinant PSI protein showed that in RNA gel shift assays, PSI exhibited multiple RNA–protein complexes in native gels independent of protein concentration, suggesting multiple modes or conformations of RNA binding rather than cooperative or multimeric binding of the protein to the P element 5′ exon (Siebel et al., 1992; M. Adams, unpubl. results; see Fig. 1C, lanes 14–15). Therefore, we tested binding of both the full-length PSI protein and a truncated protein containing only the four KH domains (KH0–3), expressed and purified from *E. coli*, to several SELEX RNAs. The relative binding affinities of the SELEX RNAs were compared to the wild-type P element 5′ exon RNA. The affinities of the different SELEX RNAs varied dramatically (Fig. 2A, upper right side). Interestingly, one SELEX RNA (Up#6) bound both the full-length PSI and the KH0–3 proteins with higher affinity than the P element 5′ exon RNA (KH0-3 $K_D$ values for wild-type and Up#6 RNA are 5.5 nM and 3.8 nM, respectively; Fig. 3A, lanes 11–15, 3B, lanes 13–18). In contrast to Up#6, one RNA (SELEX Up#2) showed very weak binding to PSI. Both the KH0–3 protein (Fig. 3B, lanes 7–12) and the full-length PSI (Fig. 3A, lanes 6–10) showed reduced binding to this SELEX RNA. These results indicate that the N-terminal four KH domains are sufficient for RNA binding by PSI and the selected sequences represent a population of high affinity binding sites for the KH domains of PSI.

### Site-specific labeling and UV crosslinking, as well as a negative in vitro selection, identifies the PSI protein binding region in the P element 5′ exon RNA

RNA footprinting and RNA modification–interference experiments to identify the RNA-binding site for PSI on the IVS3 5′ exon have not been successful. This problem is similar to that observed for binding of tra-2 protein to the dsx enhancer element (T. Maniatis, pers. comm.). This inability to map a specific binding site is possibly due to heterogeneity in RNA–protein contacts or in the modes by which the four KH domains of PSI bind to the 5′ exon RNA. To test which region of the P element 5′ exon regulatory element PSI interacts with, a differential labeling and UV crosslinking experiment was performed. Two 5′ exon RNAs were prepared in which either the 5′ half or 3′ half of the RNA was radiolabeled with [$\alpha$-$^{32}$P]UTP (Fig. 4C). UV photochemical crosslinking assays showed that PSI only crosslinked to the RNA in which the 5′ half was labeled (Fig. 4A, lanes 1–3). No crosslinking was observed to the RNA with the labeled 3′ half (Fig. 4A, lanes 4–6). Competition experiments with either unlabeled 5′ half (Fig. 4B, lanes 2–3) or unlabeled 3′ half RNA (Fig. 4B, lanes 4–6) showed that only the 5′ half of the 5′ exon containing the F1 and F2 sites (Fig. 4C) could compete for PSI binding. This experiment confirms that PSI interacts with the 5′ exon in the F1/F2 region.

An additional approach utilizing negative SELEX was used to define the specific nucleotides important for PSI binding on the IVS3 5′ exon sequence. The optimal conditions for nitrocellulose filter binding of *E. coli*-expressed PSI protein were established by comparing the binding of PSI to the 5′ exon RNA versus a nonspecific RNA, PL2 (Siebel & Rio, 1990; see Materials and Methods). Two rounds of mutagenic PCR were performed on the 5′ exon and RNA not bound by the PSI protein was recovered, amplified and retested to recover the sequences that failed to bind PSI. After four rounds of filter binding and negative selection, these RNAs showed a reduction in binding to PSI as determined by native gel analysis (data not shown). Individual down-selected variants of the P element 5′ exon RNA were further analyzed by cloning, sequencing, and gel shift analysis with the PSI protein (data not shown). Sequence analysis of several clones showed that most of the mutations were restricted to three regions of the 5′ exon RNA, the F1 and F2 pseudo-5′ splice sites and the purine-rich region adjacent to the F1/F2 site (Fig. 2B).

Negatively selected 5′ exon RNAs were tested for binding to the PSI KH0–3 protein. RNAs with mutations in the 3′ purine-rich region showed similar binding affinities for the PSI KH0–3 protein compared to that of the wild-type 5′ exon sequence (data not shown). RNAs carrying mutations in the F1 and F2 sites showed weaker binding to KH0–3 (Fig. 2C). These RNAs carry a common mutation in the F2 pseudo-5′ splice site, a U-to-C change (Fig. 2B). The single mutant RNA D#520 and the double mutant RNA D#519, with an additional mutation 3′ to the F1/F2 site (U-to-A change), show a similar reduction in PSI KH0–3 protein binding. The mutant D#424 RNA carrying the second mutation (G-to-C change) in the F1 pseudo-5′ splice site displayed a slight increase in binding activity compared to D#520 or D#519 (Fig. 2C). Among those RNAs tested, the D#48 RNA had the weakest binding (KH0–3 $K_D$ values for wild-type and D#48 RNA are 5.5 nM and 91.6 nM, respectively) and this sequence has two mutations in the U residues in the F1 and F2 pseudo–5′ splice sites (Fig. 2B and Fig. 3A, lanes 19–24). Two other single

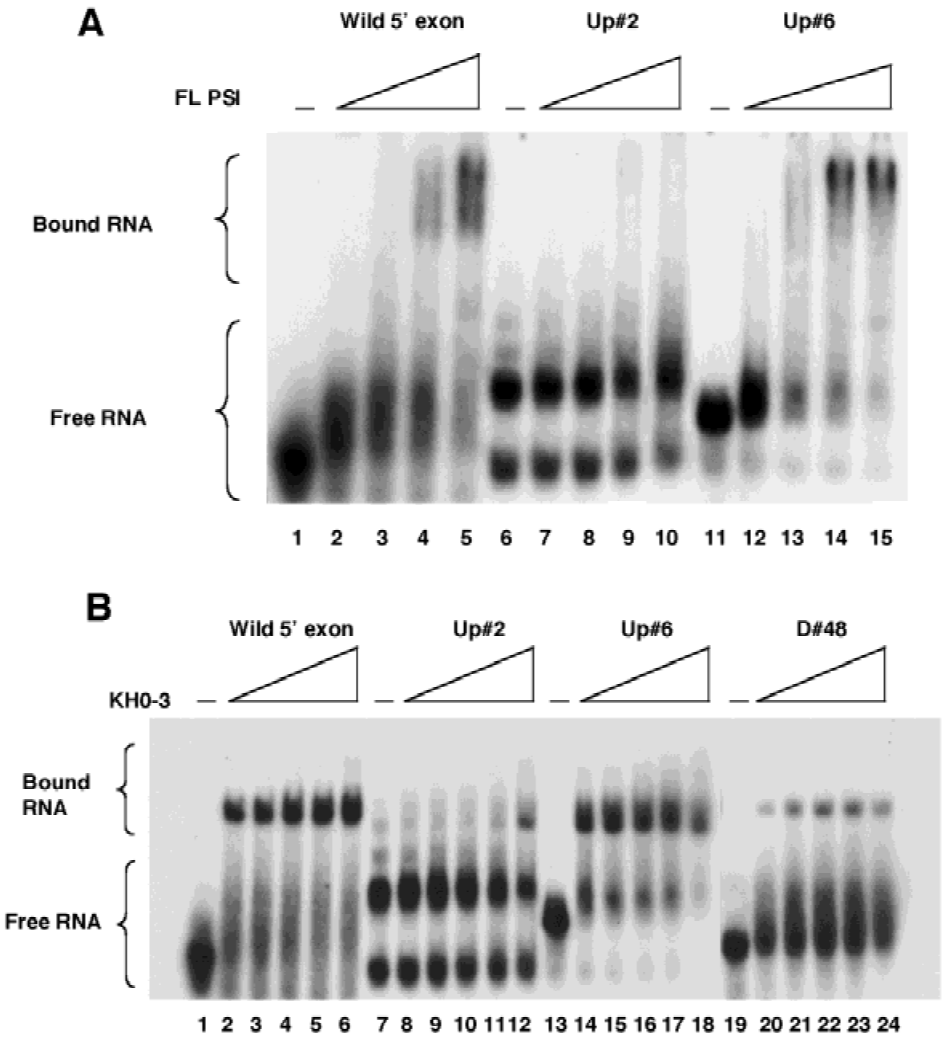



**FIGURE 3.** Analysis of SELEX RNA binding by four KH domains and full-length PSI protein. **A:** SELEX RNA binding by PSI protein. Increasing amounts of full-length PSI protein (0, 25, 50, 150, and 200 nM) purified from *E. coli* was used. Lanes 1–5: wild-type 5′ exon, lanes 6–10: SELEX Up#2, and lanes 11–15: SELEX Up#6. **B:** SELEX RNA binding by four KH domains. RNA-binding activity was tested for wild-type 5′ exon and SELEX Up#2, Up#6, and D#48 RNAs with purified N-terminal KH0–3 protein. Increasing amounts of KH0–3 protein (0, 10, 20, 40, 80, and 160 nM respectively, in lanes 1–6, 7–12, 13–18, and 19–24) were incubated with radiolabeled RNA and reactions were run on a native gel. Arrows indicate the free RNA and RNA–protein complex formed on these RNAs.

mutant RNAs with changes in the F2 region (Fig. 2B, D#44 and D#513) had no effect on PSI KH0–3 protein binding (data not shown). These data suggest that PSI specifically interacts with these two uridine residues in the F1 and F2 pseudo-5′ splice sites and that these are critical nucleotides for PSI binding.

### A high affinity RNA-binding sequence for PSI can function as a splicing inhibitor of P element IVS3 in vitro

To test the functional activities of selected SELEX RNAs in the context of P element pre-mRNA splicing substrates, the wild-type IVS3 5′ exon sequence was substituted with the PSI high affinity SELEX Up#6, the low affinity PSI binding sequence SELEX Up#2, or the negatively selected D#48 RNA sequences (all of these splicing substrates maintained an equal distance between the inserted sequence and the accurate IVS3 5′ splice site). As controls, 5′ exon mutants L#456 and F1,F2-GT (Siebel et al., 1992) were used (see Fig. 5B for sequences). The L#456 and F1,F2-GT 5′ exon mutants show activation of IVS3 splicing in vivo (Laski & Rubin, 1989) and in vitro (Siebel et al., 1992). The substrate RNAs were incubated in Kc cell splicing extracts, which exhibit little or no splicing of the wild-type P element IVS3 (Siebel & Rio, 1990; Siebel et al., 1992). A low level of splicing was observed for the wild-type IVS3 RNA (Fig. 5A, lane 1). The cleaved intron–3′ exon intermediate lariat of 205 nt and 5′ exon product of 56 nt

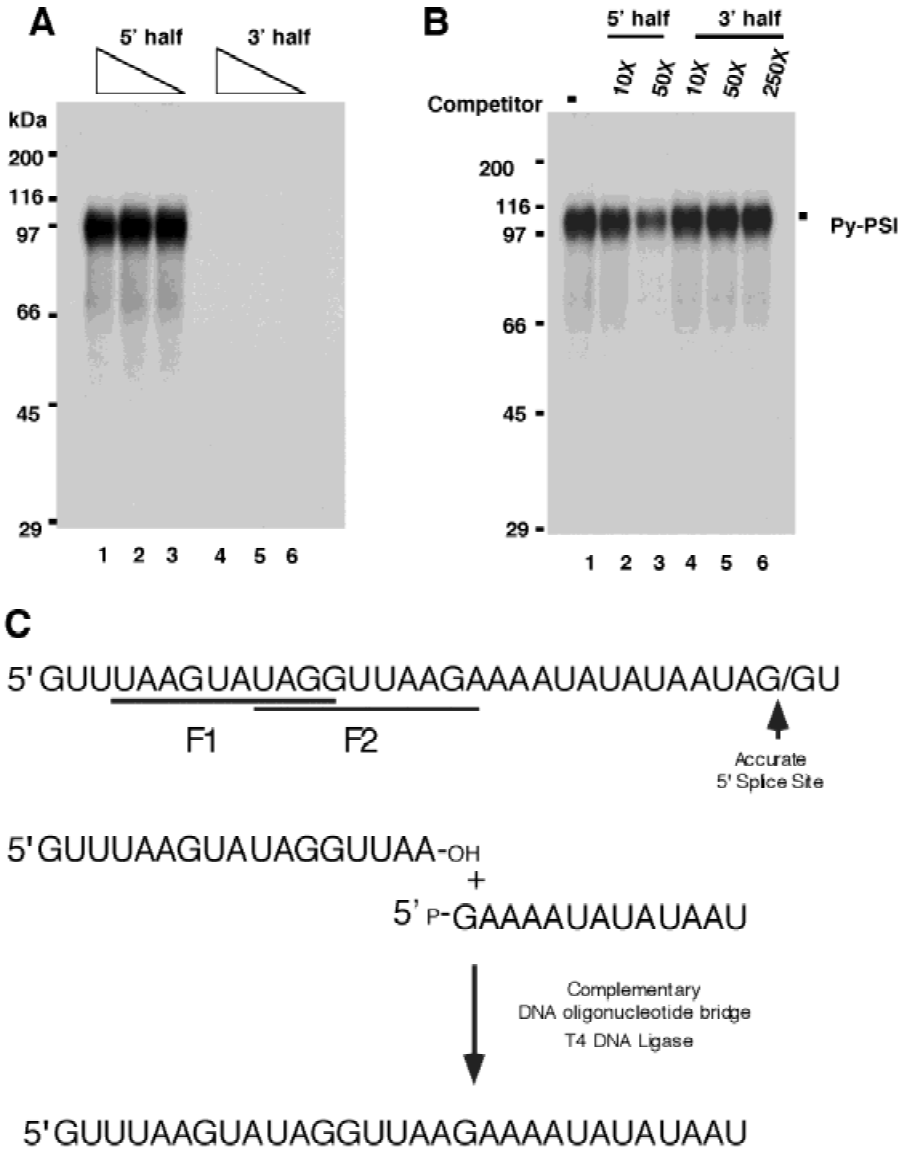



**FIGURE 4.** Binding of PSI to the differentially labeled 5′ exon RNA. **A:** Crosslinking of PSI to the radiolabeled halves of the 5′ exon regulatory element. Purified full-length PSI protein was subjected to crosslinking and increasing amounts of RNase A was added. PSI crosslinked radiolabeled 5′ half RNA (lanes 1–3) and radiolabeled 3′ half RNA (lanes 4–6) are shown. **B:** Competition assays with unlabeled halves of the 5′ exon regulatory element. Increasing amounts of unlabeled either 5′ half (lanes 2–3: 10- and 50-fold molar excess) or 3′ half (lanes 4–6: 10- 50- and 250-fold molar excess) of 5′ exon regulatory element was added to the crosslinking reaction prior to the addition of probe with labeled 5′ half of 5′ exon RNA. Lane 1: no cold competitor added. **C:** Construction of radiolabeled halves of IVS3 5′ exon regulatory element. Radiolabeled one half of the 5′ exon RNA was ligated together with the corresponding other half using a complementary DNA oligonucleotide for use in crosslinking assays.

for L#456 and F1,F2-GT substrates were clearly observed (Fig. 5A, lanes 2 and 6). Remarkably, the substrate containing the high affinity SELEX Up#6 showed a complete inhibition of IVS3 splicing (Fig. 5A, lane 3), whereas Up#2, which binds PSI poorly, showed no reduction in splicing activity comparable to the wild type (Fig. 5A, lane 4). However, this RNA may bind hrp48 (see crosslinking competition, Fig. 6D, lanes 8–10). The D#48 point mutant substrate shows a slight reduction in splicing compared to the wild-type RNA (Fig. 5A, lane 5) which might reflect better binding of hrp48 protein to the F2 region when PSI shows weaker binding to this substrate (see crosslinking Fig. 6B, lane 5). Similar inhibition of IVS3 splicing with the SELEX Up#6 substrate was seen in six independent experiments done with different Kc nuclear extract preparations (data not shown). These data suggest that the presence of a strong binding site for the PSI protein can function as an inhibitor of IVS3 splicing in vitro. However, it should be noted that the Up#6 RNA also binds other proteins

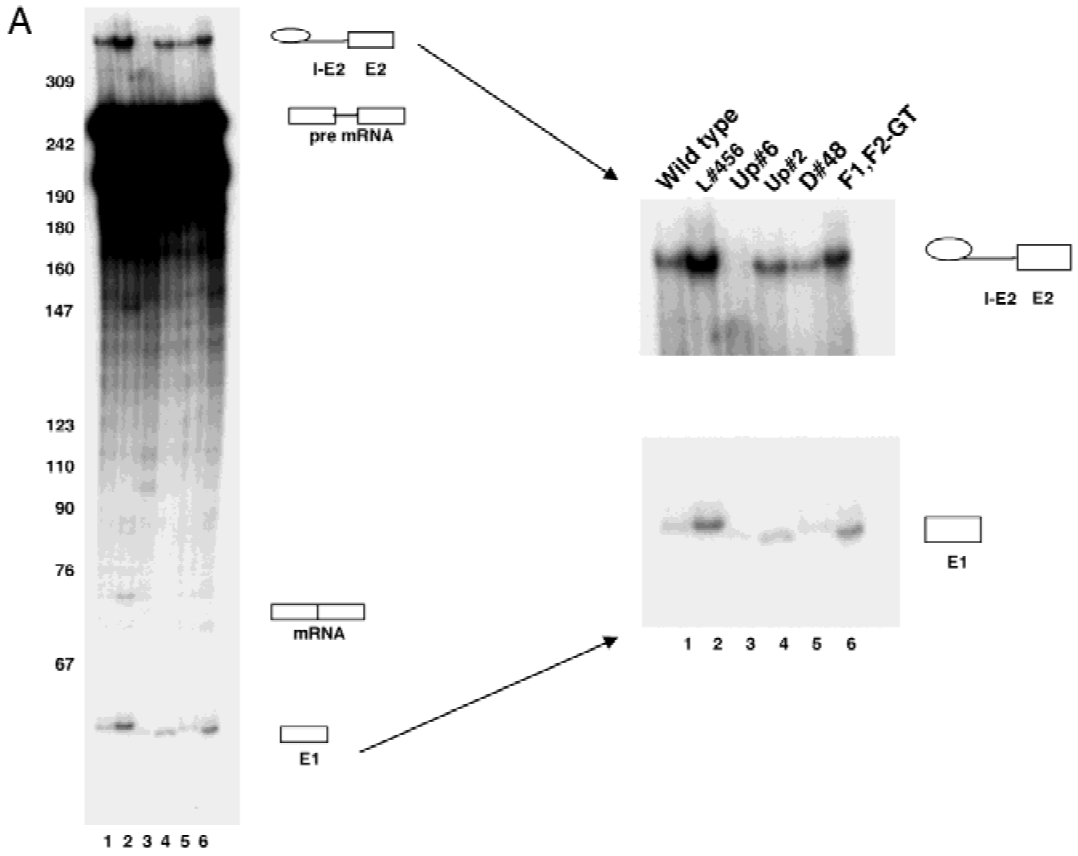
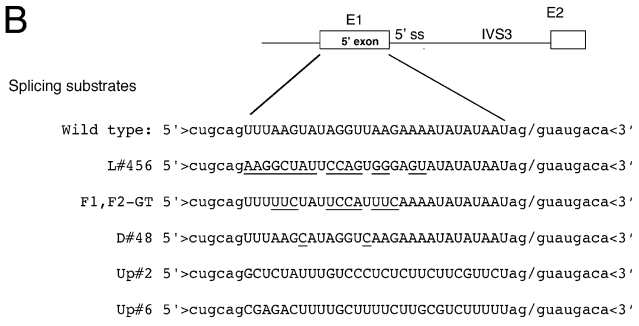



**FIGURE 5.** In vitro splicing assays with PSI SELEX RNA. **A:** P element 5′ exon sequence was substituted in splicing substrates with SELEX sequences as indicated above the lanes. Lane 1: wild-type 5′exon; lane 3: Up#6; lane 4: Up#2; lane 5: D#48; and as positive controls, lane 2: L#456; lane 6: F1,F2-GT splicing products were resolved on a 10% denaturing polyacrylamide gel. Spliced products and intermediates are schematically shown to the right of the gels. [$\alpha$-$^{32}$P]CTP labeled *Msp*I digest pBR322 molecular weight ladder is indicated by M. **B:** In vitro splicing substrate sequences. Only the 5′ end fragment of the splicing substrates are shown. The 5′ exon substituted region in these substrates is given in upper case and common sequences are given in lower case letters. Nucleotide changes in 5′ exon mutant substrates are underlined. Note that the two mutations in down SELEX D#48 RNA are wild type in the F1,F2-GT 5′ exon mutant substrate.




in the splicing extracts (see below, Fig. 6A). It has been shown that PSI is involved in the inhibition of splicing of the *Drosophila* P element third intron (IVS3) in somatic cells by binding along with several other proteins, such as hrp48, to the 5′ exon RNA (Siebel et al., 1992). Therefore, it was important to determine whether PSI protein and/or other RNA-binding proteins in the splicing extracts interact with the high affinity SELEX Up#6 RNA sequence.

### Analysis of RNA–protein complexes formed on the SELEX RNAs

To identify and characterize the RNA–protein complexes formed on the PSI-selected RNA sequences, radiolabeled RNAs were incubated in *Drosophila* Kc cell nuclear extracts under splicing conditions. One-fifth of the splicing reaction was subjected to native gel electrophoresis (Fig. 6A) and the other remaining sample was subjected to UV photochemical crosslinking and analyzed by denaturing protein gel electrophoresis (Fig. 6B). Native gel analysis shows that the wild-type 5′ exon RNA forms two or more slowly migrating RNA–protein complexes (Fig. 6A, lane 1) compared to all other RNAs tested, except the SELEX Up#6 and mutant F1,F2-GT RNAs, which formed complexes of similar mobility under these conditions in Kc cell nuclear extracts (Fig. 6A, lanes 3 and 6). The D#48 mutant and SELEX Up#2 RNAs formed weak, slowly migrating RNA–protein complexes characteristic of the wild-type

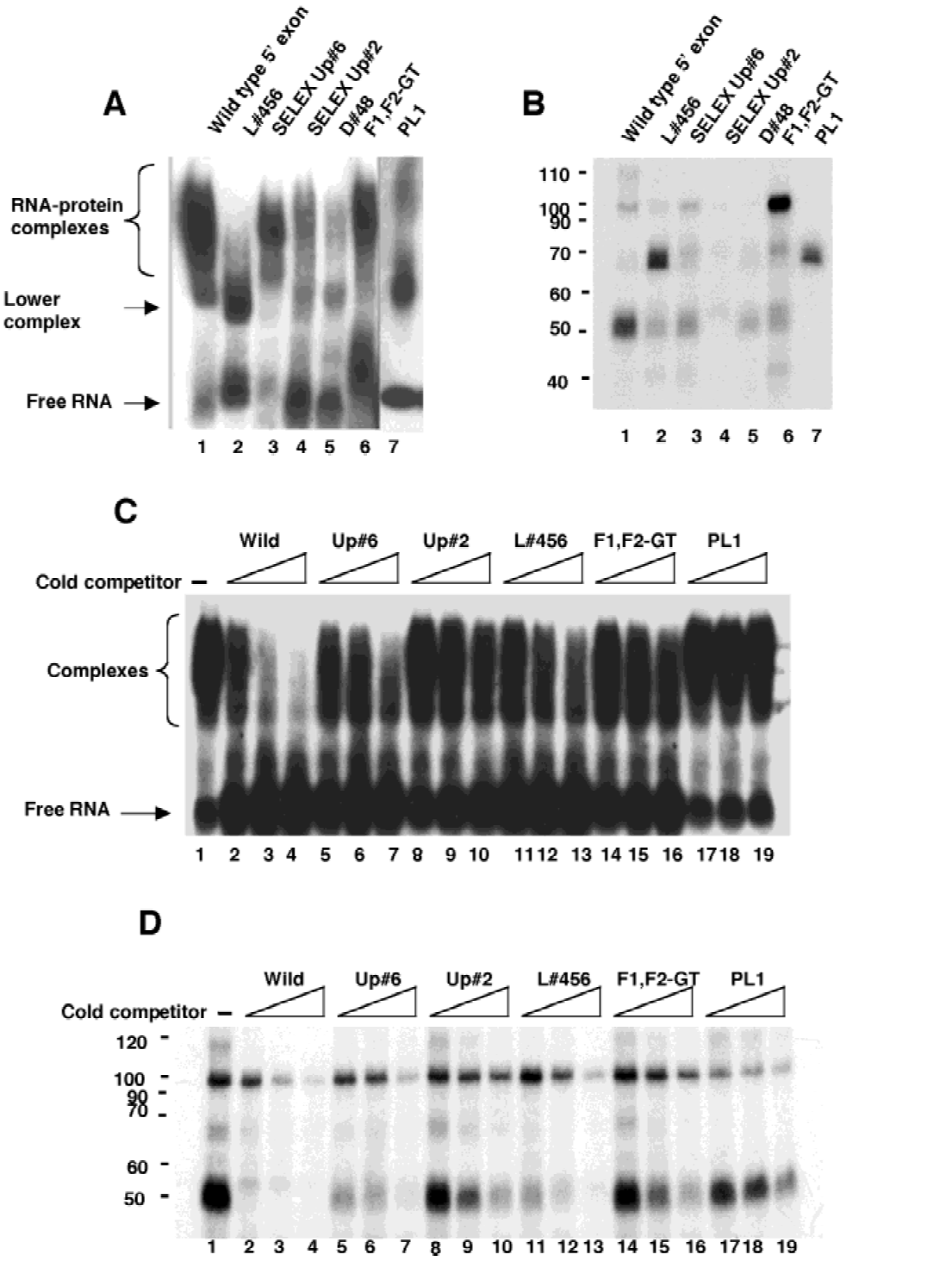



**FIGURE 6.** Analysis of RNA–protein complexes formed on SELEX RNAs. **A:** Splicing substrate templates cut with *Dra*I restriction enzyme was used to prepare short, 69-nt run-off transcripts. After incubation for 30 min with *Drosophila* Kc nuclear extracts, one-fifth of the sample was run on a native gel. The wild-type 5′ exon RNA (lane 1), L#456 (lane 2), Up#6 (lane 3), Up#2 (lane 4), D#48 (lane 5), F1,F2-GT (lane 6), and nonspecific RNA PL1 (lane 7) are shown. Arrows indicate the low molecular weight faster moving complex and free RNAs. **B:** UV crosslinking of Kc nuclear extract protein to SELEX RNAs. The remaining fraction of the reactions from **A** were used for crosslinking. Lane order as in **A**. Molecular weight markers are given at the left of the gel. **C:** Competition assay for RNA–protein complex formation. Prior to the addition of radiolabeled wild-type 5′ exon RNA, increasing amounts (50-, 100-, 150-fold molar excess) of unlabeled RNA sequences were added to the binding reaction. Lane 1 (control reaction): wild-type 5′ exon RNA probe without unlabeled RNA competitor. Lanes 2–4: wild-type RNA; lanes 5–7: SELEX Up#6; lanes 8–10: SELEX Up#2; lanes 11–13: L#456; lanes 14–16: F1,F2-GT mutant; lanes 17–19: nonspecific plasmid derived PLI RNA. **D:** Specificity of crosslinked protein to SELEX RNAs. Increasing amounts of unlabeled RNA (50-, 100-, 150-fold molar excess) were added prior to the addition of labeled RNA (lane order as in **C**) and crosslinking was done as described in Materials and Methods.

5′ exon RNA (Fig. 6A, lanes 5 and 4). The weak complex formation on these RNAs may explain the moderate level of splicing seen for Up#2 and D#48 substrates. The 5′ exon mutant L#456, as well as the plasmid-derived nonspecific PL1 RNA, did not form these slowly migrating complexes (Fig. 6A, lanes 2 and 7). These results indicate that the formation of these slowly migrating RNA–protein complexes is dependent on the presence of a PSI RNA-binding site.

The specificity of these RNA–protein complexes was tested by using these RNAs as unlabeled competitors with radiolabeled wild-type 5′ exon RNA (Fig. 6C). The various unlabeled competitor RNAs were preincubated in Kc cell nuclear extracts for 15 min prior to the addition of the radiolabeled wild-type 5′ exon RNA. The complexes formed on the wild-type RNA were specifically and strongly competed by increasing amounts of unlabeled cognate wild-type RNA (Fig. 6C, lanes 2–4), but not by the unlabeled PL1 (Fig. 6C, lanes 17–19) or Up#2 (Fig. 6C, lanes 8–10) RNA. The 5′ exon mutant F1,F2-GT RNA (Fig. 6C, lanes 14–16) and the L#456 RNA (Fig. 6C, lanes 11–13) competed weakly for complex formation (see UV crosslinking data below for an explanation). However, the Up#6 SELEX sequence competed better for complex formation than any of the other non-wild-type RNAs tested (Fig. 6C, lanes 5–7). Taken together, these data indicate that there is a correlation between the ability of these SELEX RNAs to form slowly migrating RNA–protein complexes in native gels and their activity in P element IVS3 splicing in vitro.

### Analysis of proteins UV crosslinking to the PSI SELEX sequences

To determine which proteins might be binding to the SELEX RNAs, we used UV photochemical crosslinking to [$\alpha$-$^{32}$P]UTP-labeled RNA substrates. In UV crosslinking assays, two prominent bands corresponding to the molecular weights of the previously identified PSI and hrp48 proteins (97 kDa and 50 kDa, respectively) specifically crosslinked to the wild-type 5′ exon (Fig. 6B, lane 1) and the SELEX Up#6 substrate (Fig. 6B, lane 3), whereas several of the other RNAs tested showed reduced (Fig. 6B, lane 2, L#456; lane 5, D#48) crosslinking to these two proteins. The L#456 mutant RNA shows two strong crosslinked products around 60–65 kDa and a weak crosslinked product for the hrp48 protein (Fig. 6B, lane 2). The weak competition by L#456 RNA (seen at the highest concentration tested, Fig. 6C, lane 13) for the complex formation may be due to its weak binding to hrp48. The 60–65 kDa crosslinked species were also seen with the nonspecific PL1 RNA (Fig. 6B, lane 6). The Up#2 RNA showed little or no crosslinking to PSI and hrp48 proteins (Fig. 6B, lane 4). These results are consistent with the native gel analysis with purified PSI protein (Fig. 3A) showing that Up#6 binds PSI protein better than the Up#2 RNA. The mutant substrate F1,F2-GT shows a strong crosslinked product at 97 kDa, but weak crosslinking to the product ~50 kDa, and the F1,F2-GT mutations are known to affect hrp48 binding (Siebel et al., 1994, 1995). However, the F1,F2-GT RNA is defective in competing for PSI binding (Fig. 6D), suggesting that increased UV crosslinking by the F1,F2-GT RNA reflects an altered nucleotide sequence and label transfer rather than better PSI RNA binding. This also may explain the poor competition seen with F1,F2-GT RNA for complex formation (Fig. 6C lanes 14–16). To confirm that the 97-kDa and 50-kDa crosslinked products are PSI and hrp 48, respectively, we carried out immunoprecipitation after UV crosslinking. Affinity-purified antibodies against PSI and hrp48 specifically immunoprecipitated the crosslinked proteins of the expected sizes (data not shown).

Crosslinking of proteins to RNA relies on the position of bases that absorb the shortwave UV light and their juxtaposition to amino acid side chains in the protein in a favorable geometry. To determine how specifically the crosslinked proteins from nuclear extracts bind to these RNAs, competition experiments were performed with radiolabeled wild-type 5′ exon RNA with a large molar excess of different unlabeled RNAs (Fig. 6D). The 97-kDa crosslinked product (PSI) was efficiently competed with increasing amounts of wild-type RNA (Fig. 6D, lanes 2–4) but not with the F1,F2-GT RNA (Fig. 6D, lanes 14–16) or the Up#2 RNA (Fig. 6D, lanes 8–10). The Up#6 RNA, which showed strong binding to purified PSI KH0–3 protein, competed to a moderate extent (Fig. 6D, lanes 5–7). The L#456 mutant RNA also showed some competition for PSI at the highest concentration tested (Fig. 6D, lanes 11–13). The nonspecific RNA, PL1, strongly competed for the crosslinked species around 60–65 kDa and showed a weaker but similar level of competition for the PSI and hrp48 proteins at all levels of RNA tested (Fig. 6D, lanes 17–19). The hrp48 protein is efficiently competed by the wild-type 5′ exon RNA as well as by the Up#6 and L#456 RNAs (Fig. 6D, lanes 2–4, 5–7, 11–13, respectively). The SELEX Up#2 RNA (Fig. 6D, lanes 8–10) and the F1,F2-GT RNA (Fig. 5D, lanes 14–16) showed weak competition for hrp48 at the highest concentrations tested. These competition experiments confirmed that the differences in crosslinking observed are due to the differential binding of PSI and, with exception of the F1,F2-GT RNA, are not due to differences in label transfer efficiencies between the different RNAs. These data strongly suggest that PSI protein specifically binds to the Up# 6 sequence in Kc nuclear extracts and may be directly involved in the inhibition of splicing of IVS3 in vitro. This observation is consistent with previous data where antibodies directed against the PSI protein could activate IVS3 splicing in vitro (Siebel et al., 1995).

### Antibody supershift analysis of proteins present in RNA–protein complexes formed on PSI SELEX RNAs

To further address whether PSI is present in the RNA–protein complexes that are formed on the SELEX Up#6 RNA, an antibody supershift assay was performed with Kc nuclear extracts (Fig. 7). The wild-type 5′ exon, L#456, Up#6, Up#2, D#48, and F1,F2-GT RNAs were tested in this assay in which an RNA–protein binding assay was performed and then antibody specific for either PSI or hrp48 was added and incubated for another 20 min prior to native gel electrophoresis. Because the antibody was added after RNA–protein complex formation, antibody addition should not block complex formation on these RNAs. Addition of anti-hrp48 antibody to the binding reactions, showed altered mobility of all the complexes (A, B, and C) formed on the wild-type 5′ exon RNA (Fig. 7, lane 4). However, anti-PSI antibody showed an effect only on the slowest migrating complex (complex C) present on this RNA (Fig. 7, lane 3). This is consistent with previous native gel–UV crosslinking experiments (Siebel et al., 1992). The L#456 RNA formed a better supershift complex (Fig. 7, lane 8) with anti-hrp48 antibody compared to the supershift complexes formed on the Up#6, Up#2, D#48, and F1,F2-GT RNAs. The anti-PSI antibody supershifts the complexes formed on the Up#6 RNA sequence (Fig. 7, lane 11) much better than the Up#2 RNA (Fig. 7, lane 15) and the L#456 RNA (Fig. 7, lane 7) and very weak or no supershift complexes were observed with the D#48 mutant RNA (Fig. 7, lane 19). The complex formed on the F1,F2-GT RNA was also supershifted by the anti-PSI antibody (Fig. 7, lane 23) consistent with the UV crosslinking data. These experiments indicate that PSI is a component of the complexes formed on the Up#6 SELEX RNA.

## DISCUSSION

In this study, we have identified a consensus RNA-binding sequence for the KH domain protein PSI. We showed that a high affinity PSI binding sequence is capable of acting as a splicing inhibitor in the context of a modified P element IVS3 pre-mRNA. Our data showed that PSI in Kc nuclear extracts can specifically bind to, crosslink to, and is a component of the RNA–protein complexes formed on the high affinity SELEX Up#6 RNA. These results also showed that PSI protein bound to the Up#6 RNA is involved in the inhibition of splicing in the hybrid Up#6-P element IVS3 splicing construct. Taken together with nuclear extract RNA-binding studies, these results indicate that the PSI and hrp48 proteins function together to modulate 5′ splice site selection of the P element pre-mRNA.

The multiprotein complex formed on the P element IVS3 5′ exon promotes U1 snRNP binding to a nonfunctional upstream pseudo-5′ splice site and blocks U1 snRNP binding to the downstream accurate 5′ splice site (Siebel et al., 1992). One function of PSI binding to the 5′ exon is interaction of the C terminal A and B motifs with the U1 snRNP-70K protein (Labourier et al., 2001). This interaction appears to increase U1 snRNP binding to the F1 pseudo-5′ splice site in vitro (E. Labourier, M.D. Adams, and D.C. Rio, in prep.). The

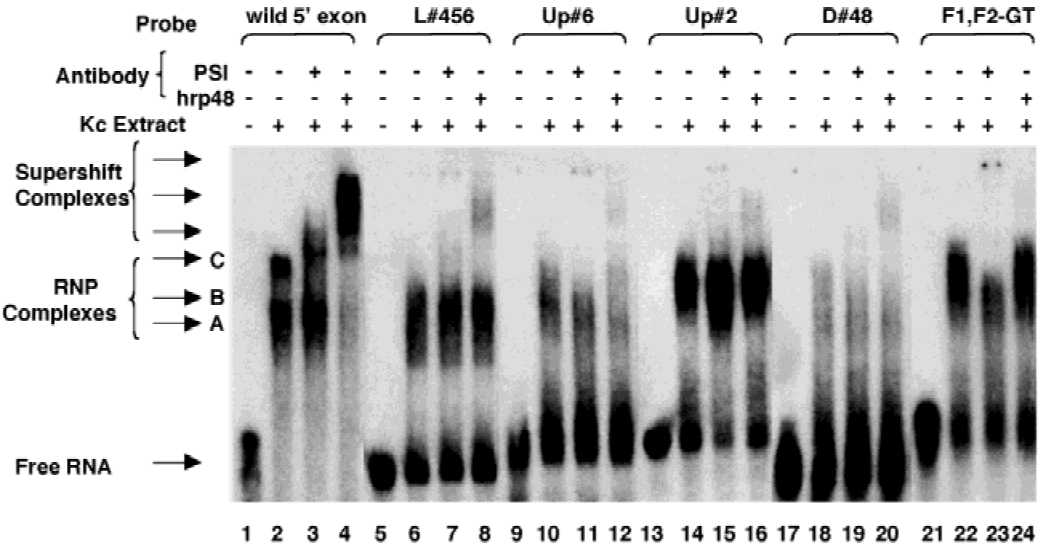



**FIGURE 7.** Antibody supershift assay. Radiolabeled RNAs, wild-type 5′ exon (lanes 1–4), L#456 (lanes 5–8), Up#6 (lanes 9–12), Up#2 (lanes 13–16), D#48 (lanes 17–20), and F1,F2-GT (lanes 20–24) were incubated with Kc nuclear extract for 30 min to allow complex assembly. Affinity purified antibody was added after 30 min to the corresponding reactions as shown above each lane. The free RNA (lanes 1, 5, 9, 13, 17, and 21) and RNA–protein complexes (A, B, and C) in the absence of antibody (lanes 2, 6, 10, 14, 18, and 22) are shown. The supershifted complexes with anti-PSI (lanes 3, 7, 11, 15, 19, and 23) and anti-hrp48 (lanes 4, 8, 12, 16, 20, and 24) antibody are marked with arrows.

RNA–protein complex formed by PSI, hrp48, and several other unidentified proteins on the high affinity Up#6 SELEX RNA may inhibit splicing by sterically blocking U1 snRNP binding to the accurate 5′ splice site and thereby preventing assembly of a functional spliceosome. This idea of a steric block to U1 snRNP binding is consistent with previous results in which the F1/F2 sites were moved away from the IVS3 5′ splice site resulting in a loss of inhibition (Siebel et al., 1992). Because PSI can directly interact with U1 snRNP through its C-terminal A and B repeats, it is also possible that stably bound PSI protein on high affinity SELEX Up#6 RNA may bring U1 snRNP to the 5′ exon RNA element in the absence of a strong U1 snRNA–pre-mRNA interaction and the presence of U1 snRNP and PSI would block the binding of U1 snRNP to the accurate 5′ splice site. Therefore, the presence of a strong PSI-binding site alone upstream of the IVS3 5′ splice site appears to be sufficient to inhibit splicing of P element IVS3 in vitro.

There are several other reports of splicing factors that recruit U1 snRNP to the pre-mRNA and stabilize U1 snRNP–pre-mRNA interactions. The yeast RNA-binding protein Mer1p, which has a single KH domain, is involved in mediating meiosis-specific alternative splicing of the MER2 pre-mRNA (Engebrecht et al., 1991). Mer1p, which is expressed only in meiosis, binds to the region downstream of the weak 5′ splice site of the MER2 transcript to facilitate meiosis-specific splicing of this intron. Mer1p associates with U1 snRNP in the absence of MER2 target RNA (Spingola & Ares, 2000). By dividing the Mer1p protein into two fragments containing either the C-terminal KH domain or the N-terminal domain, it is known that only the N-terminal domain can coimmunoprecipitate U1 snRNA and that the KH domain has no direct interaction with U1 snRNP (Spingola & Ares, 2000).

The yeast U1 snRNP-specific protein Nam8p, which is stably associated with U1 snRNA, has been shown to be required for efficient recognition of weak 5′ splice sites (Puig et al., 1999). It has been shown that Nam8p also binds to the region downstream of the weak 5′ splice site of MER2 pre-mRNA and stabilizes the U1 snRNA–pre-mRNA base pairing interaction (Puig et al., 1999). The mammalian homolog of Nam8p, the TIA-1 protein that was first identified as an effector of apoptosis (Tian et al., 1991), also has been shown to function in U1 snRNP recruitment to a weak 5′ splice site in the human apoptosis signaling receptor gene Fas pre-mRNA (Forch et al., 2000). TIA-1 interacts with the region downstream of the weak 5′ splice site in the fas gene intron 5 to promote splicing and inclusion of the downstream exon 6 yields an mRNA encoding a membrane-bound active form of the fas receptor.

Another example of potential U1 snRNP recruitment is seen in the mammalian Src pre-mRNA splicing by the KH-type splicing regulatory protein (KSRP; Min et al., 1997). KSRP contains four KH domains and is one of the mammalian homologs of the *Drosophila* PSI protein (Min et al., 1997). The tissue-specific splicing of the mouse c-src N1 exon in neurons is regulated by a multiprotein complex that forms on an intronic splicing enhancer downstream of the N1 5′ splice site (Chou et al., 2000). It has been shown that antibodies specific for the KSRP protein inhibit splicing of the N1 exon in vitro (Min et al., 1997). Therefore, during N1 exon inclusion, KSRP may be involved in recruiting U1 snRNP to the 5′ splice site by protein–protein interactions and thereby serving to activate N1 exon inclusion.

It is interesting to note that some of the splicing factors like PSI and Mer1p, identified to be involved in recruiting U1 snRNP to pre-mRNA, contain KH RNA-binding motifs. Studies of crystal structures of the KH RNA-binding domains revealed that it has a $\beta\alpha\alpha\beta\beta\alpha$ topology and the three-stranded antiparallel $\beta$ sheets oriented against the three $\alpha$ helices form a favorable surface for RNA recognition (Lewis et al., 1999). The overall specificity of this motif for RNA may be achieved through the interaction of loops and other exposed surfaces of the domain with additional interaction sites on the target RNA. Absence of a strong sequence specificity for this RNA-binding motif would explain how these different splicing factors consisting of KH motifs could recognize and bind to different or distinct regions near particular weak 5′ splice sites.

Our studies reveal a short pyrimidine-rich consensus (RCYYCUURYRC; Fig. 2A) for PSI RNA binding. Detailed RNA-binding studies on two KH domain proteins, Nova-1 and Nova-2, showed a preference for tetranucleotide UCAY sequences in their target RNAs (Yang et al., 1998; Jensen et al., 2000). The KH domain splicing factors SF1/mBBP and yBBP both recognize the YNCURAY branch point sequence (Berglund et al., 1998) and it was recently shown that the SF1/mBBP recognizes specific nucleotides in the branch point sequence (Peled-Zehavi et al., 2001). Changing the consensus branchpoint sequence UAC UAAC at position 4 from U to G had a strong negative effect on SF1/mBBP binding (Peled-Zehavi et al., 2001). These studies, in conjunction with our SELEX data for PSI indicate that KH domain-containing proteins generally recognize short pyrimidine-rich sequences in their target RNAs. The different KH domain proteins may achieve their overall affinity and specificity for target RNAs by contacting additional sequences surrounding short pyrimidine-rich core motifs. The context surrounding these core motifs therefore may play a major role in generating high affinity binding sites. The extended PSI consensus sequence we have identified (RCYYCUURYRC; Fig. 2A) shows the importance of nucleotides surrounding the CUU core in the PSI RNA-binding sites. Interestingly, the resemblance of the consensus PSI SELEX sequence to the P element 5′ exon F2 pseudo-5′ splice site also

shows the preference of PSI for the pyrimidine-rich sequences in its natural target.

The recruitment of U1 snRNP to pseudo-5′ splice sites, as seen for the P element IVS3 5′ exon, or to weak 5′ splice sites such as for MER2 and fas, could modulate the recognition of 5′ splice sites to get either inhibition or activation, respectively. Our data showed that a high affinity PSI SELEX RNA-binding sequence can function as a splicing inhibitor element and that the PSI protein is the major contributor to the inhibitory activity seen in the hybrid SELEX Up#6-P element IVS3 substrate splicing. It is possible that similar sites are found near other 5′ splice site-like sequences in PSI target cellular RNAs, in much the same way as the U-rich auxiliary regions are found near the msl-2 and fas 5′ splice sites that function in TIA-1 binding (Forch et al., 2000). Other examples of proteins with pyrimidine-rich RNA-binding sites acting as splicing inhibitors include the *Drosophila* Sex-lethal protein (Singh et al., 1995) and the hnRNPI/PTB protein (Singh et al., 1995; Chou et al., 2000). Thus, RNA-binding protein interaction with pre-mRNA sequences may act, both positively or negatively to control snRNP binding, splice site selection, and the early steps of intron recognition and spliceosome assembly.

## MATERIALS AND METHODS

### Generation of recombinant vaccinia virus vector

For recombinant virus, human 143 TK cells were infected with wild-type vaccinia virus at a multiplicity of infection equal to 0.1 in serum-free DMEM. After 1 h at 37 °C, the innoculum was removed and replaced with DMEM plus 2% fetal calf serum. The plasmid pTM1-Py-PSI was precipitated with $CaCl_2$ and the precipitates were applied to the vaccinia virus-infected cells. The infected/transfected cells were incubated for 48 h at 37 °C to allow in vivo recombination of the protein coding sequences in to the viral TK gene. The cells were harvested and the viruses were isolated. The purified viruses were amplified and titered on BSC-40 cells as described (Ausubel et al., 1987).

### Recombinant protein expression and purification

N-terminal polyoma epitope-tagged recombinant PSI was expressed from a recombinant vaccinia virus vector in human HeLa cells and purified as described below. HeLa cells were grown in suspension to a density of approximately $5 \times 10^5$ cells/mL, harvested, and resuspended in 1/10 volume serum-free media. The cells were co-infected with recombinant vaccinia viruses expressing Py-PSI and T7 RNA polymerase at a multiplicity of infection equal to 2.5 to 5 for each virus. Cells were harvested 20 h postinfection and nuclear extracts were prepared according to Dignam et al. (1983) with the following exceptions: Nuclear extraction buffer (Buffer C) contained 0.35 M NaCl and 0.15 mM DTT; the nuclei were extracted two times with Buffer C; and the nuclear extracts were not dialyzed versus Buffer D. Both nuclear extracts were spun at 70,000 × *g* for 30 min at 4 °C and the supernatants were stored at −80 °C.

For immunoaffinity purification of the Py-tagged PSI, the antibody resin was preequilibrated with buffer containing 20 mM HEPES-NaOH, pH 7.6, 10% glycerol, 0.35 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.15 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$, and 0.1% *n*-octylglucoside. The antibody resin was prepared by covalently attaching the anti-Py (Glu-Glu) antibody to protein G sepharose using dimethylpimelimidate. The nuclear extract was thawed and the clarified extract was incubated with the equilibrated immunoaffinity resin for 3 h at 4 °C, rotating. The resin was collected in a disposable Poly Prep column (BioRad) and washed with at least 50 column volumes of the following buffer: rPy-PSI, 20 mM MES-NaOH, pH 6.5, 10% glycerol, 0.35 M NaCl, 0.2 mM EDTA, 0.15 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$, and 0.1% *n*-octylglucoside. The protein was then eluted in batch with 2 column volumes of the appropriate wash buffer containing 100 μg/mL of the peptide EYMPME for 30 min at room temperature. The elution procedure was repeated three times. The elutions were brought to 0.5 mM DTT before freezing in liquid nitrogen.

Immunoaffinity purified rPy-PSI was diluted fivefold into buffer A (20 mM MES-NaOH, pH 6.5, 10% glycerol, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$), and loaded onto a POROS-SP column at 1 column volume per minute. The column was developed with a linear gradient of NaCl (from 0.07 M to 1 M) in buffer A over 20 column volumes. rPy-PSI eluted at a conductivity equal to 0.14 M NaCl.

Polyoma-tagged PSI was also expressed in *D. melanogaster* cell line L2 and purified by binding to anti-polyoma protein-G agarose beads as above. N-terminally $His_{(6)}$-tagged PSI was expressed in *E. coli* from the pRSETA plasmid (Siebel & Rio, 1990) and purified on nickel NTA-agarose followed by a MonoS column with elution in 10 mM MES-NaOH, pH 6.5, 200 mM NaCl, 10% glycerol, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSF.

### SELEX

SELEX was performed using nitrocellulose filter binding as described (Tuerk & Gold, 1990). Briefly, a single round of SELEX constituted reverse transcription of RNA using M-MLV reverse transcriptase primed with a 3′ oligonucleotide containing *Eco*RI/*Bam*HI restriction sites. With the addition of the 5′ primer T7-*Hin*dIII, sequences where amplified using 13 cycles of polymerase chain reaction (PCR). Half of the PCR product was used as template for [$\alpha$-$^{32}$P]UTP radiolabeled transcription (Gurevich, 1996). RNA was purified by phenol/chloroform extraction and ethanol precipitation then resuspended in a small volume (20 μL) of water before incubation with PSI protein in SELEX buffer (20 mM HEPES-NaOH, pH 7.6, 70 mM KCl, 5 μg/mL heparin, 0.5 mM DTT, 0.05% NP40, 0.5 mg/mL BSA) in a 2-mL volume at room temperature for 30 min. Selection was performed by filtration under vacuum through nitrocellulose filters and washing with 8 mL SELEX buffer. The percent of input RNA bound to the filter was calculated by Cherenkov counting. RNA retained on the filter was eluted by the addition of 400 μL TE-buffered phenol

and 400 μL 7 M urea and heating to 95 °C for 3 min and then ethanol precipitation.

"Down selection" was performed by two rounds of mutagenic-PCR (Cadwell & Joyce, 1992) to introduce base substitutions to the 29-nt 5′ exon of IVS3 P-element RNA followed by four rounds of negative selection by filter binding. PCR was performed using a 5′ primer, T7-*Hin*dIII to incorporate a T7 promoter site, a 3′ primer, *Eco*R1/*Bam*HI and 5′ exon of IVS3 in pGEM2 (Promega) as template (Siebel & Rio, 1990). The initial round of negative selection was performed by gel shift with either *Drosophila* (polyoma-tagged) or *E. coli*- (His-tagged) expressed full-length PSI protein. PSI protein (approximately 10 pmol) was incubated with 12.5 kcpm of freshly body-labeled [$\alpha$-$^{32}$P]UTP in 20 mM HEPES-KOH, pH 7.6, 0.05% NP-40, 100 mM KCl, 10 mg/mL heparin, 6% glycerol for 30 min at room temperature followed by separation on a 4% Tris/glycine polyacrylamide gel. Following electrophoretic separation and autoradiography, the band of unbound RNA was purified from the gel. Purified, unbound RNA was then subjected to three additional rounds of negative selection by nitrocellulose filter binding whereby the unbound RNA was recovered without washes with SELEX buffer.

### Native gel assays with SELEX RNA pool

Native gel analysis was performed as described (Siebel & Rio, 1990) in binding buffer containing 25 mM HEPES-KOH, pH 7.6, 0.05% NP-40, 100 mM KCl, 10 mg/mL heparin, 6% glycerol and separated on 4.2% (80:1) 50-mM Tris/glycine polyacrylamide gel at 140 V, at room temperature for 2.5 h.

### Splicing substrates and in vitro splicing assays

The wild-type P element splicing substrate pGEM2-IVS3wtPBP plasmid (Siebel & Rio, 1990) was used as a template for PCR-mutagenesis using appropriate primers for the SELEX sequences (Up#6, Up#2, and D#48). SELEX PCR fragments were cleaved with *Pst*I and *Xba*I restriction enzymes (Siebel et al., 1992), and subcloned in to the *Pst*I-*Xba*I sites of the pGEM2 IVS3 PBP splicing substrate. GpppG capped [$\alpha$-$^{32}$P]UTP labeled run off transcripts were synthesized using T7 RNA polymerase from *Bam*HI-linearized plasmids. Nuclear splicing extracts were prepared from Kc cells as described (Dignam et al., 1983). In vitro splicing reactions were performed essentially as described (Siebel & Rio, 1990) and products were separated on 10% polyacrylamide-urea gels.

### Native gel electrophoresis, UV crosslinking, and competition assays

The radiolabeled 69-nt-long RNA-binding substrates containing the wild-type or mutant IVS3 5′ exon or SELEX sequences were prepared from splicing substrate plasmids linearized with *Dra*I restriction enzyme. The nonspecific PL1 RNA was transcribed from the *Eco*RI-cleaved pGEM2 plasmid (Siebel & Rio, 1990). Radiolabeled RNAs were incubated in *Drosophila* Kc nuclear extracts (in the presence of 10 mg/mL heparin) under splicing conditions for 30 min and one-fifth of the sample was analyzed on a 4.2% (60:1) Tris-glycine native gel (at room temperature).

For UV crosslinking, the remaining sample was exposed to shortwave 254-nM light for 20 min. Samples were treated with RNase A (1 mg/mL) at 37 °C for 30 min and products were separated on a 8.5% SDS-polyacrylamide gel.

For competition assays, unlabeled competitor RNA was added (in molar excess) to the binding reaction prior to addition of radiolabeled RNA and reactions were continued as above.

Differentially labeled 5′ and 3′ half 5′ exon RNAs were prepared by transcription using synthetic oligonucleotide templates and T4 DNA ligase as described (Moore et al., 1993).

### Antibody supershift assays

RNA-binding reactions were performed in Kc nuclear extracts as outlined above. After the preincubation of RNA and extract, affinity-purified antibody (~200 ng) was added to the binding assay and incubated for another 20 min at room temperature. The supershifted RNA–protein complexes were separated on a 4.2% (60:1) native polyacrylamide gel at room temperature for about 3 h.

## ACKNOWLEDGMENTS

We thank Rio lab members for their support and advice, especially Emmanuel Labourier for his generous gift of reagents and Marco Blanchette for useful suggestions and clarifying the manuscript. This study was supported by National Institutes of Health Grant Number R 1 GM61987-11.

*Received May 1, 2001; returned for revision May 22, 2001; revised manuscript received May 23, 2001*

## REFERENCES

Adams MD, Rudner DZ, Rio DC. 1996. Biochemistry and regulation of pre-mRNA splicing. *Curr Opin Cell Biol 8*:331–339.

Adams MD, Tarng RS, Rio DC. 1997. The alternative splicing factor PSI regulates P-element third intron splicing in vivo. *Genes & Dev 11*:129–138.

Ausubel FM, Brent R, Kington RE, Moore DD, Seidman JG, Smith J, Struhl A. 1987. *Current protocols in molecular biology*. New York: Wiley and Sons.

Berglund JA, Fleming ML, Rosbash M. 1998. The KH domain of the branchpoint sequence binding protein determines specificity for the pre-mRNA branchpoint sequence. *RNA 4*:998–1006.

Buckanovich RJ, Darnell RB. 1997. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol Cell Biol 17*:3194–3201.

Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science 265*:615–621.

Cadwell RC, Joyce GF. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl 2*:28–33.

Chain AC, Zollman S, Tseng JC, Laski FA. 1991. Identification of a *cis*-acting sequence required for germ line-specific splicing of the P element ORF2-ORF3 intron. *Mol Cell Biol 11*:1538–1546.

Chou MY, Underwood JG, Nikolic J, Luu MH, Black DL. 2000. Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing. *Mol Cell 5*:949–957.

Dignam JD, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res 11*:1475–1489.

Engebrecht JA, Voelkel-Meiman K, Roeder GS. 1991. Meiosis-specific RNA splicing in yeast. *Cell 66*:1257–1268.

Forch P, Puig O, Kedersha N, Martinez C, Granneman S, Séraphin B, Anderson P, Valcarcel J. 2000. The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing. *Mol Cell 6*:1089–1098.

Graveley BR. 2000. Sorting out the complexity of SR protein functions. *RNA 6*:1197–1211.

Gurevich VV. 1996. Use of bacteriophage RNA polymerase in RNA synthesis. *Methods Enzymol 275*:382–397.

Jensen KB, Dredge BK, Stefani G, Zhong R, Buckanovich RJ, Okano HJ, Yang YY, Darnell RB. 2000. Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron 25*:359–371.

Kanamori H, Dodson RE, Shapiro DJ. 1998. In vitro genetic analysis of the RNA binding site of vigilin, a multi-KH-domain protein. *Mol Cell Biol 18*:3991–4003.

Labourier E, Adams MD, Rio DC. 2001. Modulation of P element pre-mRNA splicing by a direct interaction between PSI and U1 snRNP 70K protein. *Molecular Cell*, in press.

Laski FA, Rio DC, Rubin GM. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. *Cell 44*:7–19.

Laski FA, Rubin GM. 1989. Analysis of the *cis*-acting requirements for germ-line-specific splicing of the P-element ORF2-ORF3 intron. *Genes & Dev 3*:720–728.

Lasko P. 2000. The *Drosophila melanogaster* genome: Translation factors and RNA binding proteins. *J Cell Biol 150*:F51–56.

Lewis HA, Chen H, Edo C, Buckanovich RJ, Yang YY, Musunuru K, Zhong R, Darnell RB, Burley SK. 1999. Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure Fold Des 7*:191–203.

Lewis HA, Musunuru K, Jensen KB, Edo C, Chen H, Darnell RB, Burley SK. 2000. Sequence-specific RNA binding by a Nova KH domain: Implications for paraneoplastic disease and the fragile X syndrome. *Cell 100*:323–332.

Lin Q, Taylor SJ, Shalloway D. 1997. Specificity and determinants of Sam68 RNA binding. Implications for the biological function of K homology domains. *J Biol Chem 272*:27274–27280.

Lopez AJ. 1998. Alternative splicing of pre-mRNA: Developmental consequences and mechanisms of regulation. *Annu Rev Genet 32*:279–305.

Matunis MJ, Michael WM, Dreyfuss G. 1992. Characterization and primary structure of the poly(C)-binding heterogeneous nuclear ribonucleoprotein complex K protein. *Mol Cell Biol 12*:164–171.

Min H, Turck CW, Nikolic JM, Black DL. 1997. A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes & Dev 11*:1023–1036.

Misra S, Rio DC. 1990. Cytotype control of *Drosophila* P element transposition: The 66 kd protein is a repressor of transposase activity. *Cell 62*:269–284.

Moore MJ, Query CC, Sharp PA. 1993. Splicing of the precursors to mRNA by the spliceosome. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–358.

Peled-Zehavi H, Berglund J, Rosbash M, Frankel AD. 2001. Recognition of RNA branchpoint sequences by SF1/mBBP in a splicing factor complex. *Molecular and Cell Biology,* submitted.

Puig O, Gottschalk A, Fabrizio P, Séraphin B. 1999. Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection. *Genes & Dev 13*:569–580.

Siebel CW, Admon A, Rio DC. 1995. Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. *Genes & Dev 9*:269–283.

Siebel CW, Fresco LD, Rio DC. 1992. The mechanism of somatic inhibition of *Drosophila* P-element pre-mRNA splicing: Multiprotein complexes at an exon pseudo-5′ splice site control U1 snRNP binding. *Genes & Dev 6*:1386–1401.

Siebel CW, Kanaar R, Rio DC. 1994. Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI. *Genes & Dev 8*:1713–1725.

Siebel CW, Rio DC. 1990. Regulated splicing of the *Drosophila* P transposable element third intron in vitro: Somatic repression. *Science 248*:1200–1208.

Singh R, Valcarcel J, Green MR. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science 268*:1173–1176.

Siomi H, Matunis MJ, Michael WM, Dreyfuss G. 1993. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res 21*:1193–1198.

Smith CW, Valcarcel J. 2000. Alternative pre-mRNA splicing: The logic of combinatorial control. *Trends Biochem Sci 25*:381–388.

Spingola M, Ares M Jr. 2000. A yeast intronic splicing enhancer and Nam8p are required for Mer1p-activated splicing. *Mol Cell 6*:329–338.

Tian Q, Streuli M, Saito H, Schlossman SF, Anderson P. 1991. A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. *Cell 67*:629–639.

Tuerk C, Gold L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*:505–510.

Yang YY, Yin GL, Darnell RB. 1998. The neuronal RNA-binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. *Proc Natl Acad Sci USA 95*:13254–13259.