A Specific RNA Hairpin Loop Structure Binds the RNA Recognition Motifs of the *Drosophila* SR Protein B52

HUA SHI, BRYAN E. HOFFMAN, AND JOHN T. LIS*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

Received 19 December 1996/Returned for modification 21 January 1997/Accepted 10 February 1997

B52, also known as SRp55, is a member of the *Drosophila melanogaster* SR protein family, a group of nuclear proteins that are both essential splicing factors and specific splicing regulators. Like most SR proteins, B52 contains two RNA recognition motifs in the N terminus and a C-terminal domain rich in serine-arginine dipeptide repeats. Since B52 is an essential protein and is expected to play a role in splicing a subset of *Drosophila* pre-mRNAs, its function is likely to be mediated by specific interactions with RNA. To investigate the RNA-binding specificity of B52, we isolated B52-binding RNAs by selection and amplification from a pool of random RNA sequences by using full-length B52 protein as the target. These RNAs contained a conserved consensus motif that constitutes the core of a secondary structural element predicted by energy minimization. Deletion and substitution mutations defined the B52-binding site on these RNAs as a hairpin loop structure covering about 20 nucleotides, which was confirmed by structure-specific enzymatic probing. Finally, we demonstrated that both RNA recognition motifs of B52 are required for RNA binding, while the RS domain is not involved in this interaction.

Most metazoan genes are transcribed to yield pre-mRNAs containing introns that are excised and exons that are joined to create a functional mRNA (for a review, see reference 39). This pre-mRNA splicing process is catalyzed by the spliceosome, a multicomponent protein-RNA complex consisting of small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP proteins. The accurate selection and pairing of corresponding 5' and 3' splice sites occur when these factors are assembled onto the pre-mRNA in an orderly manner (for a review, see reference 32). A family of non-snRNP protein factors, termed SR proteins, have been identified as both essential splicing factors and modulators of alternative splicing that function in early spliceosome assembly (for reviews, see references 13 and 30).

SR proteins are structurally and functionally related and evolutionarily conserved. This family comprises at least six members, designated SRp20, SRp30a, SRp30b, SRp40, SRp55, and SRp75, which are conserved from *Drosophila melanogaster* to humans (50). Among them, the human SRp30a (SF2/ASF) (16, 17, 25, 26) and SRp30b (SC35) (14, 15) have been the most intensively studied. In addition, three other human SR proteins, SRp30c (38), p54 (8), and 9G8 (6), have been cloned recently. For *Drosophila*, only two SR proteins have been cloned. They are homologous to the human SRp55 and SRp20 and are designated B52 (7) and RBP1 (24), respectively.

Different SR proteins appear to have overlapping but distinct functions in constitutive and alternative splicing reactions in vitro. Although individual SR proteins are able to complement splicing-deficient cytoplasmic S100 extracts (31, 50), their activities vary when they are presented with different pre-mRNA substrates (12). In alternative splicing, one SR protein may promote preferential usage of one splice site, whereas another SR protein specifies another splice site (51). While members of the SR protein family are functionally redundant for many splicing events in vitro, the accurate splicing of a

subset of pre-mRNA in vivo may be critically dependent on one particular SR protein or combination of SR proteins.

The *Drosophila* SR protein B52 was cloned from an embryonic cDNA expression library by using a monoclonal antibody against a *Drosophila* protein associated with transcriptionally active loci on polytene chromosomes (7). It is also referred to as *Drosophila* SRp55, which was independently cloned from *Drosophila* Schneider cells by using the monoclonal antibody mAb104 (36). The activity of both *Drosophila* B52 and mammalian SRp55 in splicing was established by their ability to complement the splicing-deficient S100 extracts (31, 50). The mammalian SRp55 was also shown to promote the use of a distal alternative 5' splice site in simian virus 40 large-T splicing (52).

In vivo, B52 is an abundant protein and plays a critical role in *Drosophila* development. B52 deletion mutants are homozygous lethal at the second-instar larval stage (35), and overexpression of B52 protein causes severe developmental defects (28), evincing that B52 is an essential protein which must be expressed at the appropriate level. However, little is known about the molecular events that underlie B52's activities. Larvae deficient in B52 can still splice the five endogenous pre-mRNAs (35) and one transgenic pre-mRNA (22) tested, indicating that it is not required for general splicing in vivo. Although mammalian SRp55 has been shown to selectively promote U1 binding to the distal simian virus 40 large-T 5' splicing site, the sequence determinant for this 5' splice site selection has not been identified (52).

The SR proteins share a similar protein structure, which has at least one RNA recognition motif (RRM) at its N terminus and a C-terminally located domain rich in arginine-serine dipeptide repeats (RS domain). The RS domain has been postulated to mediate protein-protein interactions among splicing factors and regulators (48). An RRM has two highly conserved elements, referred to as RNP1 and RNP2, embedded in a less conserved sequence of approximately 80 residues. It has been found in over 200 RNA-binding proteins involved in processing and transportation of RNA (for a review, see reference 4). In many cases, proteins with RRMs have been shown to bind RNA with sequence specificity. Recent crystal-

* Corresponding author. Mailing address: Section of Biochemistry, Molecular and Cell Biology, Cornell University, 417 Biotechnology Building, Ithaca, NY 14853. Phone: (607) 255-2442. Fax: (607) 255-2428. E-mail: jtl10@cornell.edu.

lographic (33) and nuclear magnetic resonance (1) studies of the U1A snRNP protein RRM with its cognate RNA have provided insight into the RRM-RNA interaction. The RRM has a compact $\beta\alpha\beta\alpha\beta$ structure which binds single-stranded RNA on the surface of a four-stranded β -sheet. Like many other SR proteins, B52 has an extensive RS domain and two RRM, one of which conforms to the RRM consensus and the other of which is degenerate (7).

One possible mechanism by which SR proteins commit pre-mRNA to splicing requires the initial binding of a specific SR protein to the RNA. Control of alternative splicing may also be achieved by cell- or tissue-specific factors interacting with specific elements on pre-mRNA. Previously, SF2/ASF, or SRp30a, was shown to recognize purine-rich exonic splicing enhancer elements of the bovine growth hormone pre-mRNA (42). Recently, in vitro selection of high-affinity binding sites for SF2/ASF and SC35 (SRp30b) yielded distinct sets of consensus RNA sequences (43). Similar experiments done with RBP1 led to the discovery of its target sequences within two regions of the *Drosophila Dsx* pre-mRNA (21). Consistent with these findings, a single missense mutation in the first RRM of B52 resulted in a dominant second-site enhancer of RNA-processing defects (34). From these observations, it is reasonable to assume that the specific recognition of RNA sequence or structure elements may be essential for the function of B52.

In this study, we characterized the RNA-binding specificity of B52 by an in vitro approach. We first identified the specific RNA sequence recognized by B52 by using an in vitro selection technique, which yielded a relatively long consensus sequence. Unlike the RNAs that bind other SR proteins, which appear to be short and unstructured (21, 43), the B52-binding site was found to have a hairpin loop structure as predicted by a computer program by using free energy minimization and as confirmed by structure-specific enzymatic probing. We also mapped the RNA-binding site on B52 and found that both RRM are required for the interaction with the RNA aptamer. We discuss the effect of these selected B52-binding RNAs on splicing in the light of some ongoing experiments.

MATERIALS AND METHODS

Protein expression and purification. Full-length B52 protein was expressed in Sf9 cells by using the baculovirus expression system. The B52 cDNA used in this study was described previously (28) and was cloned into the vector pJVP10Z (47). The transfection, purification, and culturing of the recombinant baculovirus were performed as described previously (20, 41). The overexpressed B52 was purified by the standard SR protein purification procedure (50).

B52 deletion constructs were generated by PCR amplification of the corresponding regions of the B52 cDNA. They were then cloned into the vector pGEM-3Z (Promega Corporation, Madison, Wis.). Truncated versions of B52 were made by in vitro translation with the TNT Coupled Reticulocyte Lysate System (Promega Corporation) and L-[³⁵S]methionine (in vivo cell labeling grade; Amersham Life Science Inc., Cleveland, Ohio) according to the manufacturer's instructions. The quality of translation products was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) prior to use.

Oligonucleotides. The template-primer system consisted of three oligonucleotides identical or similar to those used by Bartel et al. (3): Temp (108 bases), 5'-ACCGAGTCCAGAGCTTGTAGTACTN₄₀GCCTAGATGGCAGTTGAATTCTCCATAGTGAGTCGTATTAC-3' (N₄₀ stands for 40 bases of random sequence); T7Univ (40 bases), 5'-GTAATACGACTCACTATAGGGAG AATTCAACTGCCATCTA-3'; and RevUniv (22 bases), 5'-ACCGAGTCCAG AAGCTTGTAGT-3'. These oligonucleotides were synthesized at a quantity of 0.2 μ M by Integrated DNA Technologies, Inc. (Coralville, Iowa).

The templates of the deletion, mutation, and antisense constructs of B52-binding sequences (BBSs) were made as oligonucleotides appended with the T7 promoter sequence. They were synthesized as Gibco BRL Custom Primers (Internet, PRIMERS@lifetech.com).

In vitro selection. (i) **Pool construction.** The gel-purified synthetic oligodeoxynucleotides Temp, T7Univ, and RevUniv were used as template and primers for a PCR (5-ml reaction mixture) carried out in 10 Eppendorf tubes. The temperature cycling was performed by manual transfer among three water baths and was stopped after four cycles when the product was still increasing nearly

twofold per cycle. The amplified DNA template was phenol extracted and ethanol precipitated.

(ii) **Cycles of selection.** RNA for each round was produced by using the T7-MEGAShortscript in vitro transcription kit for large-scale synthesis of short-transcript RNAs (Ambion Inc., Austin, Tex.) according to the manufacturer's instructions. Gel-purified RNA was quantified by spectrophotometry and diluted into 1 $\times$ binding buffer (50 mM Tris-Cl [pH 7.6], 200 mM potassium acetate, 5 mM MgCl₂, 2.5 mM dithiothreitol). The pool was heated to 70°C for 3 min and then cooled to ambient temperature over 5 min before addition of the protein. Prior to every other cycle, the folded RNA pool was also passed through an HAWP filter (Millipore Inc., New Bedford, Mass.). The binding reaction was performed in a 100- μ l volume with 30 μ M RNA and 30 nM B52 protein. Forty units of rNasin (Promega Corporation) was also included in first three cycles. The reaction mixture was incubated for 80 min with temperature cycling between 25°C (or ambient temperature) and 37°C, such that the reaction was at 25°C for three quarters of the time (2). Protein-RNA complexes were isolated by filtration and extracted as previously described (46). The selected RNAs were amplified by using the Superscript preamplification system for first-strand cDNA synthesis and Taq DNA polymerase (Life Technologies Corporation, Gaithersburg, Md.) according to the manufacturer's instructions.

(iii) **Cloning and sequencing.** The pool DNAs were cloned in the vector pGEM-3Z, between the *Eco*RI and *Hind*III sites of the polylinker region. Individual clones were sequenced by the Taq cycle sequencing method with DyeDeoxy terminators on a Applied Biosystems 737A automated DNA sequencer by the DNA Services facility at Cornell Biotechnology Programs.

RNA-protein binding assay. All binding assays were performed in 20- μ l volumes in 1 $\times$ binding buffer (see above). A typical binding assay mixture with labeled RNA contained about 20 fmol of ³²P-labeled RNA probe and different amounts (1 to 10 pmol) of B52 protein. The RNA probes were uniformly labeled with [α -³²P]UTP (Amersham Life Science Inc.) by using the T7-MAXIScript in vitro transcription kit (Ambion, Inc.) according to the manufacturer's instructions. Prior to use in a binding assay, the majority of transcripts of each RNA preparation were shown to be of the expected size by electrophoresis on an 8% polyacrylamide-7 M urea gel. A typical binding assay mixture with the truncated versions of B52 contained 1 μ l of the ³⁵S-labeled translation reaction mixture with 2 pmol of unlabeled RNA transcript (the final concentration of RNAs in this reaction was 100 nM, and the concentration of the labeled B52 proteins was much lower). The reaction mixtures were allowed to equilibrate for 15 to 20 min at ambient temperature before being subjected to filter binding, gel shift, or UV cross-linking. Gel shifts were performed at 4°C. The binding reaction mixtures were set at 4°C for 5 to 10 min before being loaded onto a 2.5% agarose gel in 0.25 $\times$ Tris-borate-EDTA buffer. The affinity of the RNA for the in vitro-translated polypeptides was estimated by comparing the intensities of the bands representing the complex on the agarose gel and those of the corresponding polypeptides run on SDS-PAGE. To cross-link proteins to RNA, the binding reaction mixtures were irradiated for 20 min from 5 cm directly above by inverting a short-wavelength UV transilluminator. The products of cross-linking were analyzed by standard SDS-PAGE.

Footprinting. The procedures used for footprinting were modifications of those previously reported (9). The RNA substrate (100 ng) was incubated with or without B52 under binding conditions. RNase T₂ (1 U) or RNase V₁ (0.35 U) was added and incubated at room temperature for 1 min. The digestion was stopped by phenol-chloroform extraction followed by ethanol precipitation. The precipitated RNAs were resuspended and analyzed by primer extension with ³²P-labeled pUC/M13 reverse sequencing primer and Superscript II reverse transcriptase (Life Technologies) and separated on a 6% (for RNase T₂ digestions) or 8% (for RNase V₁ digestions) sequencing gel.

RESULTS

In vitro selection of B52-binding RNA. In order to study the RNA-binding capacity of B52, we used, with some modification, an in vitro selection and amplification scheme (SELEX) (46). We adapted the template-primer system used by Bartel et al. (3) to produce RNA molecules having a stretch of 40 random bases sandwiched between 5' and 3' constant regions for primer annealing. The 5' constant sequence included a promoter for T7 RNA polymerase. Five hundred sixty picomoles of synthetic oligodeoxynucleotide template (20 μ g of a 108-mer) was amplified by PCR to make the original pool. Assuming that 10% of the molecules were free of chemical lesions that prevent second-strand synthesis and transcription (19), this pool contained 3.4×10^{13} different sequences.

The target protein used in the selection was B52 overexpressed from a baculovirus construct in insect cells. The B52 produced by these insect cells appears to be full-length and properly modified, as it has the same electrophoretic mobility

FIG. 1. In vitro selection of B52-binding RNA. (A) Baculovirus-expressed B52 used as target protein in the selection. Lanes 1 and 2, protein markers (Broad Range; New England Biolabs) and the B52 preparation from baculovirus-infected cells, respectively. Lanes 3 and 4, immunoblot analysis with a B52-specific antibody and either B52 produced by baculovirus-infected cells (lane 3) or a *Drosophila* nuclear extract prepared as described previously (11) (lane 4). (B) RNA sequences selected by B52 from a randomized pool. The names of the individual clones are given to the left of each selected sequence. The variable region is in capital letters, and the shared sequence motif is in boldface. Nucleotides belonging to the flanking constant region are in lowercase. Nucleotides only shared by BBS 4,14,15 and BBS 8 are underlined. The full sequence of the flanking constant region is shown at the bottom. (C) Predicted secondary structures of the B52-selected RNA sequences. Nucleotides forming the shared sequence motif are in boldface. Residues derived from the constant flanking region are in lowercase.

as the *Drosophila* protein. A Coomassie blue stain of the baculovirus B52 preparation is shown in Fig. 1A, lane 2. Immunoblot analysis (Fig. 1A, lane 3) of this same preparation with a B52-specific antibody (28) displays the identical mobility of B52 as when assayed in a *Drosophila* nuclear extract (lane 4). In addition, baculovirus-produced B52 is fully active in an in vitro splicing assay (22). Truncated versions of B52 were not used because we could not exclude a priori the involvement of the SR domain in possible sequence-specific binding.

The pool of random RNA was carried through nine rounds of selection and amplification. RNA-protein complexes were selected by binding to nitrocellulose filters. We observed a significant increase of the pool's affinity for B52 as the selection progressed (data not shown). Fractions of two final pools were cloned and sequenced. Some cloned sequences were found in duplicate or triplicate, indicating that the complexity of these selected pools was quite low. All sequences in the last pool belong to either of two families.

One family consists of four different sequences, one of which

occurred in three separate isolates (Fig. 1B). Since all six members of this family showed specific binding to B52, they are designated B52 binding sequences (BBSs). These sequences have two absolutely "conserved" hexamers separated by a variable nucleotide. There is also a "conserved" G in the flanking region of either side one base from the end of the double hexamer. This sequence motif is found in different sequence contexts and in different positions of the randomized region. In two of these sequences, BBS #8 and BBS 4,14,15, sequence similarity extends further downstream, covering a region an additional 16 nucleotides (nt) in length.

A second family of eight sequences was isolated, whose common feature is multiple G triplets. Members of this family are thus called multi-G motifs (MGMs). Most sequences in this family have four G triplets within a 25-nt region. Unexpectedly, they have no significant affinity to B52, as demonstrated below. Similar sequences were also encountered in another in vitro selection that was done simultaneously, but these sequences have not yet been fully characterized. At this

FIG. 2. Affinity and specificity of B52-binding RNA. (A) B52 binds tightly to selected sequences. Shown is a band shift assay on a 2.5% agarose gel of 20- μ l binding reaction mixtures with increasing amounts of purified B52. MGM #1 is an RNA sequence containing the MGM (5'-CGGAAGTACGGTCAAGGGATTAGGGTCA GGGGTTCCGGT-3'). G0 is a sequence randomly picked from the original pool (5'-GAGACCCACCGACACCTCGGCCGCGGGGCTTTTAGCGAG-3'). (B) Selected sequences compete for binding to B52. Reactions were performed as described for panel A except that an excess amount of cold competitor RNA was added prior to the radioactive probes.

time, we speculate that the MGM sequences were enriched via a structure formed by two or more RNAs or that they are the result of contamination from this other selection done in parallel, since RNAs in both selections were purified on the same gels. Because the individual MGMs tested do not bind B52 avidly (see Fig. 2A), this family will not be considered further in this paper.

We used the computer program MulFold (23, 53) to examine the free-energy-minimized secondary structures of the BBSs with their flanking constant region (i.e., the RNA molecules as they were selected). The conserved region in all four different sequences was predicted to fold into a common secondary structure element core (Fig. 1C): the double hexamer forms the loop and part of the stem of a hairpin loop structure, and its flanking regions form the extended stem, which may contain a bulge or internal loops. In different structures, the hairpin loop encompasses 23 to 38 nt with 6 to 12 Watson-Crick base pairs and some G-U pairs. The loops and the tops of the stems are identical in all four structures. Since the random sequence was only 40 nt in length, it is not surprising that this structural element includes sequences from the primer annealing region in some cases. Some bases in the constant region on either side of the random stretch even become part of the conserved sequence of different BBSs. These structures suggest that the common structural element may be selected for binding to B52.

Affinity and specificity of B52-binding RNA aptamers. The B52-binding RNAs were selected by repeated rounds of filter binding. When individual RNAs expressed from cloned sequences were tested by this filter binding assay, the BBS RNAs were all able to bind B52 to generate signals that were 20 times the background level produced by RNAs from the unselected pools. A band shift assay was also developed as an independent way to assess the affinities and specificities of binding of the selected sequences to B52, and these results are shown in Fig. 2. After cloning the selected sequences, we made RNA transcripts without most of the 3' constant region by cutting the template with the restriction endonuclease *ScaI*. For BBS #8 and BBS #14, this trimming did not interfere with stem formation. The band shift assay was first performed with these sequences and some members of the MGM family. Two se-

quences randomly picked from the original unselected pool were used as controls. Increasing amounts of B52 were added to 32 P-labeled RNA probes, which contain the 5' constant region, in the presence of a large excess of tRNA. Figure 2A shows a representative assay. Stable complexes were formed with BBS sequences but not with the MGM sequence or the control sequence from the unselected pool. Since only one shifted complex is observed and since B52 is a monomer (27), we assume that the interaction between B52 and RNA has a 1:1 stoichiometry. Because the concentration of the RNA probes in these reactions is negligible compared to that of B52, we were able to estimate the apparent dissociation constant of the complex as roughly equal to the protein concentration at which 50% of the RNA is bound. For example, lane 6 of Fig. 2A, where 50% of RNA is bound, indicates that the K_d of the B52-BBS #14 complex is about 50 nM. The other BBS on this gel, BBS #8, has an even higher affinity to B52, with a K_d of approximately 20 nM.

The specificity of the B52-RNA interaction was examined by subjecting it to the challenge of different specific competitors. As shown in Fig. 2B, a constant amount of B52 (5 pmol) was incubated with increasing amounts of four different unlabeled RNA competitors before the addition of either the BBS #8 or BBS 14 probe. It is obvious that both BBS #8 and BBS #14 can compete with themselves (Fig. 2B, lanes 1 to 3 and 16 to 18) and with each other (lanes 4 to 6 and 13 to 15), but the MGM and the control random sequence (G0) do not compete with these interactions (lanes 7 to 12 and 19 to 24).

To further test the specificity of the BBS, we also performed additional binding assays. RBP1, another *Drosophila* SR protein, and HSF, the *Drosophila* heat shock factor, did not show any measurable affinity for the BBS RNAs (data not shown). Also, as B52 has been shown previously to be able to cross-link with DNA in vivo (7), we tested single-stranded DNA containing the BBS motif for its ability to bind B52. We detected no B52-DNA complex in the band shift assay under conditions similar to those used for BBS RNA transcripts (data not shown). These results indicate that B52 possesses distinct RNA binding specificity for the BBS RNA.

Defining the minimum RNA sequence requirements for binding. To determine the sequence and structure require-

A			
RNA	Sequence and secondary structure	Energy	Affinity
BBS-I/Long	GGCUGG UCAACCAGGCGAC CGCCACCCGCGCGC (((---(((-----))))))-----	-13.0	++++
BBS-I	GGCUGG UCAACCAGGCGAC CGCC (((---(((-----))))))	-11.0	++++
BBS-I/NoBulge	GGCGG UCAACCAGGCGAC CGCC (((---(((-----))))))	-14.7	++++
BBS-II	GGGUAC GUAACCAGGCGAC AGUACCC (((---(((-----))))))	-10.8	++++
bbs-II	GGAC GUAACCAGGCGAC AGU --<<<<<<----->>>>>>	?	+++
AltStem	GUCAACCAGGCGAC -(((-----)))---	-3.4	+
B			
RNA	Sequence and secondary structure	Energy	Affinity
FlipBBS-II	GGCAUGAA UCAACCAGGCGAC GCAUGCC (((---(((-----))))))	-9.7	+
TransBBS-I	GGAGUG UCAACCAGGCGAC AUCC (((---(((-----))))))	-11.3	+
bbs-II/5'Stem	GGACUG UCAACCAGGCGAC AGU --(((---(((-----))))))	-8.9	+
bbs-II/3'Stem	GGACGG UCAACCAGGCGAC CGU --(((---(((-----))))))	-10.4	++++
C			
RNA	Sequence and secondary structure	Energy	Affinity
AntiBBS-I	GGCGGUCGCCUGGUAGCAGCC (((---(((-----))))))	-11.4	-
AntiBBS-II	GGACUGUGCCUGGUAGUCCU --(((---(((-----))))))	-3.5	-
UUCG/BBS	GGUCGCCUGGUAGUCC GUAACCAGGCGAC A -(((---(((-----))))))	-34.5	-

FIG. 3. Mapping of the minimal binding elements on RNA. (A) Deletion analysis defines a minimal binding site. (B) Substitution mutations define key features of the binding site. (C) Antisense sequences of the consensus do not bind B52. Short RNA transcripts were produced by in vitro transcription, and their affinities for B52 were assayed by band shift. Each + sign indicates about a threefold increase in affinity. The highest affinity is identical to that of BBS #8. The predicted secondary structure of each RNA is indicated under the sequence. Pairing bases are denoted by matching parentheses, and unpaired bases marked by dashes. The stabilities of these structures are indicated by their folding energy (in kilocalories per mole). The substituted bases in panel B are in italic. Although the base pairs represented by angle brackets were expected for the construct bbs-II as in its parental construct BBS-II, the computer program predicted an alternative structure with negligible folding energy (0.2 kcal/mol).

ments for binding by B52, we designed deletion and substitution mutations and antisense constructs of BBSs in light of the shared sequence motif and predicted secondary structures. These short RNA transcripts were produced by in vitro transcription of synthetic templates, and their affinities to B52 were assayed by band shift.

The approximate 5' and 3' termini of the aptamer binding site were determined by deletion analysis (Fig. 3A). Based on the sequence of BBS #8, we first made an RNA construct, BBS-I/Long, containing the region shared by all BBS RNAs plus the region shared only with BBS #14 but missing both 5' and 3' constant regions. It is predicted to retain the BBS #8 hairpin loop structure, which is the shortest and most stable hairpin among the four BBSs, plus a 3' unfolded tail. This construct has full binding activity. Next, we deleted the 3' tail

to make the construct BBS-I, which consists of just the hairpin loop structure, and it was also fully active. When we deleted the bulged U in the stem to make the construct BBS-I/NoBulge, the binding activity was still not compromised.

Another construct, BBS-II, which is part of the BBS #11 sequence and has the conserved hairpin loop with an internal loop in the stem, binds B52 as well as BBS-I. However, when several nucleotides on each side of BBS-II were deleted to decrease the stability of the stem, the affinity for B52 decreased almost by a factor of three. This construct, bbs-II, which is comparable to BBS-I in length and contains the sequence shared by all members of BBS family, was predicted to have no stable secondary structure in solution. However, we cannot rule out the possibility that non-Watson-Crick base pairs, base-ribose interactions, and/or base-phosphate interactions may occur in the internal loop region to produce a compact and stable structure or that binding to B52 may stabilize the remaining weak pairs to give a structure like the one in the full sequences. The last deletion construct, AltStem, contains only the double-hexamer region plus one of the two conserved Gs at the 5' end. The resulting predicted structure, a smaller and slightly different hairpin, still retained some low affinity for B52, but much less than that of bbs-II.

The data described above suggest that both the primary sequence and secondary structure of the RNAs contribute to their affinity for B52. Therefore, we made substitution variants to separate the effects of sequence and structure (Fig. 3B). Based on the sequence of BBS-II, we made the variant FlipBBS-II, in which part of the sequences in each half of the stem were interchanged. Based on the sequence of BBS-I, we made the variant TransBBS-I by two GC-to-AU transitions in the stem. Both of these constructs, in which the conserved flanking Gs were replaced, were poor binders even though their predicted secondary structures were unaffected by the changes. Then, based on bbs-II, we tried to convert the unstable internal loop region into a stem by changing sequences on either side of the loop, resulting in the variants bbs-II/5'Stem and bbs-II/3'Stem, respectively. Although both have a similarly stable hairpin loop structure, the former, which has only one conserved flanking G, was a weak binder, but the latter, which has both conserved flanking Gs, was a strong one. In making the strong binder, bbs-II/3'Stem, we generated a sequence identical to that of BBS-I between the two endmost conserved Gs.

As a negative control, we also made antisense RNAs of BBS-I and bbs-II (Fig. 3C). Both antisense constructs did not bind B52, although AntiBBS-I has a hairpin loop structure similar to that of BBS-I. In addition, we tested the importance of the single strandedness of the core BBS consensus by linking it to its antisense sequence via a UUCG tetraloop. The BBS consensus sequence trapped in an RNA duplex lost its binding activity.

These experiments demonstrated that the sequence between and including the two endmost conserved Gs is required for efficient recognition by B52. While most of the conserved sequence resides in the loop, these variants showed that a particular sequence of the stem (the conserved flanking Gs) is also critical for this interaction. The flanking regions on both sides of this conserved sequence segment can contribute to binding in cases where they help form a more stable stem, probably by prepaying some entropic cost of the interaction. Thus, a strong B52-binding site on RNA is at least 17 nt (the core BBS, G through G) or around 20 nt (the length of most hairpin loops tested) in length.

Probing the secondary structure and critical regions in the B52-binding site of RNA. In order to delineate the B52-binding

FIG. 4. Localization of the B52-binding site on RNA. (A) RNase footprinting with RNase T₂ (lanes 2 to 4) or RNase V₁ (lanes 9 to 11) identifies the hairpin loop of BBS #8 as the BBS. RNase was added to binding reaction mixtures lacking B52 (lanes 2 and 9), with 2 pmol of B52 (lanes 3 and 10), or with 20 pmol of B52 (lanes 4 and 11), and the resulting RNA products were analyzed by primer extension. The input RNA is shown to identify nicks in the RNA as well as reverse transcriptase stops (lane 1). The sequencing ladder (lanes 5 to 9) is used to identify the bases of interaction. Open circles indicate weak T₂ protection, and closed circles and squares indicate strong T₂ and V₁ protection, respectively. The single-stranded regions and the sequence in the stem of the hairpin loop structure are indicated next to the sequencing ladder by thin and thick lines, respectively. (B) Partial predicted secondary structure of the BBS #8 RNA in the context of vector sequence. The sequence of BBS #8 is indicated by the nucleotides, and the constant and vector sequences are depicted by a skeletal diagram. Symbols are as in panel A.

region on the selected RNA at a higher resolution, we performed RNase footprinting with the cloned BBS #8. Two RNases with different specificities were used to determine the region of binding in both single- and double-stranded areas of RNA. These experiments have the added benefit of confirming the predicted hairpin loop secondary structure of the selected RNA.

RNase T₂ is a single-strand-specific endoribonuclease with low specificity (a preference for adenines), while RNase V₁ cleaves RNA predominantly at double-stranded regions with no apparent sequence specificity (9). Both free RNA and RNA-B52 mixtures were treated with RNases T₂ and V₁, and the resulting RNA products were detected by primer extension. This allows an accurate identification of the protected nucleotides, since sequencing reactions with the same primer can be run alongside the footprinting reactions. A comparison of the primer extension products from free RNA with those from RNA-B52 mixtures identifies residues that are in contact with the protein. As the B52 concentration is increased, the intensities of several bands decrease (Fig. 4A, lanes 2 to 4 and 9 to 11). Comparing the positions of these bands directly to the sequencing ladder identifies bases that are protected by the binding of B52 (Fig. 4B). These B52-protected bases reside in the predicted hairpin loop region. In addition, the cleavage of the RNA substrate by these RNases in the absence of B52 confirms the predicted hairpin loop secondary structure of the selected RNA. The RNase protection results provide physical evidence that the conserved hairpin loop of BBS RNA is the target of B52 interaction.

Locating the RNA-binding site of B52 in both RRM. BBS RNAs were selected for their ability to bind full-length B52 protein. It is possible that B52 binds RNA through either one or the other RRM, the SR domain, or a combination of these domains. To identify the specific RNA-binding site on B52, we

made a set of [³⁵S]methionine-labeled B52 deletion constructs (Fig. 5A) by *in vitro* translation and tested their ability to bind the selected BBS #8 RNA in two different assays. These polypeptides also serve as a second source of target proteins in binding assays to confirm the specificity of binding to the selected BBS RNA.

The *in vitro*-translated polypeptides were examined by SDS-PAGE to verify that they were of the predicted sizes before being used in the binding assays. The integrity of two polypeptides, R1 and R12, is shown in Fig. 5C (see below). Other polypeptides were made in an identical way and had similar quality. Figure 5B shows a band shift assay with ³⁵S-labeled polypeptides individually or in combination. The binding reaction was set up as for the experiments described above but with a constant amount of BBS #8 RNA transcript that was in a molar excess relative to protein. In the presence of BBS #8 RNA, ³⁵S-labeled full-length B52 (F) showed a discrete band (Fig. 5B, lane 2), which has the same mobility as that of a B52-BBS #8 complex generated with ³²P-labeled RNA (Fig. 2). The polypeptide containing both RRMs (R12) gave rise to a band running faster than the full-length B52-BBS #8 complex (Fig. 5B, lane 4). The SR domain (S) (Fig. 5B, lane 6), RRM I (R1) (lane 10), and a polypeptide containing both RRM II and the SR domain (R2S) (lane 12) did not produce any discrete band on this native gel. When R12 and S were put together in the presence of the RNA, a band (Fig. 5B, lane 8) with mobility similar to that seen with R12 (Fig. 5B, lane 4) was observed. No shifted band was produced with the other combination, polypeptides R1 and R2S. These data indicate that the SR domain of B52 does not participate in the interaction with the RNA aptamer and that both RRMs are required for the binding. In addition, the affinity of the RNA for both *in vitro*-synthesized polypeptides F and R12 is comparable to that

FIG. 5. RNA-binding site on B52 localized to both RRM domains. ^{35}S -methionine-labeled B52 deletion constructs (A) (aa, amino acids) were made by *in vitro* translation, and their ability to bind BBS #8 RNA was examined in two different assays. The binding reactions with different ^{35}S -labeled proteins or combinations thereof were monitored by band shift assay on a native agarose gel (B) or UV cross-linking and SDS-PAGE (C). The closed arrowhead in panel C signifies the B52-R12-BBS #8 complex, whereas the open arrowhead points to where the B52-R1-BBS #8 complex would be expected to migrate.

measured previously (Fig. 2) by using baculovirus-produced B52 protein in excess over RNA (see Materials and Methods).

To confirm the requirements of both RRM domains for binding, we repeated these binding reactions and subjected the reaction mixtures to UV irradiation. The UV-cross-linked species were then analyzed by SDS-PAGE. As shown in Fig. 5C, R12 can be cross-linked to the BBS 8 RNA, forming a complex with an apparent molecular weight the sum of those of the protein and RNA components (lane 6). In contrast, R1 was not cross-linked to RNA (Fig. 5C, lane 3). In these reactions, equal molar amounts of protein were used. Besides, both RRM domains must be *cis* to bind RNA, since R2S cannot complement R1 to restore the binding even when the RNA is at much a higher concentration (data not shown).

DISCUSSION

The SR proteins are a structurally related family of proteins that play a key role in early pre-mRNA splicing events (13, 30). Not only are they critical in committing pre-mRNAs to splicing *in vitro*, but they also control the splice site choice of alternatively spliced transcripts. These functions are likely to be important in normal development, although the true *in vivo* targets of many of these factors have remained elusive. The knockout of the B52 gene in *Drosophila* (35) currently provides the only demonstration of an essential role of an SR protein family member during development. B52 possesses two RRM domains as well as a large basic domain composed of multiple strings of RS dipeptide repeats, and thus it is expected to interact directly with nucleic acid (7). In an attempt to identify sequences in RNA that are recognized by B52, we have isolated RNA aptamers from a large population of random RNAs that bind tightly to B52. The BBS we defined is longer than many consensus sequences selected by other SR proteins and non-SR proteins containing RRM domains (7, 21, 43, 45). In addition, it has some symmetry and forms a hairpin structure. We find that both RRM domains of the B52 protein act together to bind an RNA hairpin in a sequence-specific interaction.

Both RRM domains are involved in binding. *In vitro* selection (SELEX) has been used to select oligonucleotides that bind to proteins that recognize nucleic acid as their normal function (3, 29, 37, 46) as well as to proteins that do not bind nucleic acid by nature (10, 44). The B52 protein consists of three different domains that are potentially capable of binding RNA,

i.e., the two RRM domains and the RS domain. Since its large and positively charged RS domain has the potential to interact with polyanionic RNA either specifically or nonspecifically, we used full-length B52 in our selection scheme to make sure that all epitopes on the protein surface were initially available for interaction with the RNA pool. The RNA aptamer that was eventually selected bound to both RRM domains but not the RS domain of B52, consistent with the findings for other SR proteins such as SF2/ASF and SC35 (43). Moreover, the presence of the RS domain does not compromise the binding of B52 to RNA. However, in contrast with two previous *in vitro* selection experiments with SR proteins, which yielded short, unstructured consensus sequences (21, 43), the BBSs are long and structurally well defined.

By analogy to other RRM domains such as that of the U1A protein, whose structure is known, we can make some predictions about the structure and RNA-binding properties of B52. In the crystal structure of the U1A protein RRM-RNA complex, the RNA loop binds to the protein as an open structure, and the β -sheet provides a large surface of extensive interaction with splayed-out nucleotide bases (33). Bases of seven nucleotides in the loop are involved in direct or water-mediated hydrogen bonds. The core BBS is long enough to comprise two such binding sites, interacting with both RRM domains of B52 in a similar manner. Upon binding B52, the loop region of the BBS RNA may be enlarged to accommodate both RRM domains. Alternatively, sequence information may be read from the stem as well as the loop. Similar to our findings, heterogeneous nuclear ribonucleoprotein particle A1, which also contains two RRM domains, selected an RNA aptamer having two nearly identical hexamers separated by two variable nucleotides (5).

As mentioned above, the N-terminal RRM of B52 displays a very high similarity to the RRM consensus sequence, while the second RRM is considerably different from the consensus sequence. Although single RRM domains of SR proteins have been shown to bind RNA or to be able to select RNA aptamers, the overall binding specificity of the intact protein containing multiple RRM domains may not be the sum of the specificities of the individual RRM domains. The B52-binding RNA aptamer that we selected failed to bind either RRM. The two RRM domains of B52 appear to act as a single RNA-binding composite, cooperating in some way to determine both the RNA binding affinity and specificity of the intact molecule.

Selection of the B52-binding sequences. RNA molecules are generally sequence-dependent folded structures rather than simple strings of nucleotides. Current evidence concerning protein-RNA interaction suggests that proteins often make sequence-specific contacts with single-stranded residues in loops, while base-paired regions are often necessary for presentation of single-stranded nucleotides in the correct orientation (49). To maximize the structural information provided by the RNA molecule in an RNA-protein interface, the random region used in in vitro selection should be as long as possible. However, in most cases the complete representation of a variable region that is longer than 25 nt would exceed the practical limit of oligonucleotide synthesis. In addition, we found that the yield and quality of the product decrease dramatically when its total length exceeds 100 nt in most oligonucleotide synthesis facilities. Taking these factors into account, we decided to use a template with a random region of 40 nt. This size and the required constant flanking region would require the synthesis of a template slightly over 100 nt in length. Such RNAs would be able to accommodate at least one stable hairpin loop structure or be long enough to be folded into a pseudoknot.

The core BBS site that we selected is less than 20 nt in length. To cover the sequence space of a 20-mer, 4^{20} or 10^{12} different sequences are needed. This requirement was met in our selection, since the original pool contains at least 7.1×10^{14} different 20-mer sequences (i.e., 3.4×10^{13} 40-mers). Therefore, the B52-binding sequences we selected may well be the best ones available, a notion also supported by the nearly perfect consensus we observed in unrelated sequence contexts. Nine rounds of selection should be sufficiently rigorous to account for the low complexity of the final selected pool.

The potential significance of B52's RNA-binding specificity. By in vitro selection, we examined the inherent RNA-binding capacity of B52 by exposing it to a vast sequence and structure space of RNA. This does not necessarily mean that we have found the native RNA sequence and structure with which B52 interacts in vivo. It is possible that the native B52-binding sites on *Drosophila* RNA are more degenerate or shorter than the BBS we selected. Nonetheless, this RNA sequence can be used to help identify potential targets of B52 in the *Drosophila* genome. Indeed, we found the second hexamer of BBS (5'-G GCGAC-3') in an exon of *ftz* pre-mRNA, which has been used as the splicing substrate in a functional assay of B52 (22). This site is located near the 5' splice site in secondary structure, and a deletion that includes this site abolishes splicing. Our preliminary results showed that this pre-mRNA substrate is able to bind B52 in a BBS-competable manner with an affinity lower than that of the BBS. A search for BBS-like sequences in the *Drosophila* sequence database yielded several *Drosophila* genes matching the core BBS consensus sequence at 11 or more of the 14 invariant positions. The pre-mRNAs encoded by these genes are prime candidates for B52 action, and investigation of B52's possible in vivo role in splicing these pre-mRNAs should benefit from the availability of *Drosophila* B52 mutants (34, 35).

Aptamers as reagents in vivo. The aptamers we selected for B52 provide a useful tool to study B52 function both in vitro and in vivo. RNA aptamers have been shown to have a penchant to home to functional sites on protein surfaces and discriminate finely among these sites (18). The aforementioned result regarding the splicing of *ftz* pre-mRNA suggests that the BBS may interact with a functionally significant region of B52, and this binding may be strong enough to disrupt its normal interaction with pre-mRNA substrates. This has encouraged us to make use of BBS RNA as a genetic tool for dissecting B52

function in vivo, by creating *Drosophila* fly lines designed to overproduce BBS RNAs (40). We envision that the controlled and modulated expression of these RNA aptamers in living flies will serve as a specific, reversible, fast-acting B52 inhibitor that should provide insight into the function of B52 in vivo.

ACKNOWLEDGMENTS

We are grateful to Jeffrey Roberts, Mariana Wolfner, and Elizabeth Keller for their critical comments on the manuscript.

This work was supported by National Institutes of Health grant GM40918 to J.T.L.

REFERENCES

- Allain, F. H.-T., C. C. Gubser, P. W. A. Howe, K. Nagai, D. Neuhaus, and G. Varani. 1996. Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. *Nature* **380**:646-650.
- Bartel, D. P., and J. W. Szostak. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* **261**:1411-1418.
- Bartel, D. P., M. L. Zapp, M. R. Green, and J. W. Szostak. 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67**:529-536.
- Burd, C. G., and G. Dreyfuss. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**:615-621.
- Burd, C. G., and G. Dreyfuss. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* **13**:1197-1204.
- Cavaloc, Y., M. Popielarz, J. P. Fuchs, R. Gattoni, and J. Stevenin. 1994. Characterization and cloning of the human splicing factor 9G8: a novel 35kDa factor of the serine/arginine protein family. *EMBO J.* **13**:2639-2649.
- Champlin, D. T., M. Frasca, H. Saumweber, and J. T. Lis. 1991. Characterization of a *Drosophila* protein associated with boundaries of transcriptionally active chromatin. *Genes Dev.* **5**:1611-1621.
- Chaudhary, N. C., C. McMahon, and G. Blobel. 1991. Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. *Proc. Natl. Acad. Sci. USA* **88**:8189-8193.
- Christiansen, J., J. Egebjerg, N. Larsen, and R. A. Garrett. 1990. Analysis of rRNA structure: experimental and theoretical considerations, p. 229-252. In G. Spedding (ed.), *Ribosomes and protein synthesis: a practical approach*. IRL Press, New York, N.Y.
- Conrad, R., L. M. Keranen, A. D. Ellington, and A. C. Newton. 1994. Isozyme-specific inhibition of protein kinase C by RNA aptamers. *J. Biol. Chem.* **269**:32051-32054.
- 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 Acid Res.* **11**:1475-1489.
- Fu, X.-D. 1993. Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature* **365**:82-85.
- Fu, X.-D. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**:663-680.
- Fu, X.-D., and T. Maniatis. 1992. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science* **256**:535-538.
- Fu, X. D., and T. Maniatis. 1992. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proc. Natl. Acad. Sci. USA* **89**:1725-1729.
- Ge, H., and J. L. Manley. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**:25-34.
- Ge, H., P. Zuo, and J. L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell* **66**:373-382.
- Gold, L., B. Polisky, O. Uhlenbeck, and M. Yarus. 1995. Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**:763-797.
- Green, R., A. D. Ellington, D. P. Bartel, and J. W. Szostak. 1991. In vitro genetic analysis: selection and amplification of rare functional nucleic acids. *Methods* **2**:75-86.
- Groebe, D. R., A. E. Chung, and C. Ho. 1990. Cationic lipid-mediated co-transfection of insect cells. *Nucleic Acids Res.* **18**:4033.
- Heinrichs, V., and B. S. Baker. 1995. The *Drosophila* SR protein RBP1 contributes to the regulation of doublesex alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J.* **14**:3987-4000.
- Hoffman, B. E., and J. T. Lis. Unpublished results.
- Jaeger, J. A., D. H. Turner, and M. Zuker. 1989. Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. USA* **86**:7706-7710.
- Kim, Y. J., P. Zuo, J. L. Manley, and B. S. Baker. 1992. The *Drosophila* RNA-binding protein RBP1 is localized to transcriptionally active sites of chromosomes and shows a functional similarity to human splicing factor ASF/SF2. *Genes Dev.* **6**:2569-2579.
- Krainer, A. R., G. C. Conway, and D. Kozak. 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**:1158-1171.

26. **Krainer, A. R., A. Mayeda, D. Kozak, and G. Binns.** 1991. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and *Drosophila* splicing regulators. *Cell* **66**:383–394.
27. **Kraus, M. E.** 1995. The *Drosophila melanogaster* RNA splicing factor B52: overexpression studies and mechanisms of recruitment to nuclear targets. Ph.D. thesis. Cornell University, Ithaca, N.Y.
28. **Kraus, M. E., and J. T. Lis.** 1994. The concentration of B52, an essential splicing factor and regulator of splice site choice, is critical for *Drosophila* development. *Mol. Cell. Biol.* **14**:5360–5370.
29. **Levine, T. D., F. Gao, P. H. King, L. G. Andrews, and J. D. Keene.** 1993. Hel-N1: an autoimmune RNA-binding protein with specificity for 3' uridylate-rich untranslated regions of growth factor mRNAs. *Mol. Cell. Biol.* **13**:3494–3504.
30. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. *Genes Dev.* **10**:1569–1579.
31. **Mayeda, A., A. M. Zahler, A. Krainer, and M. Roth.** 1992. Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing. *Proc. Natl. Acad. Sci. USA* **89**:1301–1304.
32. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to mRNAs by the spliceosome, p. 303–357. *In* R. F. Gesteland and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Oubridge, C., N. Ito, P. R. Evans, C.-H. Teo, and K. Nagai.** 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**:432–438.
34. **Peng, X., and S. M. Mount.** 1995. Genetic enhancement of RNA-processing defects by a dominant mutation in B52, the *Drosophila* gene for an SR protein splicing factor. *Mol. Cell. Biol.* **15**:6273–6282.
35. **Ring, H. Z., and J. T. Lis.** 1994. The SR protein B52/SRp55 is essential for *Drosophila* development. *Mol. Cell. Biol.* **14**:7499–7506.
36. **Roth, M. B., A. M. Zahler, and J. A. Stolk.** 1991. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J. Cell Biol.* **115**:587–596.
37. **Schneider, D., C. Tuerk, and L. Gold.** 1992. Selection of high affinity RNA ligands to the bacteriophage R17 coat protein. *J. Mol. Biol.* **228**:862–869.
38. **Screaton, G. R., J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, and A. R. Krainer.** 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.* **14**:4336–4349.
39. **Sharp, P. A.** 1994. Split genes and RNA splicing. *Cell* **77**:805–815.
40. **Shi, H., and J. T. Lis.** Unpublished results.
41. **Summers, M. D., and G. E. Smith.** 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. Bulletin no. 1555. Texas Agricultural Experiment Station, College Station.
42. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.* **7**:2598–2608.
43. **Tacke, R., and J. L. Manley.** 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J.* **14**:3540–3551.
44. **Tian, Y., N. Adya, S. Wagner, C.-Z. Giam, M. R. Green, and A. D. Ellington.** 1995. Dissecting protein:protein interactions between transcription factors with an RNA aptamer. *RNA* **1**:317–326.
45. **Tsai, D. E., D. S. Harper, and J. D. Keene.** 1991. U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res.* **19**:4931–4936.
46. **Tuerk, C., and L. Gold.** 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**:505–510.
47. **Ueda, A., S. Kawamoto, T. Igarashi, Y. Ishigatsubo, K. Tani, T. Okubo, and K. Okuda.** 1994. Human monocyte chemoattractant protein-1 expressed in a baculovirus system. *Gene* **140**:267–272.
48. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**:1061–1070.
49. **Wyatt, J. R., and I. Tinoco, Jr.** 1993. RNA structural elements and RNA function, p. 465–496. *In* R. F. Gesteland and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
50. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**:837–847.
51. **Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth.** 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science* **260**:219–222.
52. **Zahler, A. M., and M. B. Roth.** 1995. Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. *Proc. Natl. Acad. Sci. USA* **92**:2642–2646.
53. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. *Science* **244**:48–52.