RNA (2001), *7*:29–43. Cambridge University Press. Printed in the USA.

A ribozyme selected from variants of U6 snRNA promotes 2′,5′-branch formation

THOMAS TUSCHL,[1,2,3,4] PHILLIP A. SHARP,[1,2] and DAVID P. BARTEL[1,3]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[2]Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[3]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA

ABSTRACT

In vitro selection was used to sample snRNA-related sequences for ribozyme activities, and several 2′,5′-branch-forming ribozymes were isolated. One such ribozyme is highly dependent upon an 11-nt motif that contains a conserved U6 snRNA sequence (ACAGAGA-box) known to be important for pre-mRNA splicing. The ribozyme reaction is similar to the first step of splicing in that an internal 2′-hydroxyl of an unpaired adenosine attacks at the 5′-phosphate of a guanosine. It differs in that the leaving group is diphosphate rather than a 5′ exon. The finding that lariat formation can be accomplished by a small RNA with sequences related to U6 snRNA indicates that the RNA available in the spliceosome may be involved in RNA-catalyzed branch formation.

Keywords: in vitro selection; RNA catalysis; splicing

INTRODUCTION

The constitution of the spliceosome catalytic center has not been determined. The finding that both group II self-splicing RNA and spliceosomes generate lariat RNAs has led to the speculation that the snRNA components present in the spliceosome might be catalytically active (Sharp, 1985; Cech, 1986; Guthrie, 1991; Steitz & Steitz, 1993). Because there are so few sequence and structural similarities between snRNAs and self-splicing group II introns, an RNA-based mechanism cannot be inferred from their comparison (for discussion, see Weiner, 1993; Michel & Ferat, 1995). On the other hand, the striking similarity in the stereochemistry of these splicing reactions (Maschhoff & Padgett, 1993; Moore & Sharp, 1993; Padgett et al., 1994) and additional parallels to other RNA-catalyzed reactions (Sontheimer et al., 1999, 1997) argue for snRNA-based catalysis in the spliceosome.

Significant progress has been made in defining the snRNA and pre-mRNA interactions in the catalytically competent form of the U2-type spliceosome (reviewed in Madhani & Guthrie, 1994a; Umen & Guthrie, 1995; Nilsen, 1998; Staley & Guthrie, 1998). The current model of snRNA and pre-mRNA interactions in the yeast spliceosome is shown in Figure 1. U2 and U6 snRNAs are engaged in base pairing interactions to juxtapose the splice site residues of the pre-mRNA substrate. Despite initial evidence for an important role of the conserved loop I of U5 snRNP in this process (reviewed in Newman, 1997), the loop appears to be dispensable (O'Keefe et al., 1996; Segault et al., 1999). Both U1 snRNA and U4 snRNA are released from the spliceosome before RNA chemistry occurs.

U6 snRNA and the 5′-terminal domain of U2 snRNA are conserved across diverse species (Guthrie & Patterson, 1988; Roiha et al., 1989). Mutational analyses of U2 and U6 snRNAs in yeast indicate that not all conserved residues are absolutely critical. Of the 29 phylogenetically conserved residues in yeast U6 snRNA, only 14 show a phenotype when mutated singly and even fewer residues are critical in the similarly highly conserved 5′-terminal domain of U2 snRNA. In mammalian systems, these conserved regions appear even more permissive to mutations (e.g., Datta & Weiner, 1993; Wolff et al., 1994). The most critical U6 snRNA residues cluster in two regions and are named according to their conserved sequences as the ACAGAGA-box and the AGC-triad (Fabrizio & Abelson, 1990; Madhani et al., 1990; Datta & Weiner, 1993; Wolff et al., 1994; McPheeters, 1996). Both elements have critical functions in 5′ splice site selection (Wassarman & Steitz, 1992; Kandels-Lewis & Seraphin, 1993; Lesser & Guthrie, 1993; Sun & Manley, 1995; Hwang & Cohen, 1996;

Reprint requests to: David Bartel, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA; e-mail: dbartel@wi.mit.edu.

[4]Present address: Department of Cellular Biochemistry, Max-Planck-Institute for Biophysical Chemistry, D-37070 Göttingen, Germany.

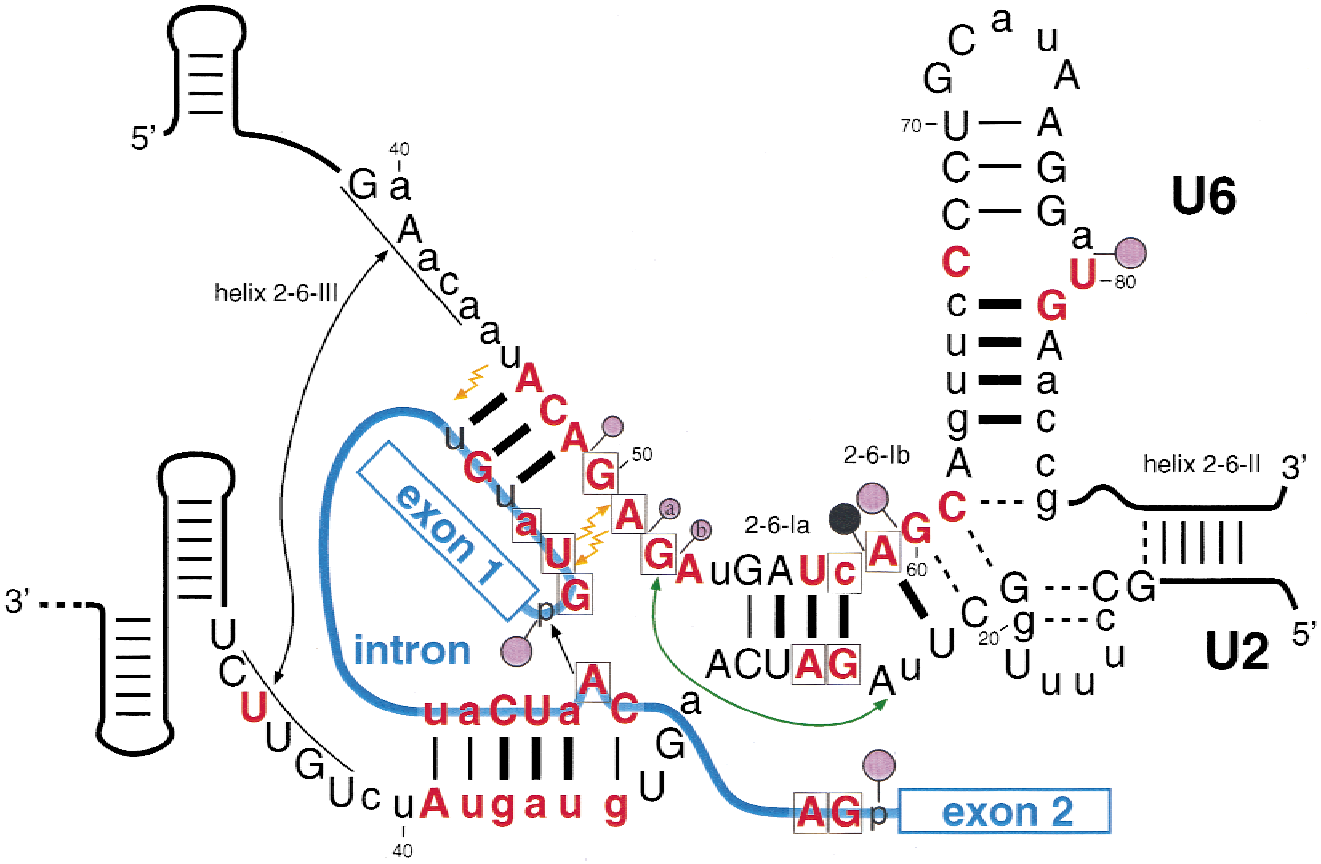

FIGURE 1. Model of the assembled spliceosome illustrating U2, U6, and pre-mRNA interactions (reviewed in Moore et al., 1993; Madhani & Guthrie, 1994a; Umen & Guthrie, 1995; Nilsen, 1998; Staley & Guthrie, 1998). Sequences and numbering refer to *Saccharomyces cerevisiae* U2 (1,171 nt) and U6 (107 nt) snRNAs. Phylogenetically conserved nucleotides (including *trans*-splicing trypanosomatidae sequences) are represented by uppercase letters. The intron–exon structure of the pre-mRNA is represented in blue. Letters in red indicate nucleotides where mutations cause a growth phenotype or a splicing defect in a biochemical reconstitution assay in yeast. Boxed nucleotides are important for step 2; mutants accumulate lariat intermediate in yeast. The pre-mRNA consensus sequences are derived from statistical analysis of yeast and human introns (Burge et al., 1999). A capital letter indicates a frequency >75%; yeast splice signals are however much more conserved than are human splice signals. Watson–Crick base pairs are symbolized by dashes. A bold dash indicates that the phenotype of a disruption can be partially or fully rescued by a compensatory change in yeast or human genetic systems, dashed lines represent base pairs that might fluctuate, and a double arrow indicates the suggested base-pairing interaction of helix 2-6 III. Spark arrows indicate crosslinks; the green double arrow depicts a tertiary base-pairing interaction, and the pins with the circle head mark critical R_P-phosphorothioate substitutions that affect splicing (large purple circle, complete block of splicing; small purple circle, partial block; black circle, block of second step only; letter 'a' in circle, effect only in yeast; letter 'b' in circle, effect only in nematode.

Luukkonen & Seraphin, 1998a, 1998b) and have been proposed to assist directly in catalysis of the splicing reaction. With inspiration from group II self-splicing introns, it was further proposed that the entire active site might be composed of RNA (Sharp, 1985; Cech, 1986; Guthrie & Patterson, 1988). However, group II introns are quite large; to date, even the smallest catalytically active derivatives are larger than 500 nt. Thus it has been unclear whether the few critical snRNA residues in the spliceosome could be sufficient for catalyzing the splicing transesterification reactions.

We have identified a 2′,5′-branch-forming ribozyme with intriguing similarities to pre-mRNA splicing and snRNA function. The ribozyme activity is dependent on a conserved U6-like sequence (ACAGAGA-box), and branch formation occurs with similar sequence specificity as in pre-mRNA splicing; an internal adenosine 2′-hydroxyl group attacks the α-phosphate of a 5′-terminal guanosine triphosphate, releasing diphosphate and forming a 2′,5′-branched lariat.

RESULTS

Selection of new lariat-forming ribozymes

Previously, we developed a combinatorial approach to explore the catalytic abilities of the U2 and U6 snRNA components in a protein-free environment (Tuschl et al., 1998). More than 10^{14} U2/U6 snRNA variants were subjected to an in vitro selection protocol to isolate

ribozymes that catalyze reactions similar to the second step of splicing, the exon ligation reaction. Although four different ribozyme activities were selected, the catalyzed reactions did not directly relate to the spliceosome-catalyzed reaction (Tuschl et al., 1998).

The most interesting ribozyme catalyzed the attack of an internal 2′-hydroxyl group at the α-phosphate of a 5′-triphosphate, and as with the spliceosome reaction, the reactive phosphate was 5′ to a G residue (Tuschl et al., 1998). However, the reaction differed from splicing in that the branch nucleotide was a guanosine, not an adenosine, and the leaving group was diphosphate, not an oligonucleotide. It was the intention to further analyze this ribozyme as a new example of a 2′,5′-branch-forming activity. A library of sequence variants of the 2′,5′-branch-forming ribozyme was synthesized at 20% partial mutagenesis (Tuschl et al., 1998). The parent construct contained an 18-nt segment based on human U6 sequence (positions 30–47 in human or positions 36–53 in yeast, Fig. 1) that included the critical ACAGAGA-box. The complexity of the starting library was 2×10^{14} different sequences, with 104 partially mutagenized positions (including the U6 segment) flanked by constant sequences at the 5′ and 3′ termini, used for primer binding.

Catalytically active sequences were isolated by iterative in vitro selection (Fig. 2). Linear starting material was separated from lariat product RNA by denaturing gel electrophoresis. Lariat-forming activity was first detected after three selection cycles, at which point about half of the pool 3 RNA reacted at an apparent rate of 0.002 min^{-1}. The selection was then continued under increasingly stringent conditions by adjusting the incubation time such that not more than 0.5% of lariat product was formed. After a total of seven rounds, the average activity approached 0.03 min^{-1}, and 73% of the input RNA formed lariats. PCR DNA from round 7 was cloned and sequenced (Fig. 3). Activities of 27 different isolates ranged between 0.004 and 0.5 min^{-1}, corresponding to a 40- to 5,000-fold rate enhancement relative to the parent sequence.

Comparative sequence analysis indicated that at least two different classes of lariat-forming RNAs were selected. These were identified by segments of invariant sequence surrounding characteristic changes from the parent sequence (Fig. 3). Eight isolates could not be categorized and may not be related in structure and function to the major classes or to each other. Surprisingly, neither of the two major classes represented the parent; an important helix of the parent ribozyme (pairing segments 71–76 and 81–86) was disrupted in both classes.

A conserved 15-nt region of one of the major classes (Fig. 3, red shading) corresponded to a highly conserved and biochemically important region of U6 snRNA (ACAGAGA-box) present at the heart of the spliceosome. The importance of this motif in this class of ribozymes and in U6 snRNA raised the question of whether the ribozymes catalyze a branch-forming reaction similar to that of the first step of pre-mRNA splicing. The isolate DL7.18, one of the more active representatives of this class of lariat-forming ribozymes, was chosen for further characterization.

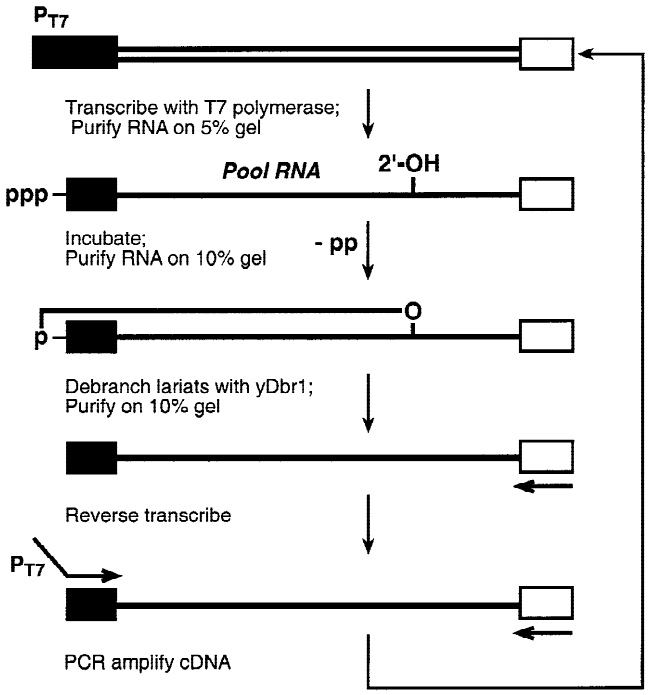

FIGURE 2. In vitro selection scheme for lariat-forming ribozymes. Abbreviations: P_{T7}: T7 RNA polymerase promoter; ppp: 5′-terminal nucleotide triphosphate; pp: diphosphate; OH: hydroxyl function; yDbr1: yeast debranching enzyme. Constant promoter and primer-binding sequences are depicted as boxes; arrows above or below those boxes symbolize the primer oligonucleotides. A single line represents single-stranded DNA or RNA, a double line represents double-stranded DNA.

As suggested by the comparative analysis, large segments of isolate DL7.18 could be deleted without disrupting branching activity. A 37-nt loop (positions 18–54) at the end of a putative stem was removed and replaced by a C(UUCG)G-tetraloop (Molinaro & Tinoco, 1995). Second, a 36-nt segment (positions 102–137) including the 3′ primer-binding sequence was removed from the 3′ end. Combining both changes resulted in a 70-nt ribozyme, DL7.18.d (Fig. 3), with activity (0.036 min^{-1}) comparable to that of full-length DL7.18 (0.064 min^{-1}).

The branch point of the deletion derivative, DL7.18.d, was mapped by partial alkaline hydrolysis and comparison of the hydrolysis ladder to a partial RNase T1 digest (Fig. 4). The branched product was generated by a *trans* reaction in which a 5′-^{32}P-labeled RNA containing the branch site (positions 55–107 from isolate DL7.18, Fig. 3) was reacted with unlabeled 5′-triphosphate-containing substrate (positions 1–17 from isolate DL7.18). RNA hydrolysis requires the presence of a

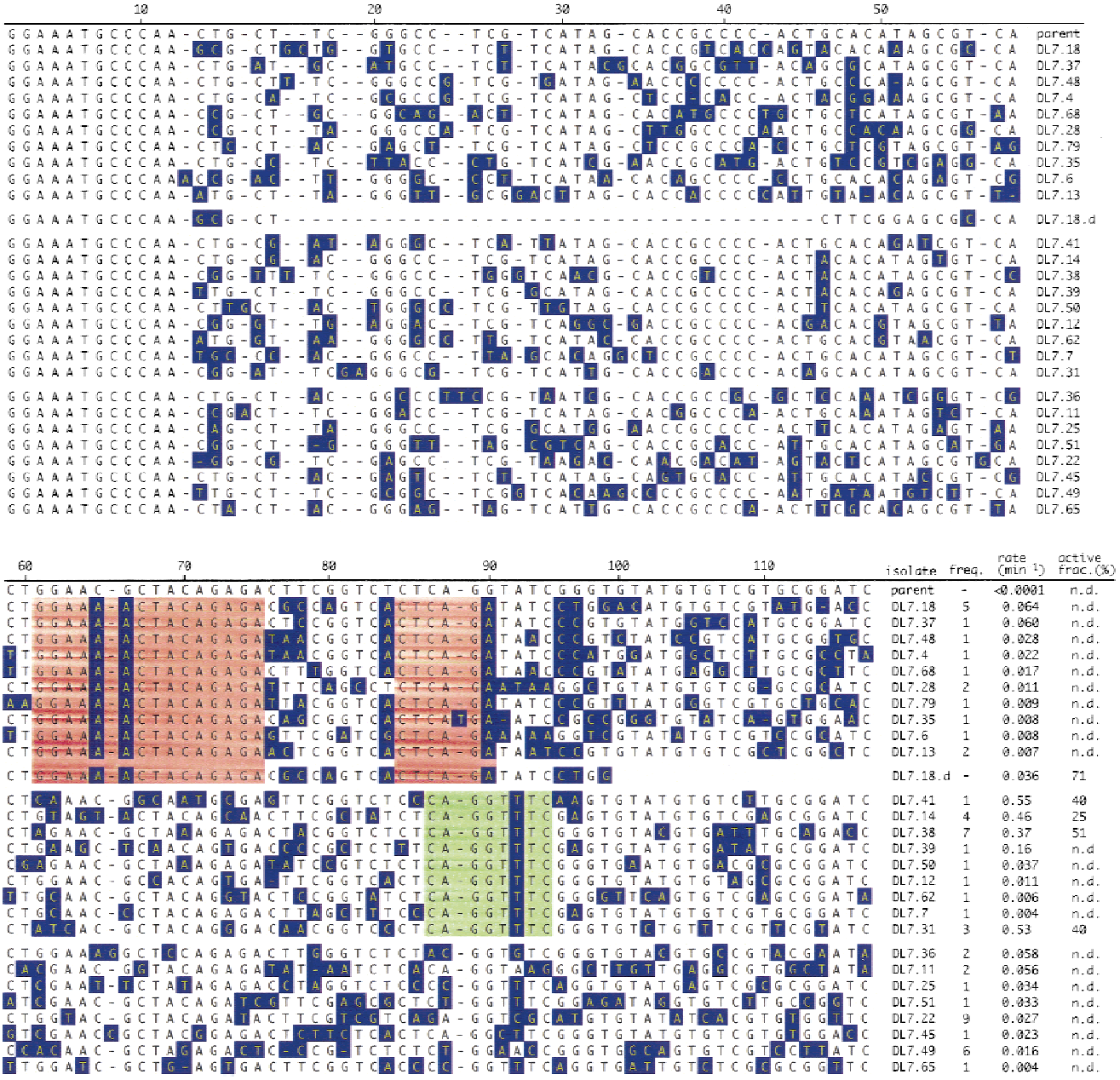

FIGURE 3. Comparative analysis of cDNAs of lariat-forming variants. The constant 5′ T7 promoter and 3′ primer-binding sequences are not shown. Residues that differ from the parent sequence are boxed in dark blue. Sequences that share a common motif (green and red boxes) and some specific nucleotide changes are grouped together and ranked by reaction rates. Eight sequences could not be classified. The sequence DL7.18.d is a deletion construct of DL7.18 in which the segment between positions 17 and 52 was replaced by a C(UUCG)G tetraloop. The frequency at which a clone was isolated is indicated (freq.), as is its rate constant for lariat formation and the fraction of molecules active; for slowly reacting isolates, the active fraction was not determined (n.d.) and assumed to be 70%.

free 2′-hydroxyl so that the last unshifted band in the lane of branched product corresponds to the nucleotide 5′ to the branch site. The branch-site hydroxyl was that of the adenosine within the conserved 4-nt segment CUCA. This position is 1 nt upstream relative to the branched guanosine identified in the parent sequence (Tuschl et al., 1998). The different branch-site residue can be considered additional evidence that a new class of lariat-forming ribozymes had been isolated.

Important sequence and structural features of the new ribozyme

It was intriguing that the ribozyme contained a conserved U6 segment and, like the spliceosome, formed a lariat involving an adenosine 2′-hydroxyl and a guanosine 5′-phosphate. To further test the sequence requirements of this ribozyme, an in vitro selection protocol was used to generate a large number of active ribo-

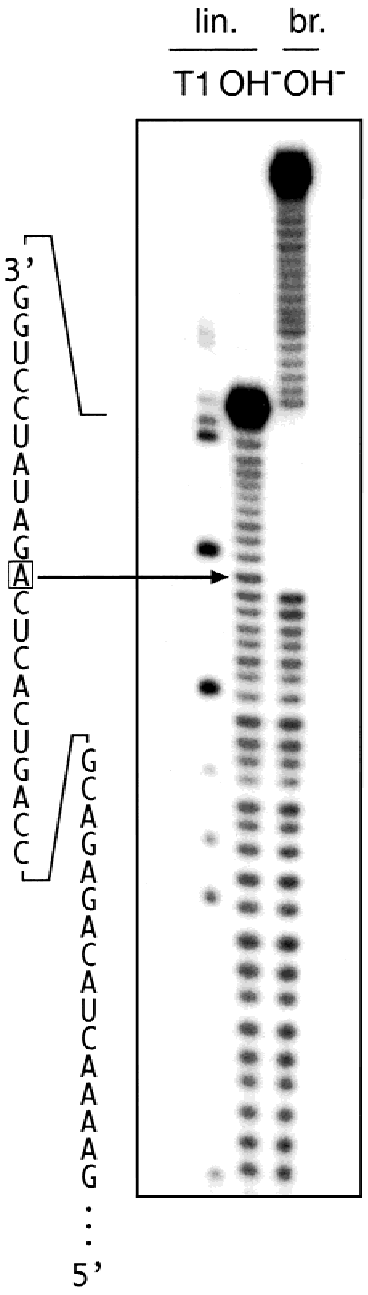

FIGURE 4. Mapping of the branch site. Partial alkaline hydrolysis (OH^-) was performed on linear (lin.) and 2′,5′-branched (br.) substrates. A reference ladder was generated by partial RNase T1 digestion, which cleaves 3′ to guanosines leaving a 3′-phosphate and a free 5′-hydroxyl. Alkaline hydrolysis requires the presence of a free 2′-hydroxyl and generates 2′,3′-cyclic phosphates. The first missing band in the hydrolysis ladder of the branched substrate corresponds to that of the branch-point nucleotide (arrow).

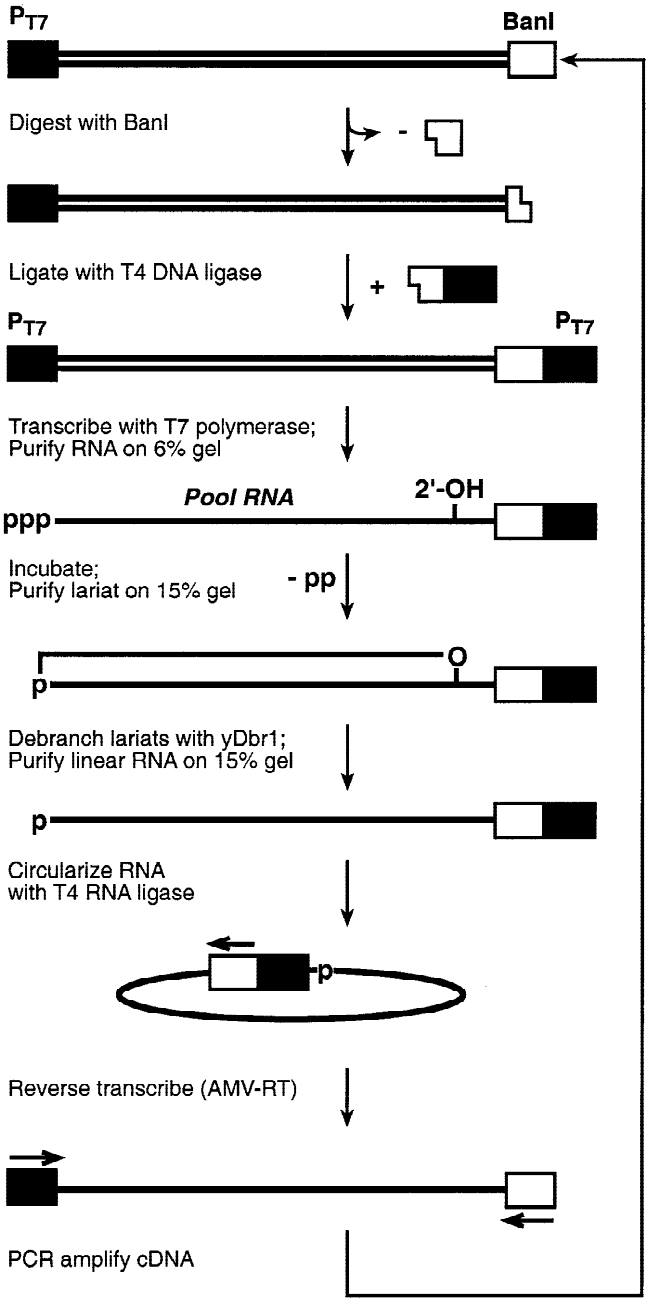

FIGURE 5. In vitro selection scheme for lariat-forming ribozymes with unconstrained 5′ sequence. To generate amplifiable cDNA with constant primer sequences at both ends, the pool RNA that had been selected and debranched was circularized, then reverse transcribed using a primer that hybridized to only the 5′ half of the constant sequence (open box). The remainder of the constant sequence (filled box), which corresponded to the T7 promoter sequence, was thus the last RNA to be reverse transcribed. Following PCR, DNA could be transcribed using this repositioned T7 promoter. The 3′-terminal T7 primer/promoter sequence was added by digestion of the PCR DNA with *Ban*I restriction endonuclease, followed by DNA ligation. Abbreviations: p: terminal 5′-phosphate; for other symbols, see Figure 2.

zyme variants. This protocol was designed to permit variation of every ribozyme residue—even the 5′-terminal residues, which cannot be varied in a typical selection protocol (Fig. 5). Ideally, the conservation of a particular residue among selected variants will reflect the importance of this residue for branch-forming activity. It should be noted, however, that despite initial randomization throughout the ribozyme, sequence preferences of protein enzymes required during amplification steps might impose additional sequence bias at

certain residues. For example, T7 RNA polymerase initiates transcription preferentially with templates encoding a guanosine and tends to abort transcription when attempting early incorporation of uridines (Milligan & Uhlenbeck, 1989). The substrate preference of the debranching enzyme Dbr1, used in our protocol to linearize selected ribozymes, might also favor ribozymes that branch using the 5′-triphosphate of a purine rather than that of a pyrimidine nucleoside (Nam et al., 1994).

After four and five rounds of selection, PCR DNA was cloned and individual isolates were sequenced and assayed for their activities (Fig. 6). Sequences of all lariat-forming ribozymes were aligned to the parent ribozyme. The majority of isolates (45/50) related to the parent structure by virtue of their invariant U6-like motif and branch-site sequence (Fig. 6A). Their reaction rates ranged between 0.12 to 0.005 min^{-1}. The most active isolate was three times faster than the starting sequence. Five isolates could not be classified (Fig. 6B). Their rates ranged between 0.04 and 0.02 min^{-1}. They have four to eight changes in the residues conserved in the main class, and two of these isolates do not appear able to form a 3′-terminal stem present in the main class, supporting the idea that these isolates are not related to the main class. They were not examined further.

The 45 isolates of the major class contained 26 conserved residues, which were concentrated in three regions. The first is a 4-nt segment, pppGGAA, at the 5′-triphosphate-containing terminus. About half of the active sequences showed a single-nucleotide insertion between the T7 promoter and the conserved GGAA 5′ end of the template DNA (Fig. 6, brown shading). Such insertions were an anticipated artifact of the selection protocol, and the additional nucleotide was most likely not transcribed in catalytically active molecules (Fig. 6 legend). The second conserved segment, AACTACA GAGA, includes the ACAGAGA U6-like sequence. The third invariant segment, ACUCAGA, includes the branch-site adenosine (underlined). This motif was also present in 7 of the 10 founding members of this class (Fig. 3); the three nonmatching sequences from the first selection differed by a change of the first A or by a U insertion after the branch-site A. The 7-nt branch-site motif is flanked by an additional four interspersed conserved residues.

Sequence analysis strongly supports a 4-bp stem (Fig. 6A, violet shading) surrounding the branch-point adenosine; numerous covariations were found in the third and fourth loop-closing base pairs. The adjacent second base pair (positions 52 and 68) was generally conserved, although one U/G covariation and six mismatches were found. The loop-closing base pair (positions 53 and 67) was always a G/C pair and flanked by a conserved U at position 54. We also identified an unusual covariation (Fig. 6A, green shading) between positions 5 and 35 (23 A:A, 15 G:C, 5 A:G, 1 A:T, and 1 C:G pair). Position 35 is immediately upstream of the conserved U6-like segment. An interaction between residues 5 and 35 might position the reactive 5′ end of the ribozyme near the conserved U6-like sequence.

Sequence comparison indicates a 28-nt variable region flanked by the unusual A5/A35 interaction. Potentially base-paired regions are located in this segment, but were not conserved in strength or position. This suggested that deletions might be active as long as the covarying A/A pair was not perturbed. Two isolates were examined (Fig. 7A). When the 17-nt loop region at the end of a potential 3-bp stem of isolate 5.33 was replaced by a stable UUCG tetra-loop (5.33.d3, Fig. 7B) the activity was unaffected. As expected, extension of the stem by one more base pair (5.33.d4) had no effect. Similarly, reduction of the internal stem-loop region of isolate 5.64 to a stable 3-bp plus a tetra-loop element (5.64.d3) did not affect the rate of lariat formation. Deletion of the 3′ half of the pairing segment (5.33.d1) caused a more than 40-fold loss in activity and resulted in a branched dimer product formation without detectable intramolecular lariat formation. Dimerization is explained by the fortuitous ability to form an intermolecular 6-bp stem after the internal deletion. In summary, the fully active deletion variants 5.33.d3 and 5.64.d3 were only 58 and 55 nt long, respectively. The 5.33.d3 variant was chosen as the prototype ribozyme that forms an RNA branch using the 2′-hydroxyl of an A and the 5′-phosphate of a G. It was renamed the 2′-5′AG1 lariat-forming ribozyme (Fig. 7B) and characterized further.

FIGURE 6 (*facing page*)**.** Comparative analysis of cDNAs of selected lariat-forming variants. **A**: Sequences with the conserved U6-like motif. **B**: Sequences without U6-like motif. Conserved regions are indicated (red bars on top of the alignment). The branch site (A62) is indicated with an arrow. An unusual covariation between positions 5 and 35 is highlighted in green. A paired region is highlighted in violet. Bases that covary are marked with dark blue vertical bars and G/U wobble pairs that maintain the predicted pairing are shaded orange. Nucleotides that differ from the parent sequence (DL7.18.d) and do not maintain pairing are boxed in dark blue. If such a change creates a mismatch within the paired segment, then the nonmutated partner is boxed in light blue. Single-nucleotide insertions between the T7 promoter and the conserved GGAA 5′ end are shaded in brown. These are an artifact of the selection-amplification protocol, generated as a consequence of an untemplated nucleotide added to the 3′ terminus of many T7 polymerase run-off transcripts (Milligan & Uhlenbeck, 1989). The inserted nucleotide is absent in the catalytically active RNA fraction (data not shown). Such inaccuracies of T7 RNA polymerase near the 5′ end have been observed previously (Helm et al., 1999). All sequences were represented by a single isolate, except sequences 4.9, 5.22, and 5.47, which were isolated twice.

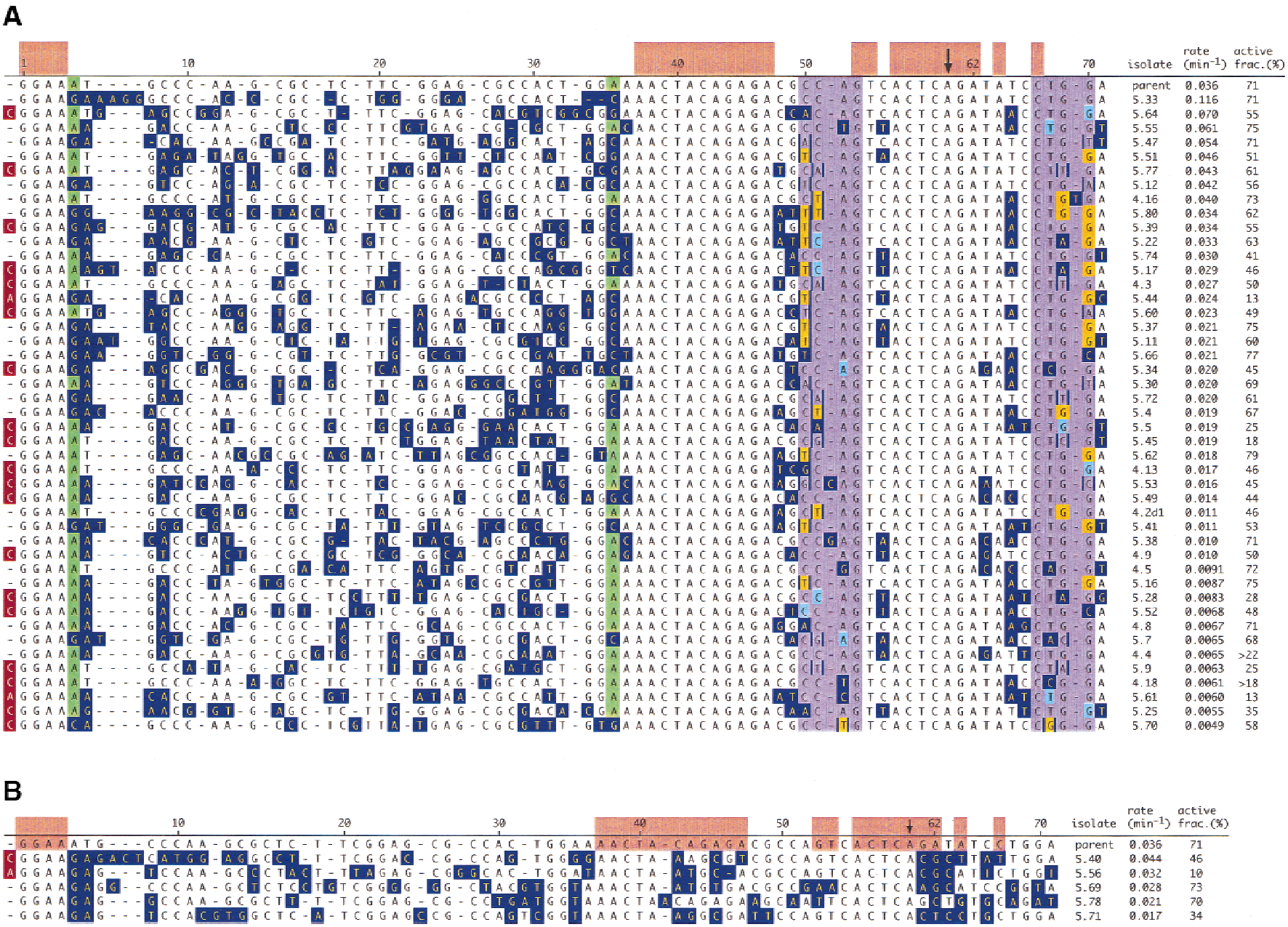

FIGURE 6. *See caption on facing page.*

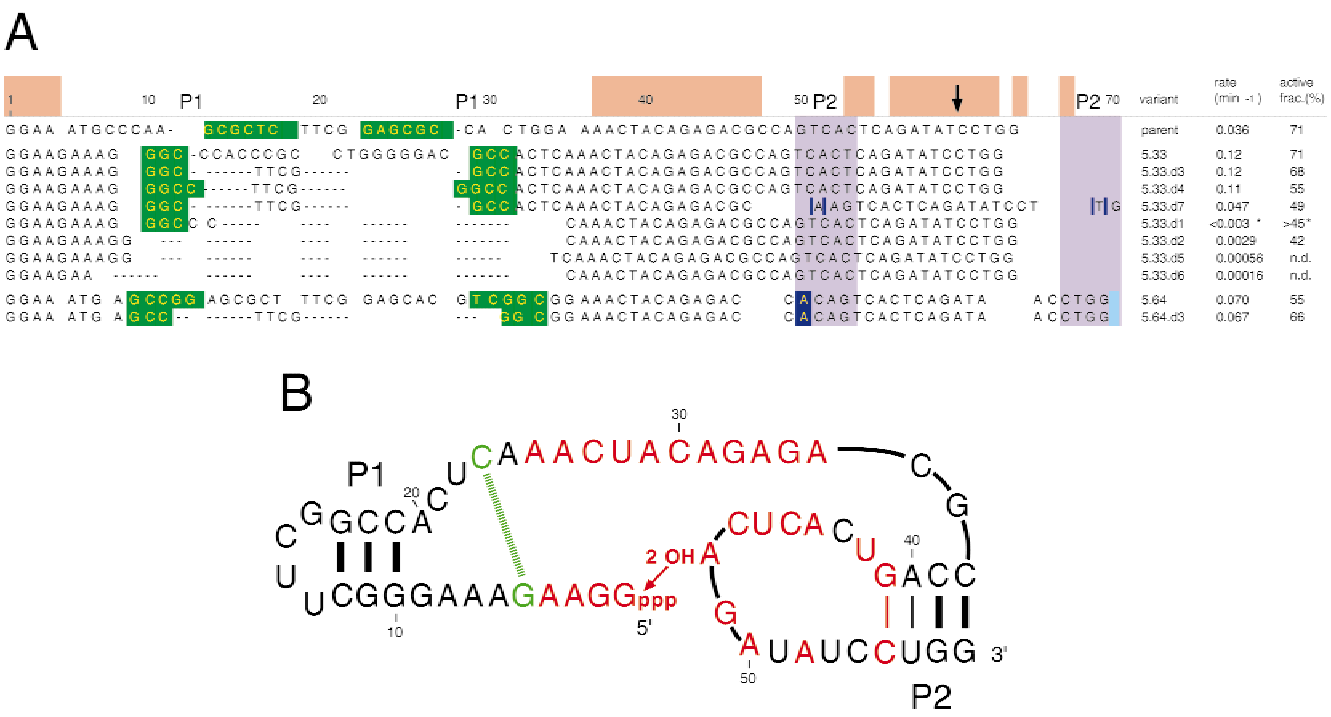

FIGURE 7. Truncation analysis of two ribozyme variants. **A**: The truncated sequences are aligned relative to the parent sequence (DL7.18.d) from the second selection. Red bars indicate conserved regions, green boxes highlight putative paired segments (P1) residing within the variable region, and the violet shaded areas mark the conserved paired region (P2). A change in a P2 base pair is indicated for construct 5.33.d7 and was tested because the A/U base-pair transversion was frequently found in selected variants. A terminal A/G mismatch in P2 is also indicated for 5.64 variants. The asterisks for variant 5.33.d1 indicate formation of a dimer rather than a lariat. **B**: 2′-5′AG1 lariat-forming ribozyme. Construct 5.33.d3 of **A** was renamed 2′-5′AG1 when it was chosen as the prototype lariat-forming ribozyme. Invariant residues are indicated in red, and the unusually covarying residues are in green. P1 and P2 denote paired regions; base pairs are symbolized by vertical lines, which are bold when covariation or additional experiments support base pairing.

Importance of the ACAGAGA segment, branch-site adenosine, and 5′ guanosine

The ACAGAGA-box is a phylogenetically conserved and functionally important region of U6 snRNA (Fig. 1). For comparison to the spliceosome, the importance of each nucleotide in the ACAGAGA-motif of the lariat-forming ribozyme (Fig. 7B, A29–A35) was tested by assaying the activity of each of the 21 single-nucleotide substitutions within this segment (Table 1). As expected, all mutants were less active than the unmodified 2′-5′AG1. The most critical positions for mutations within the ACAGAGA segment were A29, G32, and A33, where each single mutation reduced the reaction rate at least 400-fold. At the remaining positions, only the G34C and A35C mutations reduced the rate more than 400-fold, whereas other mutations were less consequential.

To test for the specificity of branch formation, the branch-site adenosine, A48 in 2′-5′AG1, was mutated. Mutation of this position to any other nucleotide diminished lariat formation to an undetectable level, corresponding to more than a 20,000-fold rate reduction. The activation or use of a cryptic branch site was not detected. Therefore, the ribozyme is specific for an adenosine branch site.

Specificity of the recognition of the 5′-terminal guanosine in ribozyme 2′-5′AG1 was tested by replacing this guanosine with adenosine. Branch formation was assayed by TLC analysis of diphosphate release from γ-^{32}P-ATP-labeled transcripts (Tuschl et al., 1998). No diphosphate was detected over a 45-h incubation period, in contrast to the rapid diphosphate release seen with γ-labeled 2′-5′AG1. Although C and U substitutions were not tested, the marked discrimination against an A substitution indicates that the 5′ G residue is specifically recognized during branch formation.

Thiophosphate substitutions

Thiophosphate interference studies have identified three sites in the ACAGAGA-box of U6 snRNA where substituting a pro-R_P oxygen with sulfur can partially inhibit splicing (Fig. 1). For the 2′-5′AG1 ribozyme, no thiophosphate interference has been detected for pro-R_P substitutions within the ACAGAGA segment, although at 13 sites outside this segment, substitution interfered with branching activity (5′-phosphates of G1-A3, A7-G9, C21, U22, A24, A25, C47-G49, our unpubl. results). A strict correlation would not necessarily be expected because substitution within U6 only partially inhibits

TABLE 1. Effects of point mutations of the ribozyme and spliceosome ACAGAGA motif.

Mutant	Rel. rate	Active fraction	Splicing in cells		Splicing in vitro	
			Yeast[a]	Human[b]	Yeast[c]	Human[d]
Unmodified	1.0	0.65	v	splicing	++	++
A29C	<0.00005	n.d.	l	splicing	+	++
A29G	0.00017	n.d.	ts	splicing	++	++
A29U	0.00005	n.d.	v (cs)	inhibition	++	++
C30A	0.080	n.d.	l	inhibition	−	+
C30G	0.0066	n.d.	l	splicing	+/−	++
C30U	0.0036	n.d.	l	splicing	+/−	n.d.
A31C	0.26	0.46	l	inhibition	+	++
A31G	0.012	n.d.	l	splicing	+	n.d.
A31U	0.0025	n.d.	l	inhibition	++	++
G32A	0.00010	n.d.	cs, ts (l)	splicing	+	+
G32C	0.00007	n.d.	v (ts)	inhibition	+	+
G32U	0.00018	n.d.	ts (l)	inhibition	+	++
A33C	0.00012	n.d.	l	inhibition	2−	2−
A33G	0.0024	n.d.	l	inhibition	2−	++
A33U	0.00007	n.d.	l	splicing	2−	2−
G34A	0.0032	n.d.	l	splicing	2+	+
G34C	0.00012	n.d.	l	splicing	2+	++
G34U	0.048	n.d.	ts	inhibition	++	++
A35C	0.0014	n.d.	v	splicing	++	++
A35G	0.17	0.22	v	splicing	++	+/−
A35U	0.041	0.15	n.d. (cs, ts)	splicing	++	++

The effects of mutations within the ACAGAGA segment of the lariat-forming ribozyme compared to published data of U6 snRNA mutants in vitro or within yeast or human cells. The numbering of point mutants refers to the prototype construct 2′-5′AG1 (Fig. 7B); the corresponding U6 snRNA ACAGAGA-box is segment A47-A53 in yeast (Fig. 1) or segment A41-A47 in human. Ribozyme rates are reported relative to the unmodified rate (0.12 min^{-1}). If a plateau could not be reached during the incubation interval, an active fraction of 0.65 was assumed and used for fitting the rate constant. Abbreviations: n.d.: not determined; v: viable; cs: cold-sensitive; ts: temperature-sensitive; l: lethal; ++: 50–100% splicing; +: 10–50% splicing; +/−: <10% splicing; −: no splicing; 2+: partial block of the second step of splicing; 2−: complete block of the second step.

[a] In vivo phenotype of yeast U6 snRNA mutants according to Madhani et al. (1990) and McPheeters (1996). Data given in parenthesis indicate differences of the second study from the first.

[b] Splicing complementation system in human cells (Datta & Weiner, 1993).

[c] In vitro splicing with reconstituted yeast U6 snRNP (Fabrizio & Abelson, 1990).

[d] In vitro splicing with reconstituted human U6 snRNP (Wolff et al., 1994).

splicing, and only one of the three sites is the same in the two systems studied, yeast and nematode.

Replacement of the reacting terminal 5′-triphosphate of 2′-5′AG1 by either the R_P or S_P α-thiotriphosphate reduced the branching reaction rate at least 500-fold (our unpubl. results). These thio effects were at least eight times greater than the 60-fold elemental effect expected solely from the decreased chemical reactivity of the triphosphate upon sulfur substitution at either of these nonbridging oxygens (Herschlag et al., 1991; Wong et al., 1991). Thus the magnitude of the thio effects suggests that the nonbridging oxygens at the reaction center each make important contacts to either a metal ion or ribozyme residue. No 2′-deoxy interference was observed within the ACAGAGA segment following incorporation of dA, dC, or dG thiophosphate (our unpubl. results).

DISCUSSION

A lariat-forming catalytic RNA has been isolated that promotes 2′,5′-branch formation between adenosine and guanosine, a reaction similar to that of the spliceosome and group II self-splicing introns. The prototype sequence (2′-5′AG1) is many times smaller than a group II intron (58 nt compared to >500 nt), the only other characterized ribozyme that promotes this reaction. The finding that an RNA of this size can promote branch formation supports the idea that the snRNA available in the spliceosome might be sufficient for catalyzing lariat formation. Indeed, the 2′-5′AG1 ribozyme contains a conserved and functionally important region of U6 snRNA. With this motif in common with U6 snRNA, the small 2′,5′-branch-forming ribozyme appears to share more sequence similarity with the spliceosomal

RNAs than does the much larger group II intron. However, it is important to note that the sequence similarities between the 2′-5′AG1 ribozyme and the snRNA are limited and might simply reflect fortuitous rather than functional similarities. The extent to which the conserved motifs of U6 snRNA and the isolated ribozyme might have similar functions is discussed below.

Function of the ACAGAGA segment in the 2′-5′AG1 ribozyme and the spliceosome

Site-specific mutagenesis of the seven positions within the ACAGAGA segment of the 2′-5′AG1 ribozyme resulted in dramatic decreases in activity (Table 1), showing that these residues are critical for 2′-5′AG1 ribozyme folding, catalysis, or both. In U6 snRNA, the ACAGAGA-motif plays at least two roles, one in spliceosome assembly and one associated with catalysis of the two splicing phosphodiester transfer reactions (Fig. 1).

During spliceosome assembly, the ACA-part of the U6 snRNA sequence base pairs, at least transiently, with positions 4, 5, and 6 of the intron (Kandels-Lewis & Seraphin, 1993; Lesser & Guthrie, 1993; Hwang & Cohen, 1996). This interaction is conserved in the evolutionary distant minor U12-type spliceosome where the GGA-part of the U6atac snRNA GGAGAGA-motif base pairs with complementary residues of the U12-type 5′ splice site (Incorvaia & Padgett, 1998). The interaction is less pronounced in the mammalian U2-type spliceosome because the mammalian 5′ splice-site sequences of the U2-type introns are much more degenerate. Sequence analysis of the 2′-5′AG1 ribozyme indicates that the ACA of the ribozyme ACAGAGA segment is not involved in a canonical pairing interaction to present the reactive 5′ end at the active site. However, the possibility that these residues utilize non-Watson–Crick interactions to position the reactive 5′ terminus has not been ruled out.

The AGAGA part of the U6 ACAGAGA-motif is conserved between the U2- and U12-type spliceosomes and its function appears more complex, as different mutant phenotypes are observed in different experimental systems (reviewed in Luukkonen & Seraphin, 1998a). Crosslinking studies (Sontheimer & Steitz, 1993; Kim & Abelson, 1996) and genetic studies in yeast (Luukkonen & Seraphin, 1998a) suggest that the center GAG-part is involved in non-Watson–Crick recognition of the 5′-most GU of the intron 5′ splice site. Genetic studies indicate that the AGAGA residues are also critical for catalysis of the first step of splicing (Li & Brow, 1996; Luukkonen & Seraphin, 1998a) and for catalysis of the second step, or alternatively, a proofreading activity prior to the second step (Fabrizio & Abelson, 1990; Madhani et al., 1990; Datta & Weiner, 1993; Kandels-Lewis & Seraphin, 1993; Lesser & Guthrie, 1993; Wolff et al., 1994). An RNA-based active site in the spliceosome is supported by the observation that U2-U6 helix Ib and a tertiary contact between U6-G52 and the bulged U2-A25 between helix Ia and Ib contribute to catalysis (Fig. 1) (Madhani & Guthrie, 1994b; Luukkonen & Seraphin, 1998b). As in the spliceosome, the 2′-5′AG1 ACAGAGA segment is flanked by paired regions (P1 and P2; Fig. 7B), which resemble the spliceosomal U2/U6 helices Ia and III (Fig. 1). Helix Ia is 4 or 5 bp long and required for splicing independent of its sequence composition (Madhani & Guthrie, 1992). Helix III is less critical and contributes in human (Sun & Manley, 1995) but not in yeast splicing (Yan & Ares, 1996).

ACAGAGA sequences of other natural ribozymes

An ACAGAGA segment is also found in particular versions of the hairpin and RNase P ribozymes, but it is not as highly conserved in these other catalytic RNAs as it is in the 2′-5′AG1 ribozyme or U6 snRNA. In the case of the sTobRV hairpin ribozyme, the first 4 nt of the segment participate in canonical base pairing to form helix I, and the identities of only the last 2 nt are critical for catalysis (Burke et al., 1996). Three residues of the ACAGAGA sequence of *Escherichia coli* RNase P RNA are not conserved in bacteria and archaea (Haas et al., 1996; Haas & Brown, 1998), and RNase P deletion variants that lack this region retain catalytic activity (Pan & Jakacka, 1996). Group II introns do not include an ACAGAGA or AGAGA segment. Instead, analogy is made between domain V of group II introns, a bulged helix with a conserved AGC trinucleotide motif, and a domain of the spliceosome that comprises the intramolecular U6 helix and adjacent the U2/U6 helix Ib, which also contains a conserved AGC segment (Michel & Ferat, 1995; Pyle, 1996; Qin & Pyle, 1998).

Combinatorial approaches to understanding complex biological enzymes

The frequency of new lariat-forming ribozymes in RNA sequence space is unknown, but an estimate based on the number of conserved positions in the 2′-5′AG1 ribozyme suggests that this catalyst should be present in typical libraries used to isolate new ribozymes from totally random sequences (Sabeti et al., 1997). Several lariat-forming ribozymes that did not conserve the ACAGAGA segment were also isolated during the course of our experiments (Fig. 3). These ribozymes have not been characterized and are probably structurally unrelated to the family of ribozymes exemplified by 2′-5′AG1. The parent ribozyme used for the library construction is probably also unrelated to the 2′-5′AG1 ribozyme as it uses a G branch site and has a 5-bp helix located 4 nt upstream of the branch site that is absent from the 2′-5′AG1 ribozyme (Tuschl et al., 1998).

Basing an RNA library on sequences presumed to be catalytically important in ribonucleoprotein complexes might be better than randomly sampling sequence space when seeking to access an ancestral-like RNA-based activity. Here, the approach resulted in a focus on a ribozyme with a sequence relationship to U6 snRNA. However, it is interesting that a peptidyl-transferase ribozyme isolated from a totally random sequence library resembles the peptidyl-transfer region of 23S ribosomal RNA in sequence and structural context (Zhang & Cech, 1998). Similar observations were made of sequence similarity with 23S rRNA in an RNA aptamer selected from random sequence to bind to CCdApPuromycin, a high-affinity ligand to the peptidyl-transferase center of the ribosome (Welch et al., 1997). The importance of individual residues within the aptamer/ribozyme motifs has not yet been tested by site-directed mutagenesis or by rerandomization and selection. Nevertheless, these examples and the lariat-forming ribozyme illustrate the potential of combinatorial methods to create model systems with functional and possibly structural similarities to complex ribonucleoproteins.

Differences between the 2′-5′AG1 ribozyme and the spliceosome

The lariat-forming reaction of the 2′-5′AG1 ribozyme is different from nuclear splicing in that the spliceosome catalyzes a transesterification reaction, whereas the ribozyme catalyzes an esterification, that is, the leaving group in the ribozyme reaction is diphosphate rather than a 5′ exon with a free 3′-hydroxyl. This difference in chemistry could suggest a fundamentally different active site in the small ribozyme, despite the necessary U6-like segment. However, it is interesting to note that group II introns can also catalyze an esterification reaction with diphosphate release (Mörl et al., 1992). The catalytic center for the first step of splicing is thought to execute this reaction (Mörl et al., 1992). Diphosphate release has not yet been observed in a nuclear splicing reaction, and the smallest 5′ exon leaving group yet observed is a 1-nt 5′ exon (Hertel & Maniatis, 1999). The thermodynamic driving force for branch formation via diphosphate release is small and must be similar to the free energy change observed when a nucleotide is added during RNA polymerization, which is within a range of −1.0 to −3.0 kcal/mol (reviewed in Kahn & Hearst, 1989).

There are probably several structural differences between the 2′-5′AG1 ribozyme and RNA interactions in the spliceosome. For example, the branch-site adenosine appears to be within a terminal loop in the 2′-5′AG1 ribozyme, whereas it is typically a bulged nucleotide in the splicing complex. Another difference between the 2′-5′AG1 ribozyme and the spliceosome is the strict requirement for an adenosine at the branch site and a guanosine at the 5′ terminus. In contrast to the 2′-5′AG1 ribozyme, mutations at the branch site of introns only reduce the efficiency of the first step of splicing but do not block it completely (McPheeters, 1996; Query et al., 1996). At the 5′ splice site, the 5′-most intron nucleotide is almost always a guanosine, but it can also be an adenosine, as seen in a subclass of U2-type and U12-type introns (reviewed in Sharp & Burge, 1997). Although G to A substitution is tolerated in the spliceosomes, it is not in the 2′-5′AG1 ribozyme. The differences in specificity could indicate that the RNA network of the spliceosome and the 2′-5′AG1 ribozyme are different, but it is also possible that branch site, splice site, or snRNA mutations are suppressed by protein components of the spliceosome that stabilize the catalytic site. In addition, more extreme preferences within the spliceosome active site would be obscured if chemistry were not the rate-limiting step in splicing.

CONCLUSIONS

The conserved sequences in the ribozyme and the U6 snRNA could have a common function but the results summarized above are far from being conclusive. Recent evidence indicates that the protein Prp8 is also closely associated in the spliceosome with the 5′ splice site, branch site, and 3′ splice site (Collins & Guthrie, 1999; Luo et al., 1999; Reyes et al., 1999; Siatecka et al., 1999). It is possible that this or another protein could be a component of the active site for the first step in splicing, perhaps in conjunction with U2 and U6 snRNA. These possibilities are difficult to investigate because the conformations of snRNAs change during the splicing process and the heart of the spliceosome is relatively inaccessible to probes. In the near term, it will be important to determine the structure of the 2′-5′AG1 ribozyme to gain insight into the specific function of the ACAGAGA sequence in lariat formation. It may then be possible to test if the analogous residues in U6 snRNA have a related function in the spliceosome.

MATERIALS AND METHODS

Pool constructions

RNA pool for first selection

A 20% degenerate oligodeoxynucleotide library of the original lariat-forming ribozyme was synthesized as previously described (Tuschl et al., 1998). The degenerate oligodeoxynucleotide, GGAAATGCCCAACTGCTTCGGGCCTCGTCATAGCACCGCCCCACTGCACATAGCGTCACTGGAACGCTACAGAGACTTCGGTCTCTCAGGTATCGGGTGTATGTGTCGTGCGGATCCTTATTTGATGTCATCCGA, contained the 114-nt ribozyme sequence (Tuschl et al., 1998) followed by a 2-nt spacer and a 19-nt 3′ primer-binding sequence (primer-binding sequences underlined). The degenerate oligonucle-

otide was amplified by large-scale PCR (Ekland & Bartel, 1995) using a 5′ primer that appended the T7 RNA polymerase promoter sequence (TAC**TAATACGACTCACTATA** GGAAATGCCCAA; bold, T7 promoter; underline, segment hybridizing to pool second-strand product). The PCR product was transcribed with T7 RNA polymerase to yield the degenerate RNA pool. The RNA was gel purified on a 5% denaturing gel, which does not discriminate between linear and lariat forms of the parent ribozyme, and thus lariats that form during transcription were not expected to be lost. Traces of copurified template DNA were hydrolyzed with RQ1 DNase (Promega).

RNA pool for second selection

An oligodeoxynucleotide library was synthesized with the entire DL7.18.d sequence (Fig. 3) mutagenized to 20% degeneracy, TAC**TAATACGACTCACTATA**GGAAATGCCCAAG CGCTCUUCGGAGCGCCACTGGAAAACTACAGAGACGCC AGTCACTCAGATATCCTGGAC*GGCACC*TTATGCGTACTT (bold, T7 promoter; underline, heterologous 3′-primer-binding sequence; italic, *Ban*I restriction site). After large-scale PCR amplification, the product was digested with the restriction endonuclease *Ban*I and ligated (3-mL reaction, 1 μM each duplex, 60,000 U T4 DNA ligase from NEB, 16 h, 22 °C) to a pre-annealed DNA duplex with an additional T7 promoter segment (top strand, *pGCACC*TTATGCGTACTTTAC**TAATA CGACTCACTATA**; bottom strand, **TATAGTGAGTCGTATT** AGTAAAGTACGCATAA*G*). Half of the ligation mix was transcribed with T7 RNA polymerase (2-mL reaction, 20 μCi α-^{32}P-UTP) to yield the degenerate RNA pool. The RNA pool was gel purified on a 6% denaturing gel, which does not discriminate between linear and lariat forms of DL7.18.d.

First selection and amplification

Lariat-forming incubation

Pool RNA (1 μM) and 3′ primer (2 μM) were incubated at 30 °C in ribozyme buffer (100 mM KCl, 25 mM $MgCl_2$, 0.25 mM EDTA, 30 mM Tris, pH 7.6); the 3′ primer was added to prevent the 3′ constant region from interacting with the ribozyme. The initial round of the selection used 0.35 mg of pool RNA and was incubated for 22 h. In subsequent rounds, the amount of input RNA was reduced 10-fold. Once lariat formation was apparent, the incubation period was shortened to allow only 0.1% product formation. The incubation was stopped with EDTA and then ethanol precipitated.

Gel selection

RNA from the incubation steps was run on 15-cm 10% denaturing acrylamide gels to separate unreacted linear RNA from lariat product. The region of the gel corresponding to the lariat form of the parent ribozyme was excised and soaked overnight in 0.3 M NaCl and 1 μM 3′ primer as carrier. Eluted RNA was ethanol precipitated and dissolved in water. At round 3, an additional gel-purification step was introduced, in which samples were repurified on 40-cm 10% sequencing gels.

Lariat debranching

Debranching was performed using Dbr1 (a generous gift of Jef Boeke; Chapman & Boeke, 1991; Nam et al., 1994), the yeast debranching enzyme (40-μL reaction, $<$20 nM lariat, $<$1 μM 3′ primer, 20 mM HEPES/KOH, pH 7.6, 125 mM KCl, 0.5 mM $MgCl_2$, 1 mM DTT, 10% (v/v) glycerol, 20 U enzyme, 30 °C, 16 h). Before adding buffer and debranching enzyme, the RNA solution was heat-treated for 1 min at 90 °C. For the first two rounds, the debranching reaction was stopped by phenol and chloroform extraction, followed by ethanol precipitation. In subsequent rounds, the reaction was stopped by the addition of an equal volume of 8 M urea/50 mM EDTA and directly loaded onto a 15-cm 10% denaturing gel. The linear product band was excised, soaked overnight in 0.3 M NaCl and 1 μM 3′ primer as carrier, and ethanol precipitated. The gel-purification step after debranching is useful for preventing enrichment of noncatalytic RNAs with unusual mobility in denaturing gels (data not shown).

Amplification

Following reverse transcription (20-μL reaction, 5 μM 3′ primer, 200 U Superscript II from Gibco, 48 °C, 30 min), RNA was hydrolyzed by adding 40 μL of 150 mM KOH/20 mM Tris base and incubating for 10 min at 90 °C. The pH was then adjusted to 8.5 with HCl. The cDNA was amplified by PCR with a hot start (80 °C) (Tuschl et al., 1998). The PCR product was ethanol precipitated and directly used for transcription. Transcription reactions were typically 200 μL, using the template from a 400-μL PCR aliquot and 20 μCi of α-^{32}P-UTP for body labeling. RNA for the next round of selection-amplification was gel purified on a 5% denaturing gel. After seven rounds of selection-amplification, cDNA was cloned (T-Vector kit, Novagen) and individual clones were tested and sequenced.

Second selection and amplification

Starting pool (0.35 mg) was incubated for 14 h under conditions identical to the first selection. After precipitation, lariat product was separated from linear starting material on a 15% denaturing gel and repurified on a 12% denaturing gel. After debranching, linear RNA was directly gel purified on a 12% denaturing gel. The linear product band was excised, the RNA eluted in 0.3 M NaCl in the absence of carrier DNA, and ethanol precipitated. Debranching generates 5′ phosphorylated linear RNA that was circularized using T4 RNA ligase (Pan & Uhlenbeck, 1992). The circularization reaction (20-μL reaction, 12 U T4 RNA ligase from Pharmacia Biotech, 50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, 0.2 mM ATP, 0.1 mg/mL acetylated BSA, 15% dimethyl sulfoxide, 1 h, 37 °C) was stopped by phenol and chloroform extraction, and ethanol precipitated in the presence of 1 μM 3′ primer as carrier DNA. Reverse transcription employed RNase H-active reverse transcriptase to cleave the RNA circle during transcription (40-μL reaction, 16 U AMV reverse transcriptase from Pharmacia Biotech, 10 μM 3′ primer GTAAAGTACGC ATAAGGTG, 0.5 mM each dNTP, 50 mM Tris, pH 8.3, 50 mM KCl, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM spermidine, 1 h, 42 °C). The remaining RNA was base hydrolyzed and the cDNA was amplified by PCR. A 50-μL aliquot of the PCR

reaction was gel purified on a 4% NuSieve low-melt agarose gel and the DNA re-amplified in a 2-mL PCR. One milliliter of the PCR was used as template for a single PCR amplification cycle done manually in an 8-mL volume. The PCR product was ethanol precipitated and digested with *Ban*I (200-μL reaction, 200 U *Ban*I from NEB, 15 h). The digestion was inactivated (65 °C for 20 min), phenol and chloroform extracted, and ethanol precipitated. After ligation to the synthetic preannealed 3′ primer/T7 promoter duplex (0.5-mL reaction, 2,000 U T4 DNA ligase from NEB, 1 μM each DNA fragment, 1 mM ATP, 1.5 h, 22 °C), the DNA was ethanol precipitated. Transcription was as described for the first selection, and RNA was gel purified on 6% denaturing gels. Activity was first detected after three rounds of selection, where more than 10% of the pool 3 RNA reacted at an apparent rate of 0.002 min^{-1}. After two more rounds of stringent selection, the pool 5 RNA reacted at an apparent rate of 0.03 min^{-1} with an active fraction of 40%. An alternative strategy for regenerating the T7 promoter that involved an intermolecular RNA ligation was also tested but found to have unacceptable variation in ligation efficiency among ribozyme variants.

Characterization of 2′,5′-branch-forming ribozymes

RNA preparation

Transcription templates for lariat-forming ribozymes were prepared by amplifying the ribozyme inserts of plasmid clones, using the PCR primers of the preceding selection. The PCR DNA was transcribed for 5 min (20-μL reaction, 200 U T7 RNA polymerase, 1 mM of each NTP, 4 μCi α-^{32}P-UTP (3,000 Ci/mmol), 10 mM $MgCl_2$, 1 mM spermidine, 5 mM DDT, 0.1% Triton X-100, 40 mM Tris, pH 7.9, 0.3 μM template DNA, 37 °C), and reactions were quenched by addition of 30 μL of 6.7 mM EDTA and freezing; no lariat formation was detectable after transcription. The crude solution was diluted fivefold for kinetic studies. Templates for truncation and internal deletion constructs were also generated by PCR using appropriate primer pairs. All RNA sequences shorter than 60 nt were prepared by run-off transcription from synthetic DNA templates (Milligan & Uhlenbeck, 1989). (During the course of this study it was noticed that transcription of lariat-forming ribozymes could also be performed under standard conditions without detectable lariat formation after a 30-min transcription. The RNA could also be gel purified, eluted and ethanol precipitated without detecting lariats.)

Reaction kinetics for lariat formation

The lariat-forming ribozyme (<1 μM) was incubated in water for 1 min at 90 °C and then equilibrated at 30 °C for 15 min. If the 3′ primer-binding sequence was not deleted, 3′ primer (1 μM) was added prior to the 90 °C incubation. The reaction was initiated by the addition of ribozyme buffer. Aliquots were removed from the reaction at appropriate time points, quenched by the addition of 1 vol 8 M urea/50 mM EDTA stop solution, and analyzed by denaturing PAGE. Reaction rates were determined by fitting fraction of product, f, at a given time, t, to the equation $f = f_{t0} + f_{t\infty}(1 - e^{-kt})$, where f_{t0} is the fraction of lariat present at zero time, $f_{t\infty}$ the fraction of lariat at the endpoint of the reaction, and k the rate constant for lariat formation (Stage-Zimmermann & Uhlenbeck, 1998). In many cases, the endpoint of reaction was not determined (n.d. in Figs. 3 and 7 and Table 1) and $f_{t\infty}$ was assumed to be 0.7, the value for efficiently reacting variants.

It is interesting to note that ribozymes isolated from the second selection reacted efficiently only when either the 3′ primer-binding sequence was removed or hybridized to the complementary 3′ primer oligodeoxynucleotide. The 3′ primer-binding sequence appears to interfere with catalysis when unpaired (Ekland et al., 1995; Robertson & Ellington, 1999).

Mapping of the branchpoint

Branched RNA was prepared from incubation of triphosphate-containing RNA (pppGGAAAUGCCCAAGCGCUC, derived from DL7.18) with 5′-^{32}P-labeled branch site-containing RNA (GGAGCGCCACUGGAAAACUACAGAGACGCCAGUCAC UCAGAUAUCCUGG, derived from DL7.18) under standard selection conditions. Partial alkaline hydrolysis was performed as described previously (Tuschl et al., 1998).

ACKNOWLEDGMENTS

We thank Jef D. Boeke for the generous gift of recombinant yeast debranching enzyme and Peter Unrau, Chris Burge, and Mike Lawrence for comments on the manuscript. T.T. was supported by postdoctoral fellowships from the Deutsche Forschungsgemeinschaft and Merck/MIT. This work was supported by U.S. Public Health Service MERIT award R37-GM34277 and RO1-AI32486 from the National Institutes of Health to P.A.S., partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute, and by a grant from the Searle Scholars Program/The Chicago Community Trust to D.P.B.

Received August 4, 2000; returned for revision September 8, 2000; revised manuscript received October 17, 2000

REFERENCES

Burge CB, Tuschl T, Sharp PA. 1999. Splicing of precursors to mRNAs by the spliceosomes. In: Gestland RF, Cech TR, Atkins JF, eds. *The RNA world*, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 525–560.

Burke JM, Butcher SE, Sargueil B. 1996. Structural analysis and modification of the hairpin ribozyme. In: Eckstein F, Lilley DMJ, eds. *Nucleic acids and molecular biology*, vol. 10. Berlin: Springer-Verlag. pp 129–143.

Cech TR. 1986. The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. *Cell 44*:207–210.

Chapman KB, Boeke JD. 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell 65*:483–492.

Collins CA, Guthrie C. 1999. Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes & Dev 13*:1970–1982.

Datta B, Weiner AM. 1993. The phylogenetically invariant ACAGAGA and AGC sequences of U6 small nuclear RNA are more tolerant of mutation in human cells than in *Saccharomyces cerevisiae*. *Mol Cell Biol 13*:5377–5382.

Ekland EH, Bartel DP. 1995. The secondary structure and sequence

optimization of an RNA ligase ribozyme. *Nucleic Acids Res 23*: 3231–3238.
Ekland EH, Szostak JW, Bartel DP. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science 269*:364–370.
Fabrizio P, Abelson J. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science 250*:404–409.
Guthrie C. 1991. Messenger RNA splicing in yeast: Clues to why the spliceosome is a ribonucleoprotein. *Science 253*:157–163.
Guthrie C, Patterson B. 1988. Spliceosomal snRNAs. *Annu Rev Genet 22*:387–419.
Haas ES, Armbruster DW, Vucson BM, Danials CJ, Brown JW. 1996. Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res 24*:1252–1259.
Haas ES, Brown JW. 1998. Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res 26*:4093–4099.
Helm M, Brule H, Giege R, Florentz C. 1999. More mistakes by T7 RNA polymerase at the 5′ ends of in vitro-transcribed RNAs. *RNA 5*:618–621.
Herschlag D, Piccirilli JA, Cech TR. 1991. Ribozyme-catalyzed and nonenzymatic reactions of phosphate diesters: Rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. *Biochemistry 30*:4844–4854.
Hertel KJ, Maniatis T. 1999. Serine-arginine (SR)-rich splicing factors have an exon-independent function in pre-mRNA splicing. *Proc Natl Acad Sci USA 96*:2651–2655.
Hwang DY, Cohen JB. 1996. U1 snRNA promotes the selection of nearby 5′ splice sites by U6 snRNA in mammalian cells. *Genes & Dev 10*:338–350.
Incorvaia R, Padgett RA. 1998. Base pairing with U6atac snRNA is required for 5′ splice site activation of U12-dependent introns in vivo. *RNA 4*:709–718.
Kahn JD, Hearst JE. 1989. Reversibility of nucleotide incorporation by *Escherichia coli* RNA polymerase, and its effect on fidelity. *J Mol Biol 205*:291–314.
Kandels-Lewis S, Seraphin B. 1993. Involvement of U6 snRNA in 5′ splice site selection. *Science 262*:2035–2039.
Kim CH, Abelson J. 1996. Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5′ splice site. *RNA 2*:995–1010.
Lesser CF, Guthrie C. 1993. Mutations in U6 snRNA that alter splice site specificity: Implications for the active site. *Science 262*:1982–1988.
Li Z, Brow DA. 1996. A spontaneous duplication in U6 spliceosomal RNA uncouples the early and late functions of the ACAGA element in vivo. *RNA 2*:879–894.
Luo HR, Moreau GA, Levin N, Moore MJ. 1999. The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. *RNA 5*:893–908.
Luukkonen BG, Seraphin B. 1998a. Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. *RNA 4*:167–180.
Luukkonen BG, Seraphin B. 1998b. A role for U2/U6 helix Ib in 5′ splice site selection. *RNA 4*:915–927.
Madhani HD, Bordonne R, Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev 4*:2264–2277.
Madhani HD, Guthrie C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell 71*:803–817.
Madhani HD, Guthrie C. 1994a. Dynamic RNA–RNA interactions in the spliceosome. *Annu Rev Genet 28*:1–26.
Madhani HD, Guthrie C. 1994b. Randomization-selection analysis of snRNAs in vivo: Evidence for a tertiary interaction in the spliceosome. *Genes & Dev 8*:1071–1086.
Maschhoff KL, Padgett RA. 1993. The stereochemical course of the first step of pre-mRNA splicing. *Nucleic Acids Res 21*:5456–5462.
McPheeters DS. 1996. Interactions of the yeast U6 RNA with the pre-mRNA branch site. *RNA 2*:1110–1123.
Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Annu Rev Biochem 64*:435–461.
Milligan JF, Uhlenbeck OC. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol 180*:51–62.
Molinaro M, Tinoco I Jr. 1995. Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: Thermodynamic and spectroscopic applications. *Nucleic Acids Res 23*:3056–3063.
Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNA by the spliceosome. In: Gestland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.
Moore MJ, Sharp PA. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing [see comments]. *Nature 365*:364–368.
Mörl M, Niemer I, Schmelzer C. 1992. New reactions catalyzed by a group II intron ribozyme with RNA and DNA substrates. *Cell 70*:803–810.
Nam K, Hudson RH, Chapman KB, Ganeshan K, Damha MJ, Boeke JD. 1994. Yeast lariat debranching enzyme. Substrate and sequence specificity. *J Biol Chem 269*:20613–20621.
Newman AJ. 1997. The role of U5 snRNP in pre-mRNA splicing. *EMBO J 16*:5797–5800.
Nilsen TW. 1998. RNA–RNA interactions in nuclear pre-mRNA splicing. In: Simons RW, Grunberg-Manago M, eds. *RNA structure and function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 279–307.
O'Keefe RT, Norman C, Newman AJ. 1996. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell 86*:679–689.
Padgett RA, Podar M, Boulanger SC, Perlman PS. 1994. The stereochemical course of group II intron self-splicing. *Science 266*:1685–1688.
Pan T, Jakacka M. 1996. Multiple substrate binding sites in the ribozyme from *Bacillus subtilis* RNase P. *EMBO J 15*:2249–2255.
Pan T, Uhlenbeck OC. 1992. In vitro selection of RNAs that undergo autolytic cleavage with Pb^{2+}. *Biochemistry 31*:3887–3895.
Pyle AM. 1996. Catalytic reaction mechanisms and structural features of group II introns. *Nucleic Acids Mol Biol 10*:75–107.
Qin PZ, Pyle AM. 1998. The architectural organization and mechanistic function of group II intron structural elements. *Curr Opin Struct Biol 8*:301–308.
Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. *EMBO J 15*:1392–1402.
Reyes JL, Gustafson EH, Luo HR, Moore MJ, Konarska MM. 1999. The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site. *RNA 5*:167–179.
Robertson MP, Ellington AD. 1999. In vitro selection of an allosteric ribozyme that transduces analytes to amplicons. *Nat Biotechnol 17*:62–66.
Roiha H, Shuster EO, Brow DA, Guthrie C. 1989. Small nuclear RNAs from budding yeasts: Phylogenetic comparisons reveal extensive size variation. *Gene 82*:137–144.
Sabeti PC, Unrau PJ, Bartel DP. 1997. Accessing rare activities from random RNA sequences: The importance of the length of molecules in the starting pool. *Chem & Biol 4*:767–774.
Segault V, Will CL, Polycarpou-Schwarz M, Mattaj IW, Branlant C, Lührmann R. 1999. Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol Cell Biol 19*:2782–2790.
Sharp PA. 1985. On the origin of RNA splicing and introns. *Cell 42*:397–400.
Sharp PA, Burge CB. 1997. Classification of introns: U2-type or U12-type. *Cell 91*:875–879.
Siatecka M, Reyes JL, Konarska MM. 1999. Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes & Dev 13*:1983–1993.
Sontheimer EJ, Gordon PM, Piccirilli JA. 1999. Metal ion catalysis during group II intron self-splicing: Parallels with the spliceosome. *Genes & Dev 13*:1729–1741.
Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science 262*:1989–1996.
Sontheimer EJ, Sun S, Piccirilli JA. 1997. Metal ion catalysis during splicing of premessenger RNA. *Nature 388*:801–805.
Stage-Zimmermann TK, Uhlenbeck OC. 1998. Hammerhead ribozyme kinetics. *RNA 4*:875–889.
Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell 92*:315–326.
Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci USA 90*:6498–6502.
Sun JS, Manley JL. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes & Dev 9*:843–854.

Tuschl T, Sharp PA, Bartel DP. 1998. Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J 17*:2637–2650.

Umen JG, Guthrie C. 1995. The second catalytic step of pre-mRNA splicing. *RNA 1*:869–885.

Wassarman DA, Steitz JA. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science 257*:1918–1925.

Weiner AM. 1993. mRNA splicing and autocatalytic introns: Distant cousins or the products of chemical determinism? *Cell 72*:161–164.

Welch M, Majerfeld I, Yarus M. 1997. 23S rRNA similarity from selection for peptidyl transferase mimicry. *Biochemistry 36*:6614–6623.

Wolff T, Menssen R, Hammel J, Bindereif A. 1994. Splicing function of mammalian U6 small nuclear RNA: Conserved positions in central domain and helix I are essential during the first and second step of pre-mRNA splicing. *Proc Natl Acad Sci USA 91*:903–907.

Wong I, Patel SS, Johnson KA. 1991. An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. *Biochemistry 30*:526–537.

Yan D, Ares M Jr. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol Cell Biol 16*:818–828.

Zhang B, Cech TR. 1998. Peptidyl-transferase ribozymes: *Trans* reactions, structural characterization and ribosomal RNA-like features. *Chem & Biol 5*:539–553.