Nucleic Acids Research, 1995, Vol. 23, No. 5 849–855

Novel system for analysis of group I 3′ splice site reactions based on functional *trans*-interaction of the P1/P10 reaction helix with the ribozyme's catalytic core

Bharat M. Chowrira[+], Alfredo Berzal-Herranz[§] and John M. Burke*

Markey Center for Molecular Genetics, Department of Microbiology and Molecular Genetics, The University of Vermont, Burlington, VT 05405, USA

Received October 7, 1994; Revised and Accepted January 19, 1995

ABSTRACT

A group I intron from a bacterial tRNA precursor has been converted into an RNA enzyme that catalyzes the efficient polymerization of oligoribonucleotide analogs of tRNA exons using a reaction scheme consisting of multiple cycles of reverse and forward exon ligation reactions. Here, we present results showing that this system represents a novel and useful tool for the analysis of 3′ splice site reactions of group I ribozymes. First, analysis of variant substrates containing base substitutions in group I secondary structure elements P1, P9.0 and P10 confirms that exon polymerization is dependent on these structures, and therefore constitutes an appropriate and relevant model system for studying the exon ligation step of splicing. Second, to probe interactions between the intron's catalytic core and the bases and backbone of the P1/P10 reaction helix, two successful strategies for separating the internal guide sequence from the intron core were devised. One such strategy uses a construct in which the reaction helix interacts functionally with the catalytic core using only tertiary contacts. Further stabilization of this interaction through the inclusion of a 7 bp intermolecular P2 helix generates increased reaction efficiency. Third, when provided with two reaction helices, the ribozyme synthesizes mixed polymers through a mechanism that involves sequential binding and release of the duplexes. Fourth, in these reactions, turnover of the external guide sequence requires unwinding and annealing of the P2 helix, suggesting that P2 unwinding may occur during group I splicing. These results provide novel experimental tools to probe the relatively poorly understood 3′ splice site reactions of group I introns, and may be relevant to ribozyme-catalyzed assembly and recombination of oligomers in prebiotic scenarios.

INTRODUCTION

Group I intron RNA molecules mediate transesterification reactions that are required for expression of intron-containing genes (for review, see ref. 1). The development of ribozymes that catalyse 5′ splice site reactions in *trans* (2,3) has proven to be instrumental in analyzing the structural basis and detailed kinetic mechanism for 5′ splice site recognition and reactions (4). In contrast, group I constructs that direct 3′ splice site reactions in *trans* are very much under-developed. Bimolecular reactions of oligonucleotide analogs of the 5′ exon with unspliced precursor RNAs are well characterized (5). We have developed two unique approaches to the development of bimolecular systems for analysis of 3′ splice site reactions. In one approach, we reconstructed the *Tetrahymena* intron (I) so that the 5′ exon (E1) was located downstream of the intron, and showed that this I · E1 construct can attack oligonucleotide analogs of 3′ splice sites in *trans* through a reaction that generates ligated exons in the form I · E1 · E2 (6).

In the second approach, utilized in the work described here, we have converted a self-splicing group I intron from a $tRNA^{Ile}$ gene of the bacterium *Azoarcus* (7) into an RNA enzyme that catalyzes the synthesis of long RNA polymers using a novel mechanism (8). The reaction (Fig. 1A) is based on reverse and forward 3′ splice site reactions of the self-splicing intron, and so is distinct from previously described polymerization activities of group I introns that are based on 5′ splice site reactions (2,9–13). In our system, RNA substrates function as analogs of ligated exons (E1 · E2) that associate with the intron enzyme (I) by forming duplexes with an internal guide sequence (14), located at the 5′ end of the construct (Fig. 1B). Repeated reverse exon ligation and displacement steps then generate a series of covalent enzyme-linked reaction intermediates of the structure I · $(E2)_n$, where n = 1 to ≥19. Exon 1 (E1) accumulates during the course of the reaction, and attacks the intermediate at sites to the 5′ side of exon 2 (E2), generating reaction products of the structure E1 · $(E2)_m$, where $m \leq n$. The key feature of this reaction pathway is that E2

* To whom correspondence should be addressed

Present addresses: [+]Ribozyme Pharmaceuticals Inc., 2950 Wilderness Place, Boulder, CO 80301, USA and [§]Instituto de Parasitologia Biomedicina 'Lopez Neyra', Ventanilla 11, 18001 Granada, Spain

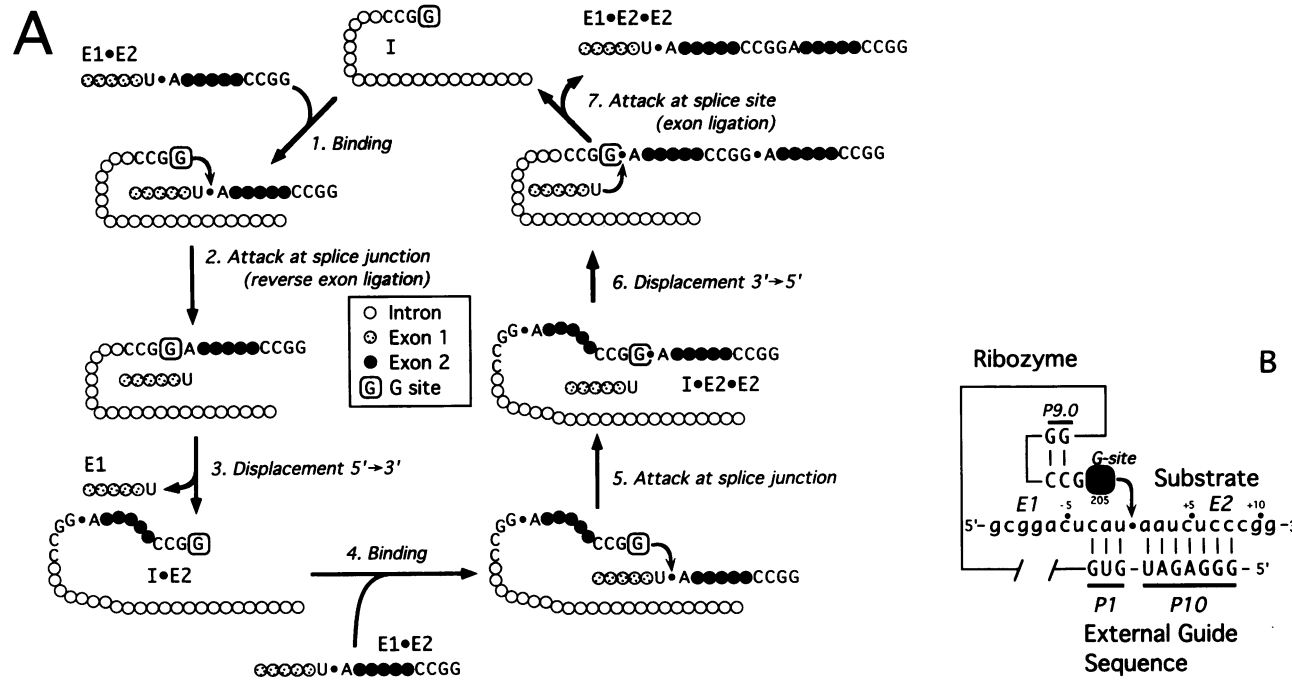

Figure 1. Exon polymerization reaction catalyzed by the modified *Azoarcus* group I intron. (**A**) Mechanism for exon polymerization reaction (8). 1. The intron RNA enzyme (I) binds the ligated exon substrate (E1 · E2) through alignment on the internal guide sequence (IGS). 2. The 3′-terminal guanosine attacks the splice junction in a transesterification reaction that represents the reverse of the exon ligation step of group I splicing, generating the covalent intermediate I · E2. 3. Guanosine at the 3′ end of the intron leaves the active site, and is replaced by the 3′-terminal G of I · E2. Exon 1 (E1) is released. 4. I · E2 binds a second E1 · E2 molecule. 5. A second reverse exon ligation step yields I · E2 · E2. 6. Displacement in the 3′→5′ direction brings the authentic 3′ splice site into the active site of the enzyme. 7. Attack of E1 on the 3′ splice site generates the reaction product E1 · E2 · E2 and regenerates the enzyme. Note that steps 3–5 can be repeated many times to generate intermediates of the structure I · $(E2)_n$, and reaction products of the structure E1 · $(E2)_n$. (**B**) Alignment structure of ligated exons on internal and external guide sequences. The ligated exon substrate (E1 · E2) binds to the internal guide sequence through the complementary pairings P1 and P10. Both internal and external guide sequences are used in the work described here. The 3′ end of the intron is aligned in the active site of the enzyme through P9.0 (19,20) and binding of the terminal G to the G-site (21,22).

is bifunctional, serving as both a 3′ exon and an intron terminus (Fig. 1).

In the work described here, we show that the reaction is dependent upon substrate sequences known to be important for recognition and reactions of 3′ splice sites, and so constitutes a valid system for analysis of the structural basis for 3′ splice site reactions or group I splicing. We demonstrate that the ribozyme can be engineered to use external guide sequences (that is, internal guide sequences that have been separated from the 5′ end of the intron), including those in which the P1/P10 reaction helix interacts with the catalytic core of the intron using only tertiary contacts. Because P1/P10 reaction helices consisting of RNA generated by solid phase synthesis can be utilized, modified nucleotides can be introduced for probing tertiary interactions between the catalytic core, guide sequences and substrates. Interestingly, when provided with two external guide sequences and cognate substrates, the ribozyme synthesizes mixed RNA polymers via a reaction pathway that involves multiple turnover events involving formation and unwinding of an intermolecular P2 helix. Together, our results provide a novel and useful model system for analysis of the structural basis for 3′ splice site reactions that may mimic some aspects of prebiotic polyribonucleotide synthesis and recombination.

MATERIALS AND METHODS

RNA synthesis

Transcriptional templates to generate the modified intron RNA were generated by PCR methods as follows. Plasmid pBGB1.1 (7) was digested with restriction endonuclease *Sph*I, which releases a 0.7 kb fragment containing the *Azoarcus* group I intron. The restriction fragment was gel purified and used as a template for PCR. The template for the modified ribozyme containing an internal guide sequence was prepared by PCR using 5′ primer (5′-TAATACGACTCACTATAGGGAAGCGAAAGAGCAGGGAGATGTGCCTTGCGCC-3′) and 3′ primer (5′-CCGGTTTGTGTGACTTT-3′). The 5′ primer incorporates a promoter for T7 RNA polymerase. Following *in vitro* transcription with T7 RNA polymerase, a modified intron is produced wherein exon 1 and five nucleotides at the 5′ end of the intron (positions 1–5) are deleted and a 18 nt unrelated sequence was substituted. The intron transcript terminates with G_{205}. Transcriptional templates used to synthesize P2-*cis* IVS and P2-*trans* IVS were prepared by PCR using two modified versions of the upstream primer, T7P2-*cis* primer (5′-TAATACGACTCACTATAGGGCCTTGCGCCGGGAAACCA-3′) and T7P2-*trans* IVS (5′-TAATACGACTCACTATAGGGGAAACCACGCAAGGGA-3′).

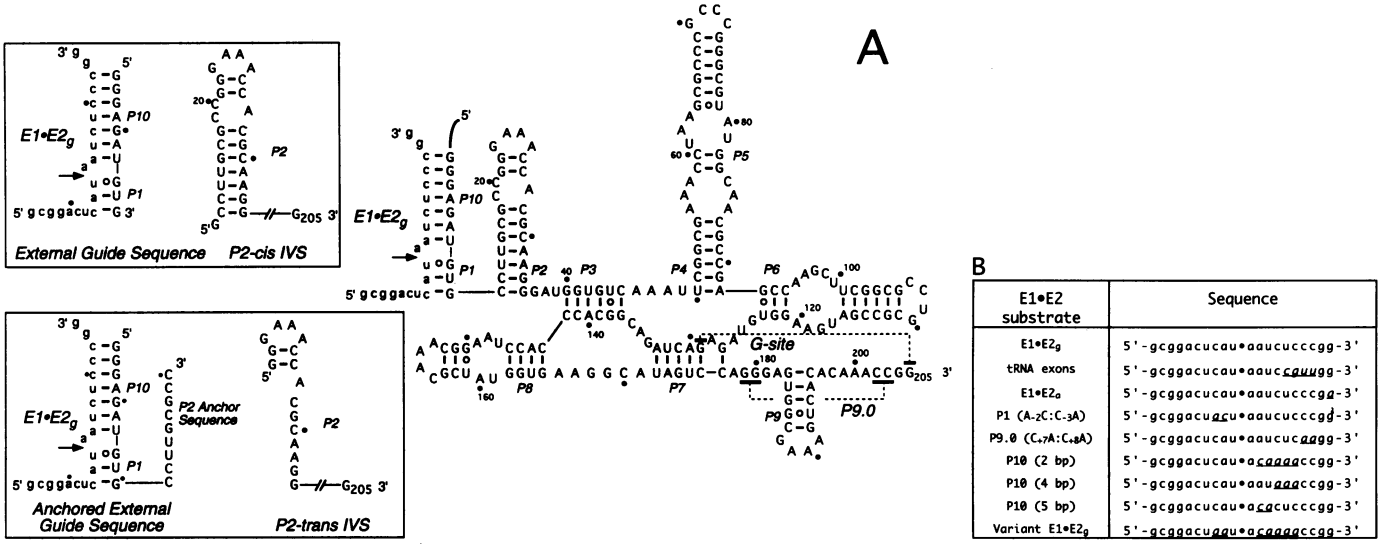

Figure 2. Ribozymes, internal and external guide sequences, and substrates. (**A**) Intron structure and modifications to utilize external guide sequences. The structure of the *Azoarcus* $tRNA^{Ile}$ intron (7) is shown. Modifications were introduced into the 5′ end of the intron to extend and stabilize P10 (8). Exon sequences are shown in lower case. The guide sequence is the sequence GGGAGAUGUG near the 5′ end of the intron. When covalently linked to the intron, this sequence is termed the 'internal guide sequence' (IGS, shown at right). When separated from the intron, it is termed the 'external guide sequence' (boxed constructs). The two strategies used for separating the internal guide sequence [P1(3′) and P10 (3′)] from the intron enzyme are shown in boxes. Intron RNA was transcribed from a DNA template using T7 RNA polymerase as described. Transcription products were separated on a 10% denaturing polyacrylamide gel and RNA was visualized by UV shadowing, excised, then eluted. (**B**) Sequences of ligated exon substrates (E1 · E2) used in this work. The reference sequence is the molecule E1 · $E2_g$ derived from the tRNA exon sequences as described (8). Base substitutions relative to the reference sequence are indicated. E1 · $E2_g$ and other RNAs containing 20 or fewer nucleotides were synthesized using solid phase RNA phosphoramidite methods, deprotected, and gel-purified as described (35,36). RNA phosphoramidites were obtained from Chemgenes, Inc.

Exon polymerization reactions

For polymerization reactions using an internal guide sequence, 0.4 μM unlabeled IVS RNA and 20 μM 3′-end labeled E1 · $E2_g$ RNA were denatured at 95°C for 1 min and renatured on ice for 1 min in a buffer containing 40 mM Tris–HCl (pH 7.5) and 100 mM $(NH_4)_2SO_4$. The mixture was preincubated at 50°C and initiated by adding $MgCl_2$ to a final concentration of 30 mM. After 30 min, reactions were quenched by adding an equal volume of formamide loading buffer and freezing on crushed dry ice. The samples were resolved on 10% denaturing polyacrylamide gels. Gels were dried and quantitated by radioanalytic imaging with a Betascope instrument (Betagen).

For reactions using external guide sequences, 3′ end-labeled E1 · $E2_g$ (10 μM) was hybridized to unlabeled external guide sequence (12 μM) by heating to 90°C for 1 min followed by renaturation at 50°C for 10 min in the presence of 40 mM Tris–HCl pH 7.5, 30 mM $MgCl_2$, 100 mM $(NH_4)_2SO_4$ and 10 mM spermidine. Similarly, P2-*cis* IVS (0.4 μM) was denatured and renatured separately at 50°C for 10 min. Reactions were initiated by mixing the enzyme with the substrate-guide sequence complex at 50°C in the presence of 0.1% ethanol. No polymerization was observed in the absence of spermidine (data not shown).

RESULTS

The exon polymerization reaction is a faithful reporter of group I 3′ splice site reactions

The 3′ splice site reactions of group I introns are dependent on four known RNA structural elements: (i) P1, a short helix between the intron's internal guide sequence (IGS) and E1 (14,15); (ii) P10, a helix between the IGS and E2 (16–18); (iii) P9.0, 1–2 bp involving intron sequences at the 3′ splice site and within the catalytic core (19,20); (iv) binding of the invariant G at the 3′ end of the intron to its binding site in the core (21,22; Figs 1B and 2A). To substantiate that our exon polymerization reaction is an appropriate model system for the study of 3′ splice site structure and reactions, it was important to demonstrate that exon polymerization is dependent upon sequences that form these structures. Therefore, we carried out reactions in which the P1, P9.0 and P10 structures were systematically destabilized by base substitutions within the appropriate sequences of the E1 · E2 polymerization substrate. Figure 3A shows that mutations that destabilize P1 ($A_{-2}C$, $C_{-3}A$) and P9.0 ($C_{+7}A$, $C_{+8}A$) strongly inhibit the formation of polyerization reaction products. In addition, the P1 mutation also inhibits the initial reverse exon ligation reaction that generates IVS · E2. It is important to note that the mutations in P9.0 and the 3′-terminal G (G_{+10}) interfere with polymerization but do not interfere with the initial attack on E1 · E2, because P9.0 and the terminal G reside within the intron for the first cycle, and within E2 for subsequent reactions (this work and ref. 8). The importance of P10, the pairing between the internal guide sequence and exon 2, was tested by the use of E1 · E2 substrates capable of forming P10 structures consisting of 2, 4, 5 and 7 bp (Fig. 3B). Results showed that destabilizing P10 resulted in a progressive decrease in the extent of exon polymerization (Fig. 3B), although production of IVS · E2 was essentially undiminished. The finding that the initial step of the polymerization reaction, reverse exon ligation, is insensitive to mutations in P10 is consistent with previous results obtained using the *Tetrahymena* system (23). Together, these results verify that reactivity in the exon polymerization assay is sensitive to

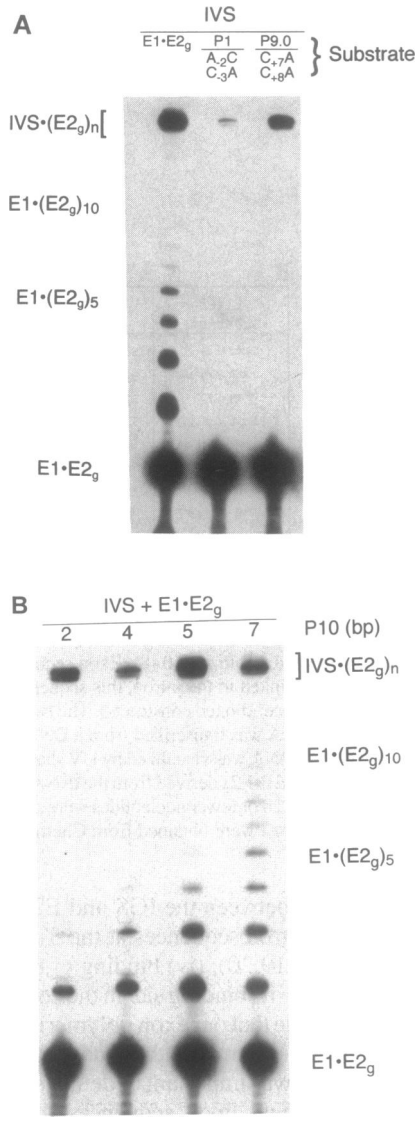

Figure 3. Exon polymerization is sensitive to mutations in sequences responsible for 3′ splice site reactions. (**A**) Exon polymerization reaction with E1 · E2 variants containing mutations in P1 and P9.0. The P1 mutant contains the two base substitution $A_{-2}C$, $C_{-3}A$. The P9.0 mutant contains the two base substitution $C_{+7}A$, $C_{+8}A$. Reactions were carried out for 30 min. IVS · $(E2g)_n$, covalent reaction intermediate or intermediates as described in Figure 1. E1 · $(E2g)_n$, exon polymerization products. Intron enzymes were unlabeled, and E1 · E2 molecules were 3′-end labeled with [^{32}P]pCp to trace levels, as described (8). (**B**) Exon polymerization reaction with E1 · E2 variants containing mutations in P10. The number of base pairs in P10 for each E1 · E2 variant is shown. Reaction conditions were as described (8). In both parts A and B, the origin of the gel is not shown; all radiolabeled species have migrated well into the gel during electrophoresis.

perturbations of the correct 3′ splice site structure, and so is an appropriate assay for 3′ splice site reactions of this group I intron.

Functional *trans*-interactions between catalytic core and reaction helix

The exon polymerization reaction demonstrates that an RNA enzyme is capable of assembling oligonucleotides into rather long polymers through a mechanism that requires a guide sequence resembling, in some respects, an RNA 'template'. This system may recapitulate important features of protein-free RNA replication and/or recombination in an 'RNA world' scenario (12,24–26). In this regard, and also to develop the ability to probe functional groups within the guide sequence, we asked if the ribozyme could utilize an external guide sequence instead of an internal one. Previously, Szostak and co-workers pioneered the concept of external guide sequences in the group I system (27), showing that they could function in the reverse of the group I 5′ splice site splicing reaction, i.e. ligation of oligoribonucleotide analogs of E1 and proximal intron sequences, accompanied by guanosine release. It should be noted that the system developed by Szostak uses a guide sequence that can form P1 (guide sequence paired with exon 1) but not P10 (guide sequence paired with exon 2), while the reaction employed here uses an external guide sequence capable of forming both P1 and P10.

Two experimental strategies were employed to separate the internal guide sequence from the catalytic core of the enzyme; both are depicted in Figure 2A. First, we split the intron between P1 and P2, and supplied the ribozyme (P2-*cis*-IVS, Fig. 2A) with substrate (E1 · E2) pre-annealed to the EGS (ribozyme segments of P1 and P10). The reaction yielded enzyme-linked intermediates and polymerization products up to at least E1 · $(E2)_4$ (Fig. 4A); the length distribution of polymerization products was significantly shorter than for the internal guide sequence. Nevertheless, the reaction efficiency is surprisingly high when it is noted that no base pairs exist between the P1/P10 reaction helix and the ribozyme; binding must take place exclusively through tertiary contacts. We expect that this construct will prove to be particularly valuable for investigations of tertiary contacts between the catalytic core and the reaction helix.

In a second construct, we split the intron from the internal guide sequence by bisecting P2, thus stabilizing binding of the reaction helix to the ribozyme through an intermolecular P2 stem (P2-*trans*-IVS, Figs 2A and 4B). Results showed that the anchored guide sequence served to direct exon polymerization as effectively as the internal guide sequence (Fig. 4C).

Polymerization-specificity changes directed by the external guide sequence

We next asked if we could change the specificity of the RNA sequence that was polymerized by altering the sequence of the EGS. To accomplish this, we changed two of three bases in P1 and five of seven bases in P10 in our variant E1 · E2 substrate and made compensatory changes in the EGS (Fig. 4B). The exon polymerization assay (Fig. 4C) showed, as expected, that no reaction took place when the variant substrate and mismatched anchored EGS (EGS A) were provided. However, when the variant substrate and its cognate EGS (EGS B) were added together, the exon polymerization reaction proceeded with an efficiency similar to that of the original substrate-EGS pair. These results clearly show that we can engineer major changes in the sequence-specificity of the exon polymerization reaction; we anticipate relatively few restrictions on the sequences that can be polymerized (see Discussion).

Synthesis of heterogeneous polymers and turnover of the external guide sequences on the ribozyme via P2 unwinding and reannealing

Because the ribozyme can use alternative guide sequences to direct the polymerization of different RNA sequences, we used

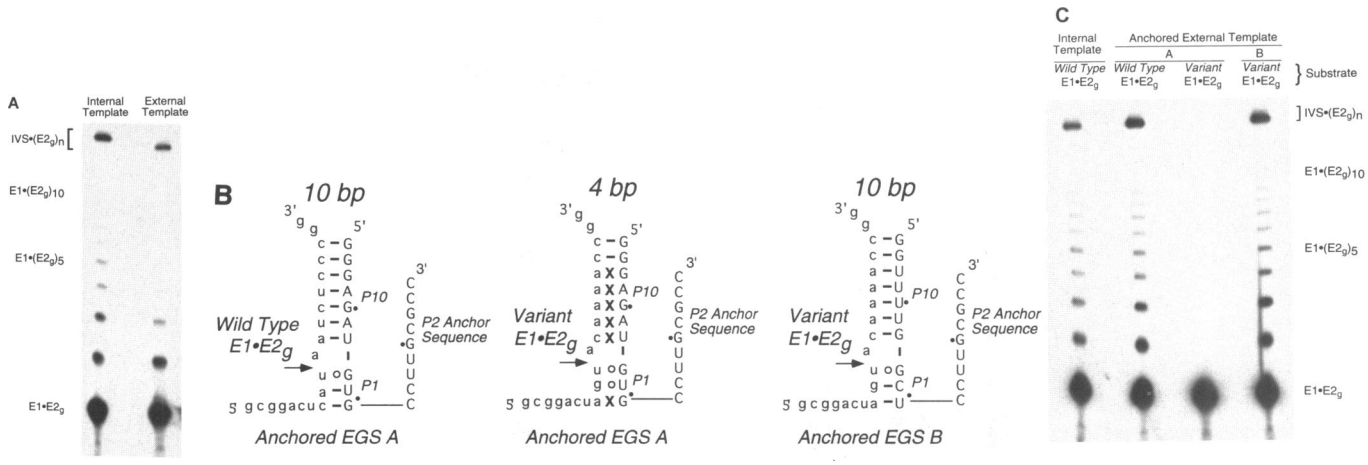

Figure 4. Exon polymerization using external guide sequences. (**A**) Exon polymerization reaction using P2-*cis* ribozyme and a 10 nt unanchored external guide sequence, as shown in Figure 2A (top left). The mobilities of the high molecular weight intermediates derived from the ribozyme using the EGS are higher than those obtained with the IGS because the EGS-dependent ribozyme is 19 nt smaller. (**B**) Anchored external guide sequences and substrates used in experiments of part C and Figure 5. Arrow indicates reaction site (splice junction). (**C**) Exon polymerization reactions using P2-*trans* ribozyme and 19 nt external guide sequences that anchor the EGS to the ribozyme by forming a bimolecular P2 stem, as shown in Figure 2A (bottom left).

the anchored EGS system described above to investigate the consequences of including more than one EGS-substrate pair in an exon polymerization reaction. Two outcomes seemed possible. First, the EGS that initially bound to the ribozyme might remain in place, such that each ribozyme molecule would react with only a single substrate sequence to generate a homogeneous polymer. Alternatively, the guide sequences might turn over on the ribozyme, such that each ribozyme molecule would react with multiple guide sequences to synthesize mixed polymers. Note that turnover of the EGS on the ribozyme requires dissociation of the 7 bp intermolecular P2 helix, followed by P2 formation using a different EGS molecule.

We conducted the following experiment to test for EGS turnover and the synthesis of mixed RNA polymers. The ribozyme was incubated with an equimolar mixture of the two substrates and external guide sequences used in the experiment of Figure 4. The most convincing demonstration of the synthesis of mixed polymers would involve direct sequence analysis of the RNA products. To make sequencing possible, we restricted the complexity of the RNA products in the following manner. First we labeled only one substrate on its 3′ end with trace quantities of [^{32}P]pCp. Second, we used a strategy analogous to the use of 2′,3′-dideoxynucleotides in DNA sequencing, in that we converted the same E1 · E2 substrate into a chain-terminating oligomer by changing G_{+10} to A. In this manner, the ^{32}P-labeled oligomer can be added to the 3′ end of a growing chain, but the reaction product will be unable to undergo further elongation. In the figures, oligoribonucleotide substrates ending in a 3′-terminal A are labeled E1 · $E2_a$; those ending in G are labeled E1 · $E2_g$.

The ensuing reaction yielded polymerization products up to E1 · $(E2)_{11}$ (Fig. 5A), although the length distribution of the products was curtailed by the use of the chain-terminating oligonucleotide. We excised the 30 nt reaction product and determined its sequence using enzymatic methods (Fig. 5B). The sequencing gel showed clearly that this band was a pure sequence mixed polymer of the structure $E1_{var}$ · $E2_{var}$ · $E2_a$. Knowing the polymerization mechanism (8) and the reaction specificity of this intron, we can deduce the reaction pathway through which this elongation product was formed (Fig. 5C). Several steps are required: (i) binding of the variant substrate-EGS duplex, (ii) transesterification to yield I · $E2_{var}$ and free $E1_{var}$, (iii) dissociation of the $E1_{var}$–EGS duplex, (iv) binding of the E1 · $E2_a$–EGS duplex, (v) transesterification to yield I · $E2_{var}$ · $E2_a$ and free E1, (vi) dissociation of the E1–EGS duplex, (vii) binding of the $E1_{var}$–EGS duplex, and finally (viii) transesterification to yield $E1_{var}$ · $E2_{var}$ · $E2_a$ and the regenerated intron enzyme. The overall reaction for synthesis of the observed 30 nt RNA product, then, is:

$$E1_{var} \cdot E2_{var} + E1 \cdot E2_a \rightarrow E1_{var} \cdot E2_{var} \cdot E2_a + E1$$

A unique and interesting feature of this pathway is that two EGS turnover events are required, first to link $E2_a$ to the 3′ end of the enzyme-linked intermediate I · $E2_{var}$, and second to deliver $E1_{var}$ to the active site for formation of the final reaction products.

It should be noted that the reaction pathway that generates the 30 nt RNA product that was sequenced represents only one of a number of possible reactions. In particular, if the ribozyme is presented with multiple EGS-substrate combinations capable of being elongated, a highly heterogeneous population of RNA products would be expected.

DISCUSSION

In the work presented here, we report several new and significant findings that extend previously reported results obtained using the *Azoarcus* exon polymerization reaction (8,28). First, we have used mutational analysis to verify that the exon polymerization reaction is a faithful reporter of group I 3′ splice site reactions; indeed it appears to be more sensitive to structural changes in the reacting RNA species than the simple reverse exon ligation reaction (this work and ref. 23). Second, functional reconstitution of the reaction after covalent separation of the internal guide sequence from the ribozyme is demonstrated. Third, we show that the P1/P10 reaction helix can functionally interact with the catalytic core utilizing only tertiary contacts. Fourth, P2 unwinding and reannealing occurs during the cycles of reverse and forward exon ligation reactions.

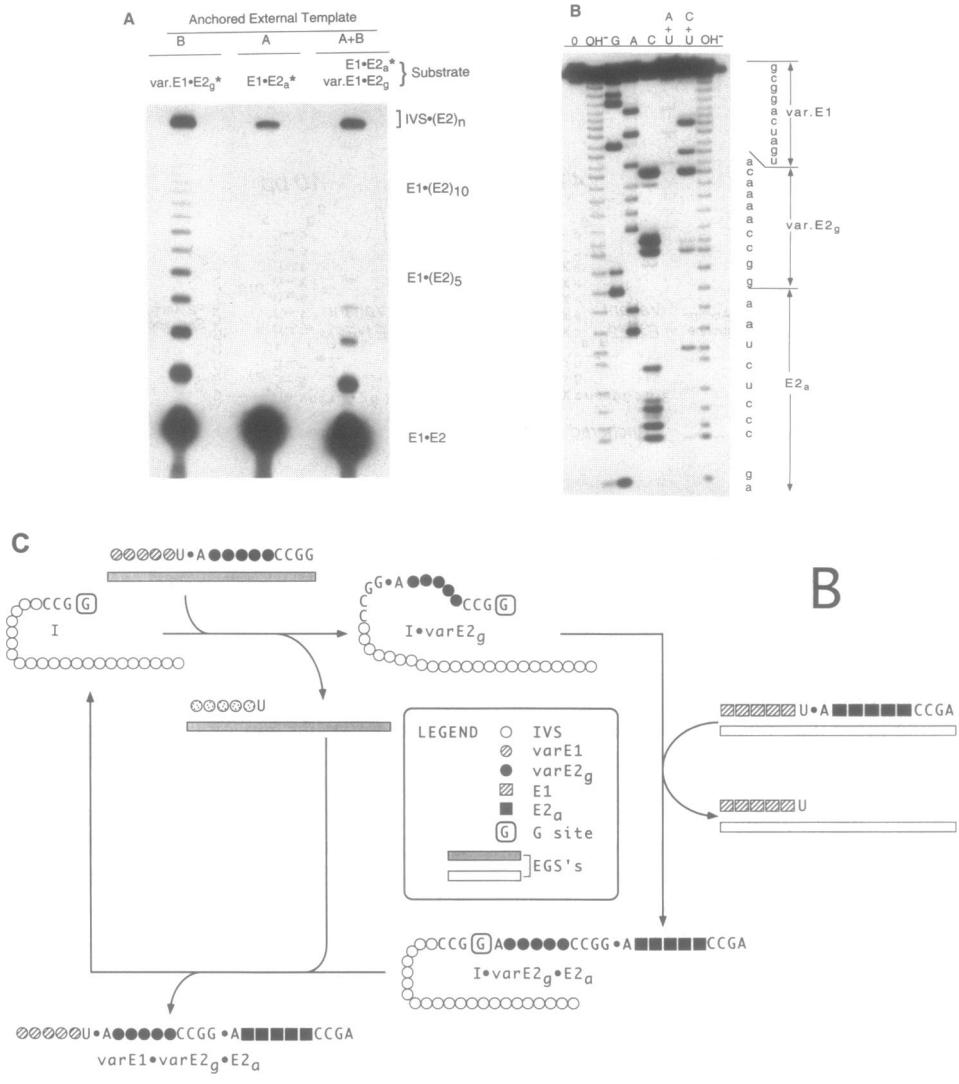

Figure 5. Synthesis of mixed polymers using multiple external guide sequences. (**A**) Polymerization using two substrates and external guide sequences. Exon polymerization reactions were set up with the variant E1 · $E2_g$ substrate and EGS B (left), with the E1 · $E2_a$ substrate and EGS A (center), and with both substrate-EGS combinations (right). Note that the $G_{+10}A$ substitution in E1 · $E2_a$ blocks further elongation after formation of I · $E2_a$, so that no polymerization products are observed. Reaction conditions were as described in Figure 4A. In the reaction with two substrate-EGS combinations (right), E1 · $E2_a$ was 3′-end labeled, and was the only labeled RNA species present in the reaction. (**B**) RNA sequence analysis of the 30 nt extension product. The 30 nt extension product from the reaction with two substrate-EGS combinations (part A, right) was isolated from the gel and sequenced using sequence-specific endoribonucleases T_1 (G), PhyM (A and U), *B.cereus* (C and U), U_2 (A) and Cl_3 (C) (37). Enzymes were obtained from United States Biochemicals, Inc. Autoradiogram of a 20% sequencing gel is shown. (**C**) Reaction pathway for synthesis of the 30 nt extension product, varE1 · $varE2_g$ · $E2_a$.

Fifth, we find that this RNA enzyme can use more than one template in an exon polymerization reaction to generate co-polymers of mixed oligoribonucleotide substrates.

An important feature of this system is that oligoribonucleotide analogs of exons and guide sequence can be generated by solid-phase synthesis. Combined with our functional reconstitution based on non-Watson–Crick interactions, this system promises to be very useful for probing tertiary contacts between the P1/P10 reaction helix and the catalytic core. Such experiments are currently being initiated in our laboratory. A disadvantage of the exon polymerization reaction is that it is not highly amenable to kinetic analysis, because of the multitude of reaction substrates and products, and the fact that reaction products themselves become substrates for both forward and reverse reactions in what is essentially a disproportionation reaction.

We have shown that a drastic change in reaction specificity can be accomplished without a detectable decrease in polymerization efficiency. Therefore, those parts of the reaction helix that dock with the catalytic core of the *Azoarcus* intron in 3′ splice site reactions appear to do so through contacts with the conserved G–U wobble pair in P1 and with the sugar–phosphate backbone; similar findings for forward and reverse 5′ splice site reactions of the *Tetrahymena* intron have been reported (29,38–40). Sargueil and Tanner have shown that a pseudoknotted substrate containing exon 1, the internal guide sequence, and exon 2 can function as a substrate for both splicing steps of the *Tetrahymena* intron (30).

With regard to the synthesis of the mixed polymer $E1_{var} \cdot E2_{var} \cdot E2_a$, it is of interest that only $E1_{var}$ was identified at the 5′ end of the polymerization product, as determined by sequencing. This indicates that there is a strong preference for the use of $E1_{var}$ over E1 in the forward exon ligation step that forms the last step in the synthesis of this product. This is an interesting result, because the E1–EGS duplex is bound to the ribozyme prior to the synthesis of the last reaction intermediate, $I \cdot E2_{var} \cdot E2_a$. In order for the observed reaction product to be generated, the E1 · EGS duplex must be released from the ribozyme and then be replaced with the cognate $E1_{var}$ · EGS duplex (Fig. 5C). This result clearly demonstrates that unwinding and re-formation of the 7 bp intermolecular P2 helix can occur during the course of 3′ splice site reactions, but does not tell us whether or not such a conformational change is essential for the second step of splicing. We attribute the strong preference for $E1_{var}$ over E1 in this last step to the fact that formation of a stable P10 complex for the last transesterification reaction to generate the observed product requires the use of the EGS to which $E1_{var}$ is bound, rather than that to which E1 is bound.

What are the limitations on the sequences that can be polymerized using the reaction described here? Efficient group I reactions require that the nucleotide at the 3′ end of E1 must be U and that at the 3′ end of E2 must be G, although it is possible that *in vitro* selection methods might be used to isolate highly active ribozymes that are not subject to these constraints (31–34). The sequence at all other positions in the E1 · E2 substrate should be essentially unrestricted, because sequences in P1, P9.0 and P10 are found to vary widely in both naturally-occurring group I introns and synthetic constructs. We anticipate that the lengths of sequences that can be polymerized may be reduced slightly or expanded considerably. Because E2 must function as both a 3′ exon and an intron terminus, there is likely to be a lower limit on the length of the sequence that can be polymerized, which we estimate to be 5–7 nt. The upper limit on the length of E2 is unknown; however, we expect that E2 may be lengthened considerably as long as it does not form a structure that would inhibit intron folding or limit access of the 3′ terminus of E2 to the catalytic core of the enzyme.

An interesting feature of the exon polymerization reaction is that the ribozyme recognizes and utilizes ribonucleotides adjacent to the splice site, and at the 3′ end of exon 2. This raises the possibility that non-nucleic acid materials could be inserted into the E2 sequence between P10 and P9.0 (Fig. 1B), in such a way that the ribozyme could catalyze the synthesis of polymers that were, in large part, non-nucleic acids.

The finding that an RNA enzyme can utilize multiple external guide sequences to direct the synthesis of RNA polymers with new sequences suggests a mechanism through which single-stranded oligonucleotides could have been assembled into novel polynucleotides during prebiotic evolution. While modern-day group I introns are unlikely to directly resemble such postulated prebiotic ribozymes, one can envision several possible mechanisms, both synthetic and recombinogenic, through which reactions similar to those described here could have contributed to the emergence of primitive self-sustaining RNA replicons.

ACKNOWLEDGEMENTS

This work was supported by grant GM36981 from the NIH. JMB is a member of the Vermont Cancer Center. We thank Johanna Polsenberg and Charles Keller for the synthesis of RNA and DNA oligonucleotides used in this work, Michele Millham for the purification of T7 RNA polymerase, Art Zaug for the polymerase purification protocol, David Shub for the plasmid clone containing the *Azoarcus* intron, and Andrew Ellington for stimulating discussions.

REFERENCES

1 Cech,T.R. (1990) *Annu. Rev. Biochem.*, **59**, 543–568.
2 Zaug,A.J. and Cech,T.R. (1986) *Science*, **231**, 470–475.
3 Zaug,A.J., Been,M.D. and Cech,T.R. (1986) *Nature*, **324**, 429–433.
4 Cech,T.R., Herschlag,D., Piccirilli,J.A. and Pyle,A.M. (1992) *J. Biol. Chem.*, **267**, 17479–17482.
5 Inoue,T., Sullivan,F.X. and Cech,T.R. (1985) *Cell*, **43**, 431–437.
6 Tasiouka,K. and Burke,J.M. (1994) *Gene*, **144**, 1–7.
7 Reinhold-Hurek,B. and Shub,D.A. (1992) *Nature*, **357**, 173–176.
8 Chowrira,B.M., Berzal-Herranz,A. and Burke,J.M. (1993) *EMBO J.*, **12**, 3599–3605.
9 Been,M.D. and Cech,T.R. (1988) *Science*, **239**, 1412–1416.
10 Doudna,J.A. and Szostak,J.W. (1989) *Nature*, **339**, 519–522.
11 Bartel,D.P., Doudna,J.A., Usman,N. and Szostak,J.W. (1991) *Mol. Cell. Biol.*, **11**, 3390–3394.
12 Green,R. and Szostak,J.W. (1992) *Science*, **258**, 1910–1915.
13 Doudna,J.A., Usman,N. and Szostak,J.W. (1993) *Biochemistry*, **32**, 2111–2115.
14 Been,M.D. and Cech,T.R. (1986) *Cell*, **47**, 207–216.
15 Waring,R.B., Towner, P., Minter,S.J. and Davies,R.W. (1986) *Nature*, **321**, 133–139.
16 Davies,R.W., Waring,R.B., Ray,J.A., Brown,T.A. and Scazzocchio,C. (1982) *Nature*, **300**, 719–724.
17 Partono,S. and Lewin,A.S. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8192–8196.
18 Ritchings,B.W. and Lewin,A.S. (1988) *Nucleic Acids Res.*, **20**, 2349–2353.
19 Burke,J.M., Esherick,J.S., Burfeind,W.R. and King, J. L. (1990) *Nature*, **344**, 80–82.
20 Michel,F., Netter,P., Xu,M.-Q. and Shub,D.A. (1990) *Genes Dev.*, **4**, 777–788.
21 Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J.W. (1989) *Nature*, **342**, 391–395.
22 Been,M.D. and Perrotta,A.T. (1991) *Science*, **252**, 434–437.
23 Woodson,S.A. and Cech,T.R. (1989) *Cell*, **57**, 335–345.
24 Gilbert,W. (1986) *Nature*, **319**, 618.
25 Cech,T.R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4360–4363.
26 Gesteland,R.F. and Atkins,J.F. (1993) in *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA World.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
27 Doudna,J.A., Couture,S. and Szostak,J.W. (1991) *Science*, **251**, 1605–1608.
28 Berzal-Herranz,A., Chowrira,B.M., Polsenberg,J.F. and Burke,J.M. (1993) *Biochemistry*, **32**, 8981–8986.
29 Pyle,A.M. and Cech,T.R. (1991) *Nature*, **350**, 628–631.
30 Sargueil,B. and Tanner,N.K. (1993) *J. Mol. Biol.*, **233**, 629–643.
31 Green,R., Ellington,A.D. and Szostak,J.W. (1990) *Nature*, **347**, 406–408.
32 Beaudry,A.A. and Joyce,G.F. (1992) *Science*, **257**, 635–641.
33 Szostak,J.W. (1992) *Trends Biochem. Sci.*, **17**, 89–93.
34 Burke,J.M. and Berzal-Herranz,A. (1993) *FASEB J.*, **7**, 106–112.
35 Chowrira,B.M. and Burke,J.M. (1991) *Biochemistry*, **30**, 8518–8522.
36 Scaringe,S.A., Francklyn,C. and Usman,N. (1990) *Nucleic Acids Res.*, **18**, 5433–5441.
37 Kuchino,Y. and Nishimura,S. (1989) *Methods Enzymol.*, **180**, 154–163.
38 Bevilacqua,P.C. and Turner,D.H. (1991) *Biochemistry*, **30**, 10632–10640
39 Herschlag,D., Eckstein,F. and Cech,T.R. (1993) *Biochemistry*, **32**, 8299–8311.
40 Strobel,S.A., Cech,T.R. (1993) *Biochemistry*, **32**, 13593–13604.