Trans Insertion−Splicing: Ribozyme-Catalyzed Insertion of Targeted Sequences into RNAs†

Ashley K. Johnson, Joy Sinha, and Stephen M. Testa*

Department of Chemistry, University of Kentucky, Lexington, Kentucky 40506

Received March 15, 2005; Revised Manuscript Received June 17, 2005

ABSTRACT: A group I intron-derived ribozyme from *Pneumocystis carinii* has been previously shown to bind an exogenous RNA substrate, splice out an internal segment, and then ligate the two ends back together (the trans excision−splicing reaction). We demonstrate that this same ribozyme can perform a trans insertion−splicing (TIS) reaction, where the ribozyme binds two exogenous RNA substrates and inserts one directly into the other. Reactions were optimized for both yield and rate, with optimum reactions carried out in 10 mM $MgCl_2$ for 2 h. Reaction products are stable, with no visible loss at extended times. The ribozyme recognizes the two substrates primarily through base pairing and requires an ωG on the ribozyme and an ωG on the sequence being inserted. We give evidence that the reaction mechanism is not the reverse of the trans excision−splicing reaction, but is composed of three steps, with intermediates attached to the ribozyme. Surprisingly, the internal guide sequence of the ribozyme is utilized to sequentially bind both substrates, forming independent P1 helices. This is an indication that ribozymes with essentially the native intron sequence can catalyze reactions significantly more dynamic and complex than self-splicing. The implications of group I intron-derived ribozymes being able to catalyze this unique reaction, and via this mechanism, are discussed.

Self-splicing group I introns are catalytic RNAs that excise themselves from RNA transcripts. These introns bind and spatially orient their substrates (i.e., exons) for catalysis, at least initially by base pairing, usually through three major molecular recognition contacts. The resultant helices, called P1, P9.0, and P10, define the precise sites of catalysis within the bound substrates, which for the self-splicing reaction consists of two distinct catalytic steps: 5′-cleavage and exon ligation. To study the folding and catalytic activity of these biological catalysts in particular, and RNAs in general, group I introns are routinely synthesized and studied in vitro using constructs that lack covalently attached substrates (i.e., as ribozymes) (*1*−*3*). Synthetic substrates that mimic the exons can then be added to the ribozyme to isolate and analyze individual aspects of the self-splicing reaction.

Group I intron-derived ribozymes can bind multiple substrates, as there are three independent base pairing elements. Moreover, each of the two catalytic steps is reversible, giving four different catalytic reactions. Therefore, group I intron-derived ribozymes, by judicious choice of substrates (and in some cases modest modifications to the ribozyme), can catalyze reaction pathways that are distinct from self-splicing. Examples include trans splicing ribozymes, whereby either terminal end of the RNA transcript is replaced with a new terminal end (*4*, *5*). Furthermore, the polymerization of one (*6*) and two (*7*) molecules has been seen with an *Azoarcus* group I intron-derived ribozyme, leading to progressively longer polymerization products. Recombination of RNA substrates using *Tetrahymena* and *Azoarcus* group I intron-derived ribozymes has produced hybrid RNAs with the 5′-end of one RNA attached to the 3′-end of another RNA (*8*). A *Tetrahymena* ribozyme has even been shown to cleave and rejoin pentacytidylic acid, synthesizing a polycytidylic acid product (*1*).

We previously reported that a *Pneumocystis carinii* group I intron-derived ribozyme can catalyze a trans excision−splicing (TES)[1] reaction, whereby the ribozyme binds an exogenous RNA substrate, removes a section from the middle, and splices the ends back together (Figure 1A) (*9*−*11*). We now report that the *P. carinii* ribozyme can also catalyze the trans insertion−splicing (TIS) reaction, whereby one RNA substrate is inserted into a second RNA substrate (Figure 1B). The mechanism of the reaction was determined using modified substrates to trap isolated reaction steps, and surprisingly, the three-step insertion reaction is not the reverse of the excision reaction. The ribozyme utilizes the internal guide sequence (IGS) to separately bind both exogenous substrates, forming two independent, sequential P1 helices. In addition, the ribozyme appears to consecutively position the ωG of the ribozyme and the 3′-G on the insert (attached to the ribozyme and presumably acting as the ωG of the ribozyme) into the guanosine binding site (GBS) for catalysis. Apparently, group I intron-derived ribozymes are capable of more dynamic and complex molecular interactions than previously demonstrated. The proposed mechanism for the TIS reaction and implications for group I intron-derived ribozyme reactions are discussed.

† This work was supported by the Kentucky Research Challenge Trust Fund and by the Muscular Dystrophy Association.

* To whom correspondence should be addressed. Telephone: (859) 257-7076. Fax: (859) 323-1069. E-mail: testa@uky.edu.

[1] Abbreviations: TIS, trans insertion−splicing; TES, trans excision−splicing; IGS, internal guide sequence; GBS, guanosine-binding site.

10.1021/bi0504815 CCC: $30.25
Published on Web 07/16/2005

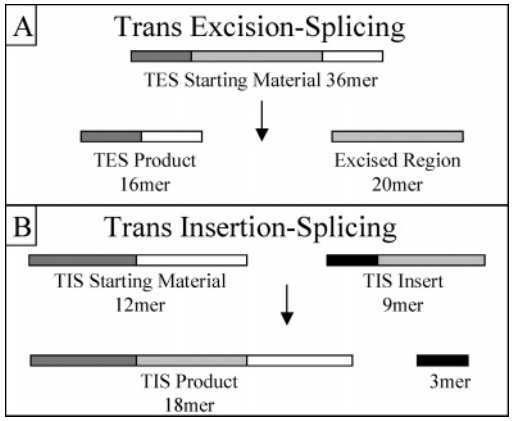

FIGURE 1: Trans excision−splicing reaction and trans insertion−splicing reaction. (A) The 36mer TES starting material reacts with ribozyme rPC to give the 16mer TES product and 20mer excised region (*9*). (B) The 12mer starting material and 9mer insert react with ribozyme rPC to give the 18mer TIS product.

MATERIALS AND METHODS

Synthesis and Preparation of Oligonucleotides. RNA oligonucleotides were purchased from Dharmacon, Inc. (Lafayette, CO), and deprotected according to the manufacturer's instructions. Oligonucleotide concentrations were calculated on the basis of UV absorption measurements using a Beckman DU 650 UV spectrophotometer (Beckman Coulter, Inc., Fullerton, CA). Designated oligonucleotides were 5′-end radiolabeled and purified by gel electrophoresis as described previously (*3*). Designated oligonucleotides were 3′-end radiolabeled by ligating 5′-end radiolabeled Cp to the 3′-end of the oligonucleotide. The Cp was 5′-end radiolabeled in 10 μL consisting of 5 μL of 250 μCi of [^{32}P]γATP (Amersham-Pharmacia, Piscataway, NJ), 2 μL of 50 mM cytidine 5′-monophosphate (CMP) (Sigma, St. Louis, MO), 1 μL of 10 units/μL T4 polynucleotide kinase (New England Biolabs, Beverly, MA), and 1 μL of 10× polynucleotide kinase buffer (supplied by New England Biolabs). The reaction was run for 90 min at 37 °C, and then the kinase was deactivated by incubation at 65 °C for 15 min. 5′-end radiolabeled pCp (5′-*pCp) was ligated to the 3′-end of the oligonucleotide in a 10 μL reaction mixture consisting of 1 μL of DMSO, 2 μL of 10 μM RNA, 2 μL of 20 units/μL T4 RNA ligase (New England Biolabs), 4 μL of 5′-*pCp (approximately 20 μM), and 1 μL of 10× T4 RNA ligase buffer (New England Biolabs). The reaction mixtures were incubated for 16 h at 4 °C. The 3′-radiolabeled oligonucleotides were purified by gel electrophoresis as described for the 5′-end radiolabeled oligonucleotides. The substrate names and sequences are shown in Table 1.

Ribozyme Preparation. The *P. carinii* (PC) ribozyme plasmid (proposed secondary structure of the rPC ribozyme shown in Figure 2) was linearized in a 50 μL reaction mixture consisting of 8 μg of plasmid, 50 units of *Xba*I (Invitrogen, Grand Island, NY), and 1× React 2 buffer at 37 °C for 2 h. Linearization was confirmed by visualization on a 1% agarose gel. The linearized DNA was purified using a QIAquick PCR purification kit (Qiagen Inc., Valencia, CA) and eluted in water. Five other ribozymes with varying 3′-ends were made from the PCR products derived from the PC plasmid (Table 1). The upstream PCR primer for all five ribozymes was 5′CTCTAATACGACTCACTATAGAGGG3′.

Table 1: Ribozyme, Starting Material, and Insert Sequences[a]

Ribozyme	3′-End Sequence
rPC	~~~~CUAG
rPC-ωA	~~~~CUA**A**
rPC-1	~~~~**UGU**G
rPC-3	~~~~CUAGCUCGUG
rPC-3ωA	~~~~CUA**A**CUCGUG
rPC-4	~~~~CUAGCUCGUGAAACAU

Starting Material	Sequence
12mer	AUGACUAAACAU
12mer-dU	AUGAC**dU**AAACAU
6mer	AUGACU
6mer-dU	AUGAC**dU**

Insert	Sequence
9mer	GCUCUCGUG
9mer-dU	GC**dU**CUCGUG
9mer-ωA	GCUCUCGU**A**
7mer-minus-3′UG	GCUCUCG_ _
7mer-minus-5′GC	_ _UCUCGUG

[a] The sequences highlighted with gray boxes represent the substrates and ribozyme used in the standard TIS reaction. The bold and underlined nucleotides indicate the positions that differ from the standard reaction components. Note that rPC-3, rPC-3ωA, and rPC-4 are the intermediate ribozymes with part of the starting material and insert attached to the 3′-end. The symbol wavy line represents the 5′-end of the ribozyme that is the same for all the ribozymes that were tested. The complete ribozyme sequence is shown in Figure 2.

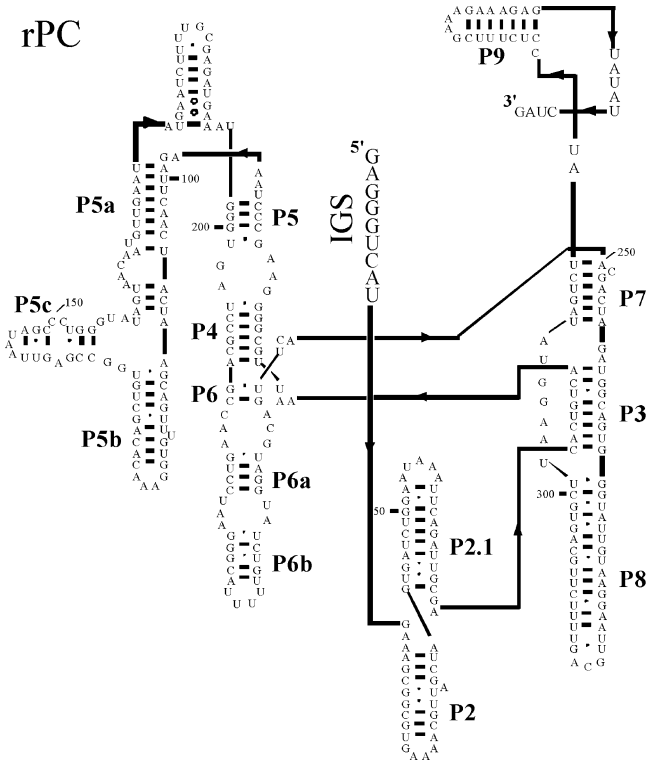

FIGURE 2: Proposed secondary structure of a *P. carinii* group I intron-derived ribozyme (rPC).

The following sequences are the downstream primers for each ribozyme (the variable region is underlined): 5′CACAATATACTCTTTCTTTCGAAAGAGG3′ for rPC-1, 5′TTAGATATACTCTTTCTTTCGAAAGAGG3′ for rPC-

ωA, 5′CACGAGCTAGATATACTCTTTCTTTCGAAAGAGG3′ for rPC-3, 5′CACGAGTTAGATATACTCTTTCTTTCGAAAGAGG3′ for rPC-3-ωA, and 5′ATGTTTCACGAGCTAGATATACTCTTTCTTTCGAAAGAGG3′ for rPC-4. The PCR products were gel purified using a Qiagen gel extraction kit (Qiagen Inc.). Runoff transcription was performed for 2 h in 100 μL reaction mixtures consisting of 1−2 μg of linear DNA, 50 units of T7 RNA polymerase (New England Biolabs), 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, 5 mM spermidine, 1 mM rNTP mix, and 62.5 μg/mL BSA. The ribozymes were purified using a Qiagen RNeasy purification kit (Qiagen Inc.).

Trans Insertion−Splicing Reaction. The extent of the TIS reaction was optimized over a range of rPC concentrations (from 10 to 1000 nM), 9mer insert concentrations (from 10 to 3000 nM), $MgCl_2$ concentrations (from 2 to 15 mM), times (from 1 to 180 min), and temperatures (from 37 to 50 °C) for the reaction with 5′-end radiolabeled 12mer starting material AUGACUAAACAU. Ultimately, optimum reactions were found to consist of 2 h at 44 °C with 200 nM rPC, 1 μM 9mer insert, and 10 mM $MgCl_2$. All reactions were carried out using approximately 1 nM 5′-end radiolabeled 12mer starting material in a buffer consisting of 50 mM Hepes (25 nM Na^+) and 135 mM KCl at pH 7.5.

In the reaction mixtures, first 3 μL of ribozyme in appropriate buffer was preannealed at 60 °C for 5 min. The ribozyme was then slowly cooled to 44 °C. Reactions were initiated by adding 2 μL of precombined substrates (radiolabeled starting material and cold insert). Optimum TIS reactions were run with 5′-end radiolabeled 12mer starting material, but 3′-end radiolabeled starting material and 5′- and 3′-end radiolabeled insert were also used in reactions to elucidate the mechanism. Reactions were terminated after 2 h by adding 5 μL of stop buffer (10 M urea, 3 mM EDTA, and 0.1× TBE). The reaction mixtures were denatured for 1 min at 90 °C and then separated on a 12% polyacrylamide−8 M urea gel. The gel was transferred to chromatography paper and dried under vacuum. The bands were visualized and quantified on a Molecular Dynamics Storm 860 Phosphorimager. The observed rate constant, k_{obs}, for the TIS reaction was obtained from the plot of the percent TIS product formed over time (*3*).

Product Isolation and Identification. The TIS product was gel purified and sequenced by partial nuclease digestion along with the synthetic version of the expected product. The TIS reaction was scaled up 50-fold (250 μL) and carried out as described above using the optimized conditions, except the reactions were conducted in five 50 μL volumes, concentrated to 30 μL after 2 h, and terminated with 10 μL of stop buffer. The product band was cut out of the gel and eluted from the gel matrix for 40 min by crushing with a magnetic stir bar in 400 μL of elution buffer [0.3 M sodium acetate, 5 mM EDTA, 10 mM Tris-HCl (pH 7.5), and 0.1% SDS]. The eluent was decanted, and a second round of elution was performed. The product was ethanol precipitated overnight and concentrated to 30 μL. The radiolabeled TIS product and synthetic 18mer were enzymatically sequenced using RNA nucleases T1, U2, CL-3, and *Bacillus cereus* (Research Unlimited, Wellington, New Zealand) essentially as described previously (*9*). *B. cereus* reactions used 0.33 unit of *B. cereus* in 33 mM sodium citrate (pH 5.0) and 1.7 mM EDTA.

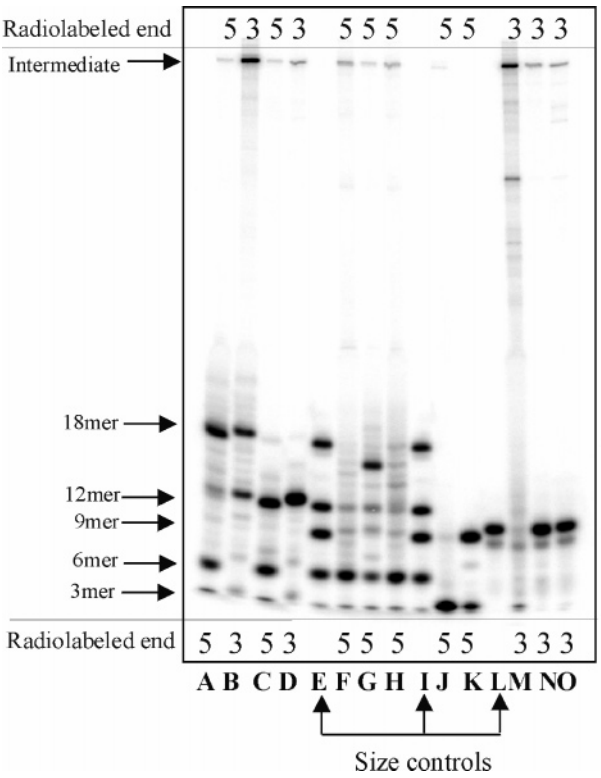

FIGURE 3: Evidence for the first nucleophilic attack in the TIS reaction. Polyacrylamide gel showing reactants and products of the TIS reaction using 200 nM ribozyme, 1 μM insert, 1 nM starting material, and 10 mM $MgCl_2$ run for 2 h at 44 °C. Lanes E and I contained 5′-end radiolabeled 6mer, 9mer, 12mer, and 18mer (sequences in Table 1) run as a size control. Lane L contained 3′-end radiolabeled 9mer run as a size control. Lanes A and B contained the standard TIS reaction run with rPC ribozyme, 9mer insert, and 5′-end radiolabeled 12mer (lane A) or 3′-end radiolabeled 12mer (lane B). The rest of the lanes are described in the text. The starting material was radiolabeled on either the 5′- or 3′-end for reactions run in lanes A−D and F−H. The insert was radiolabeled on either the 5′- or 3′-end for reactions run in lanes J, K, and M−O and are not at optimum concentrations for TIS. The radiolabeled end is indicated at the top of the gel.

RESULTS

The P. carinii Group I Intron Catalyzes the TIS Reaction. Work on the TIS reaction began by simply trying to reverse the TES reaction (Figure 1A). Instead of just one substrate, the TIS reaction requires two substrates (called starting material and insert; see Figure 1B). Reactions with 5′-end radiolabeled 12mer starting material, nonlabeled 9mer insert, and the rPC ribozyme (secondary structure shown in Figure 2) gave a large radiolabeled product, indicative of a catalytic reaction (Figure 3, 18mer in lane A). To analyze whether an insertion reaction occurred, the reaction was carried out as described above, except using 3′-end radiolabeled 12mer starting material (Figure 3, 18mer in lane B). In this reaction, we also produced a larger radiolabeled product. Note that because the 3′-end radiolabeled starting material is one nucleotide larger than the 5′-end radiolabeled starting material (due to the method of labeling), the 3′-radiolabeled product band is one nucleotide larger than the 5′-radiolabeled product band. Apparently, this product contains both the 5′- and 3′-ends of the starting material, and so is likely the result of an insertion reaction. For confirmation, the product band was excised from the polyacrylamide gel and its sequence was confirmed by enzymatic sequencing (Figure 4). The 18mer

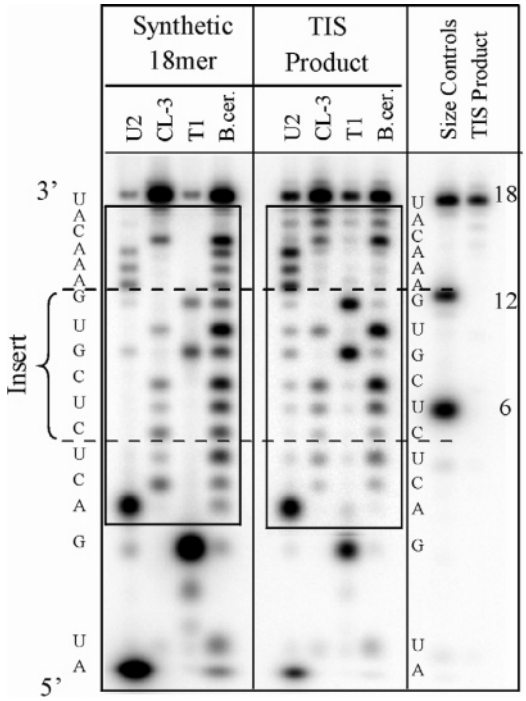

FIGURE 4: Sequencing of the 5′-end radiolabeled TIS product. The 18mer TIS product was isolated from a large reaction of rPC, 9mer insert, and 5′-end radiolabeled 12mer starting material. Both synthetic 18mer product and isolated 18mer TIS product were enzymatically sequenced next to each other using U2, CL-3, T1, and *B. cereus* endonucleases. The dotted line shows the position of the inserted region (CUCGUG) between the 5′- and 3′-ends of the 12mer starting material. Nuclease U2 is specific for adenosine, CL-3 for primarily cytidine, T1 for guanosine, and *B. cereus* for primarily cytidine and uridine. Note that the intensity of the bands in the boxed regions was enhanced relative to the rest of the gel for easier visualization.

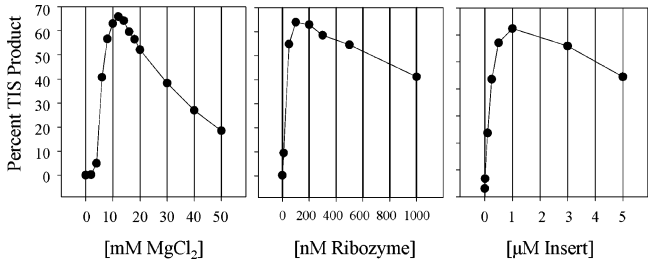

FIGURE 5: TIS optimization graphs. Except for the changing variable, TIS reactions were run under the optimum reaction conditions of 200 μM rPC ribozyme, 1 μM 9mer insert, and 10 mM $MgCl_2$ using 1 nM 5′-end radiolabeled 12mer starting material for 2 h at 44 °C. Each graph represents the average of two independent assays. The standard deviations for each point in the graphs were <20%.

product AUGACUCUCGUGAAACAU appears to be the result of insertion of the underlined 6mer region from the 9mer insert, GCUCUCGUG, into the middle of the 12mer starting material, AUGACUAAACAU. Therefore, although the exact product was unexpected, this is the first evidence that group I intron-derived ribozymes can catalyze a TIS reaction.

The reaction was optimized in terms of the yield (Figure 5) and the observed rate constant (Figure 6). Optimum reaction conditions for TIS are 200 nM ribozyme (rPC), 1 μM 9mer insert, and 10 mM $MgCl_2$ using 1 nM radiolabeled 12mer starting material. The TIS reaction produces 62.9 ± 1.3% 18mer TIS product (62.9% is the average from five

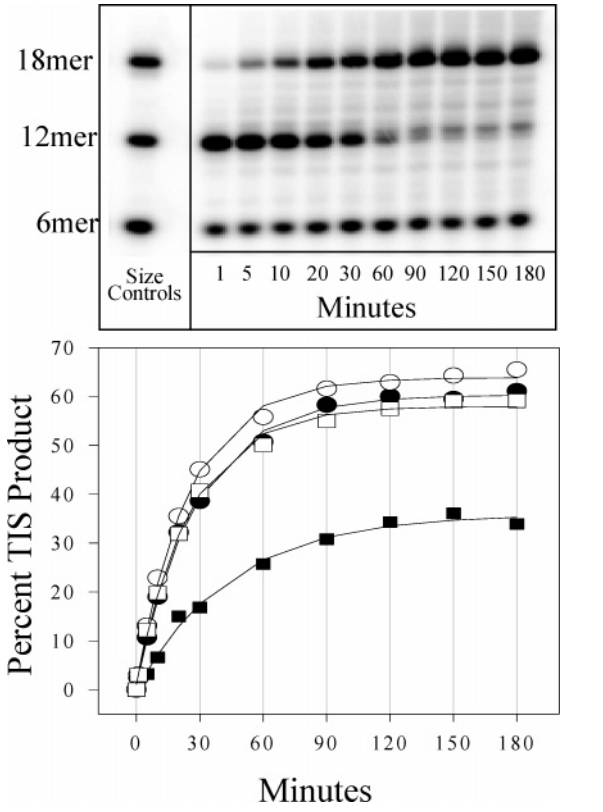

FIGURE 6: Representative time study gel and observed rate constants for the TIS reaction. The polyacrylamide gel shows reactants and products of the TIS reaction using 200 nM ribozyme (rPC), 1 μM insert (9mer), 1 nM 5′-end radiolabeled starting material (12mer), and 10 mM $MgCl_2$ run over 3 h at 44 °C. The graph shows four plots of the average amount of 18mer TIS product from two independent time studies run in 6 (■), 10 (●), 14 (○), or 18 mM $MgCl_2$ (□). The k_{obs} values are 0.021 ± 0.007 min^{-1} in 6 mM $MgCl_2$, 0.035 ± 0.004 min^{-1} in 10 mM $MgCl_2$, 0.040 ± 0.003 min^{-1} in 14 mM $MgCl_2$, and 0.039 ± 0.003 min^{-1} in 18 mM $MgCl_2$. The standard deviations for each point in the graphs were <20%.

independent TIS reactions, and the error is the standard deviation). The reaction is complete after 2 h. Times studies were performed by removing aliquots of the reaction at various time points and placing them in stop buffer. Time studies were carried out at four different $MgCl_2$ concentrations, and the observed rate constants, k_{obs}, were determined (Figure 6). The k_{obs} value for the TIS reaction is 0.04 min^{-1}, and is the same at 10, 14, and 18 mM $MgCl_2$. The k_{obs} value when using 6 mM $MgCl_2$, however, was 0.02 min^{-1}, which is half the k_{obs} for the optimum condition and results in half the yield. Apparently, $MgCl_2$ concentrations as low as 6 mM inhibit the ability of the ribozyme to catalyze the reaction.

The mechanism of the TIS reaction was elucidated by analyzing the effects of modifying functional groups in the substrates and ribozyme, and by initiating the reaction using intermediates as the starting material. Key to the elucidation of the mechanism was the discovery that a large intermediate product, indicative of the substrate attached to the ribozyme, was observed only when substrates were 3′-end radiolabeled, not when they were 5′-end radiolabeled. This suggested that the ωG on the ribozyme was involved in catalysis and that intermediates were attached to the ribozyme. The proposed TIS mechanism is shown in Figure 7.

Step 1: The ωG on the Ribozyme Performs a Nucleophilic Attack on the Insert. The sequencing results show that the

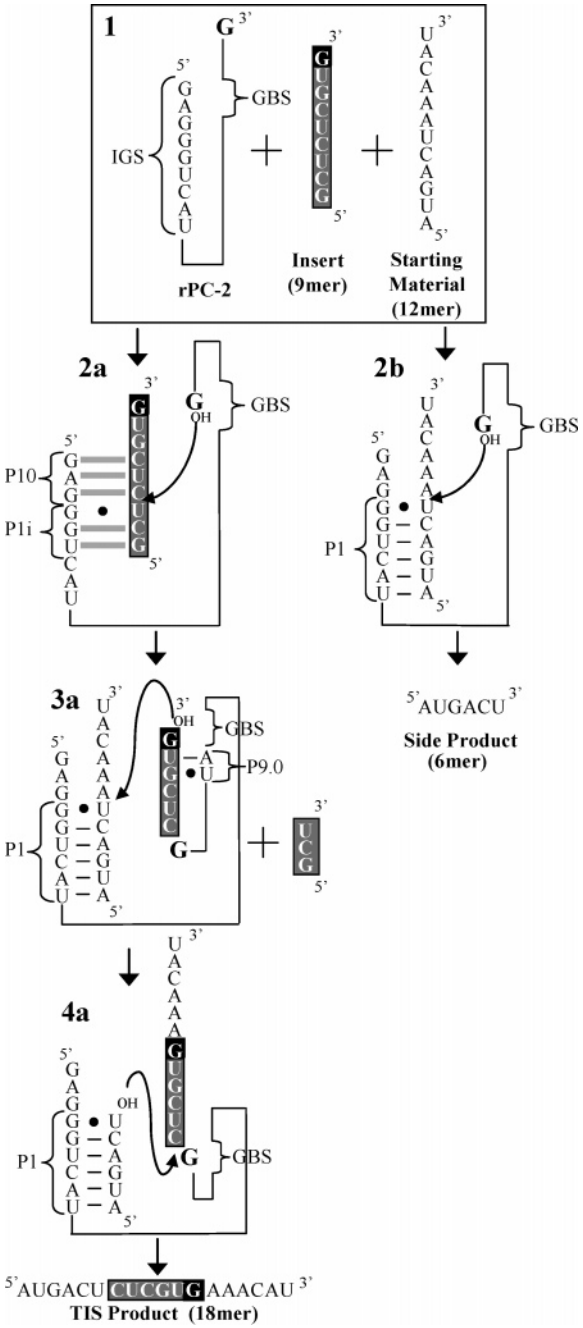

FIGURE 7: Proposed mechanism for the trans insertion−splicing reaction. The rPC ribozyme is represented by black lines, and the IGS sequence and ωG (in bold) are shown. The insert sequence is shown in white letters with a gray background. The ωGi in the insert is distinguished with a black background. (1) The rPC ribozyme binds the 9mer insert (pathway a). Alternatively, the 12mer starting material (pathway b) can bind the ribozyme to give the dead-end products. (2a) The insert (highlighted in gray and black) forms P1i and P10 helices with the IGS of the ribozyme. The ωG in the ribozyme (in bold) attacks at the 5′-splice site in the 9mer insert. (3a) The 12mer starting material displaces the insert fragments from P1i and forms a second P1 helix. A nucleophilic attack by the ωGi on the insert (attached to the ribozyme, shown with a black background) occurs at the 5′-splice site in the 12mer starting material. (4a) A nucleophilic attack of the 3′-U from the 5′-half of the starting material at the ωG in the ribozyme (in bold) produces the 18mer TIS product (inserted region with gray and black background between the 5′- and 3′-halves of the 12mer starting material). (2b) This alternative pathway involves the 12mer starting material forming the P1 helix first. The ωG on the ribozyme (in bold) can attack at the 5′-splice site in the 12mer, leading to 6mer product and the 3′-half of the 12mer attached to the ribozyme.

9mer insert is losing its 5′-GCU (GCUCUCGUG) during the TIS reaction. The loss of this GCU 3mer from the insert suggests that the insert might first be forming a P1 helix (called P1i) with the IGS of the ribozyme (Figure 7, structure 2a). This interaction would place a G-U wobble pair at the 5′-splice site where the ωG on the ribozyme could attack the insert.

To see if the ribozyme's ωG is involved in the mechanism proposed in Figure 7, the ωG in ribozyme rPC was replaced with an ωA (rPC-ωA). Although an ωA has been shown to give the second step of splicing in a group I intron from *Anabaena* (*12*), it does not result in a functional interaction with the GBS of *P. carinii* ribozymes (*11*). We expect that if the ωG in the ribozyme were important, this alteration would prevent the first reaction step, as well as TIS product formation. In this reaction, using 5′- or 3′-end radiolabeled 12mer starting material, no appreciable 18mer TIS product formed (Figure 3, lanes C and D). Note that in the 5′-end radiolabeled reaction (Figure 3, lane C), production of 6mer is occurring through ribozyme-mediated hydrolysis at the 5′-splice site, as seen in the TES reaction with the same ribozyme (*9*). Also, in the 3′-end radiolabeled reaction (Figure 3, lane D), much less of the large intermediate band forms compared to that in the reaction using rPC (compare to Figure 3, lane B), as expected if the ribozyme cannot perform the first step of the TIS reaction. Note that transcription using T7 RNA polymerase has been shown to add extra nucleotides to the 3′-end of transcripts and that a faint 18mer or intermediate band might be due to the addition of Gs to the 3′-end of the ribozyme in that process.

To test that the 9mer insert is forming a P1i helix (as depicted in Figure 7, structure 2a), the 9mer insert was shortened on either the 5′- or 3′-end. The 7mer-minus-$^{3'}$UG insert (GCUCUCG_ _) is shortened by two nucleotides on its 3′-end. This insert can still form a P1i helix, but should give a shortened TIS product. Reactions run with 5′-end radiolabeled 12mer starting material, rPC ribozyme, and a 7mer-minus-$^{3'}$UG insert (Figure 3, lane G) show production of a shorter TIS product, as expected. Likewise, an insert was shortened by two nucleotides on its 5′-end, 7mer-minus-$^{5'}$GC insert (_ _UCUCGUG), which should disrupt P1i formation. If relatively stable P1i formation is required, then this insert will not form TIS product. Reactions run with 5′-end radiolabeled 12mer starting material, rPC ribozyme, and 7mer-minus-$^{5'}$GC insert (Figure 3, lane F) show no TIS production, as expected, which is consistent with formation of a P1i helix in the TIS reaction.

Since the 9mer insert loses its $^{5'}$GCU and appears to form a P1i helix, an insert with a 2′-deoxy-U at the resulting 5′-splice site (9mer-dU, GCdUCUCGUG) was tested to see if this change inhibits 18mer production. The catalytic activity of self-splicing group I intron-derived ribozymes is substantially reduced when the 2′-OH group at the 5′-splice site is removed from the site of catalysis (*13*, *14*). Since the individual chemical steps in the TIS reaction are thought to be equivalent to those in self-splicing group I intron reactions, substitution of a 2′-deoxy-U at the site of catalysis should greatly inhibit the TIS reaction. Indeed, the reaction run with 5′-end radiolabeled 12mer starting material, rPC ribozyme, and the 9mer-dU insert gives greatly inhibited TIS production (Figure 3, lane H), as expected.

To directly show that the ribozyme is attacking the 9mer insert, reactions were run with 5′-end radiolabeled insert instead of radiolabeled starting material. These reactions are run with much less insert compared to the standard reaction and, therefore, are not under optimum TIS conditions. Nevertheless, when the reaction is run with 5′-end radiolabeled 9mer (GCUCUCGUG), all of the 9mer is cleaved to a smaller product (Figure 3, lane J), which appears to be the expected 3mer side product GCU. Running the same reaction using the ribozyme rPC-ωA and 5′-end radiolabeled 9mer gives a marked reduction in the extent of 3mer formation (Figure 3, lane K), as expected. The 3mer produced in this reaction is presumed to be mostly through ribozyme-mediated hydrolysis since the ωA in the ribozyme cannot attack at the splice site (*11*).

When the reaction is run with 3′-end radiolabeled 9mer insert (also not run with optimum TIS insert concentrations), a large intermediate band forms (Figure 3, lane M). This expected band is an intermediate formed by the insert attaching to the 3′-end of the ribozyme (shown in Figure 7, structure 3a). The formation of this large intermediate band is greatly inhibited when 3′-end radiolabeled 9mer is run with rPC-ωA (Figure 3, lane O). Likewise, this inhibition also occurs when the reaction is carried out with 3′-end radiolabeled 9mer-dU (Figure 3, lane N), which puts the inactive 2′-deoxy-U at the 5′-splice site. These results indicate that the 9mer loses the 3mer ($^{5'}$GCU) from its 5′-end and the 3′-portion of the insert becomes ligated to the 3′-end of the ribozyme in the process (Figure 7).

Step 2: ωGi on the Insert Performs a Nucleophilic Attack on the Starting Material. Once the insert fragment is attached to the 3′-end of the ribozyme, the 12mer starting material binds the ribozyme and forms a P1 helix. The 3′-G from the insert (called ωGi) now acts as the nucleophile and can attack the G-U wobble pair at the new 5′-splice site in the starting material (Figure 7, 3a).

The importance of the ωGi on the 9mer insert was tested by changing ωGi to an ωAi (GCUCUCGUA, called 9mer-ωAi). If ωGi is interacting with the GBS and acting like the ωG on the ribozyme, then changing this base to an ωA should prevent the TIS reaction. Running the TIS reaction with 5′-end radiolabeled 12mer starting material, rPC ribozyme, and the 9mer-ωA insert gives no 18mer TIS product (Figure 8, lane C). For comparison, the standard TIS reaction is shown in Figure 8, with the 5′-end radiolabeled reaction in lane A and the 3′-end radiolabeled reaction in lane B. The reaction run with 3′-end radiolabeled 12mer starting material, rPC ribozyme, and 9mer-ωA insert also shows no TIS product (Figure 8, lane D). In addition, the TIS reaction was also prevented when the ωG on the insert was substituted with a 2′-deoxy-G (data not shown): this substitution is known to impede the binding of ωG to the GBS (*13*). Evidently, ωGi on the insert is critical to the TIS reaction. Since the rPC ribozyme has an ωG, the 6mer (Figure 8, lane C) and the large intermediate (Figure 8, lane D) are likely due to the competing reaction depicted in part 2b of Figure 7.

The TIS reaction was initiated at the second step of the three-step reaction by reacting 12mer starting material with an intermediate ribozyme (rPC-3) that has the insert fragment already attached to its 3′-end (Figure 7, structure 3a). This TIS reaction was run exactly like the standard TIS reaction, except no 9mer insert was added. As seen in lane F of Figure 8, the reaction of the rPC-3 ribozyme and 5′-end radiolabeled 12mer starting material produces the 18mer TIS product. Since no insert is added to the reaction, the inserted sequence must be coming from the 3′-end of the ribozyme. The lower product yields are not surprising, probably because the insert sequence on the ribozyme intermediate (rPC-3) may not be positioned properly for the second nucleophilic attack without previously undergoing the first reaction step.

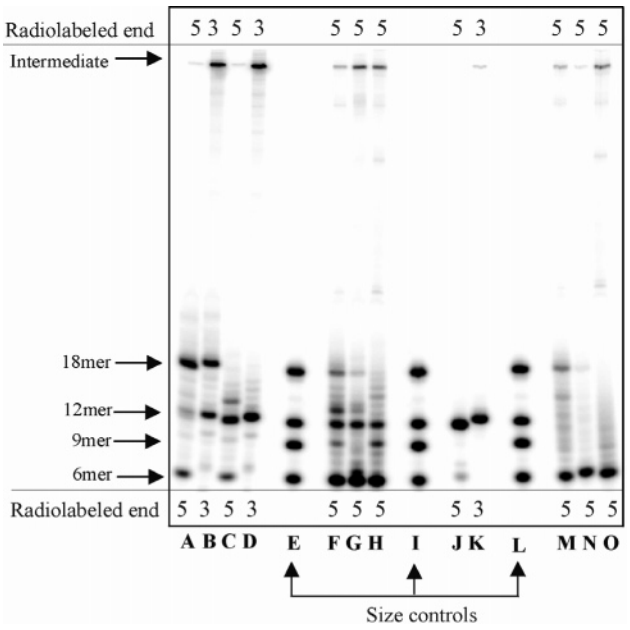

FIGURE 8: Evidence for the second and third nucleophilic attacks in the TIS reaction. Polyacrylamide gel showing reactants and products of the TIS reaction using 200 nM ribozyme, 1 μM insert, 1 nM starting material, and 10 mM $MgCl_2$ run for 2 h at 44 °C. Lanes E, I, and L contained 5′-end radiolabeled 6mer, 9mer, 12mer, and 13mer (sequences in Table 1) run as a size control. Note that reactions in lanes F−H, J, K, and M−O were run without added insert since the insert sequence was attached to the ribozyme intermediate. Lanes A and B contained the standard TIS reaction run with rPC ribozyme, 9mer insert, and 5′-end radiolabeled 12mer starting material (lane A) or 3′-end radiolabeled 12mer starting material (lane B). The rest of the lanes are described in the text. The starting material was radiolabeled on either the 5′- or 3′-end for all the reactions. Note that lanes M−O were run with the intermediate 6mer starting material. The radiolabeled end is indicated at the top of the gel.

In our proposed mechanism, the second step of the TIS reaction occurs via a nucleophilic attack on the 5′-splice site of the 12mer starting material. This was tested by changing the ribo-U at the 5′-splice site to a 2′-deoxy-U (12mer-dU, AUGACdUAAACAU). In the reaction with 5′-end radiolabeled 12mer-dU starting material, the ribozyme intermediate rPC-3, and no added 9mer insert, no TIS product forms (Figure 8, lane J). Likewise, the same reaction run with 3′-end radiolabeled 12mer-dU shows no TIS product (Figure 8, lane K). Clearly, changing the ribo-U to a 2′-deoxy-U at this position in the 12mer starting material (AUGACdU-AAACAU) suggests that this position forms the second 5′-splice site.

Step 3: The 3′-U on the 5′-Half of the Starting Material Attacks the ωG of the Ribozyme, Forming the TIS Product. After the second step of the TIS reaction, the 3′-half of the starting material, as well as the insert region, is attached to the 3′-end of the ribozyme (Figure 7, structure 4a). The 3′-U

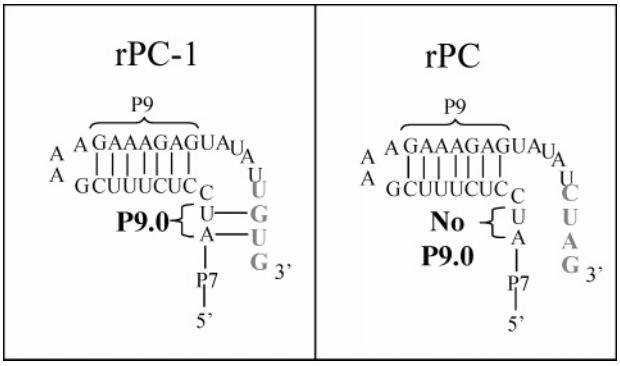

FIGURE 9: Secondary structure at the 3′-end of the rPC-1 and rPC ribozymes. The rPC-1 ribozyme is the exact sequence of the *P. carinii* group I intron, capable of forming a P9.0 helix. The rPC ribozyme used in the TIS reaction differs from rPC-1 by three nucleotides and cannot form a P9.0 helix.

on the 5′-half of the starting material can then mimic the second step of self-splicing and attack at ωG (and not at ωGi) of the ribozyme. To initiate the reaction at the third reaction step, we synthesized the ribozyme intermediate rPC-4, which has the insert and 3′-half of the starting material on the 3′-end of the ribozyme (Figure 7, structure 4a), and ran the reaction using 5′-end radiolabeled 6mer (AUGACU, the 5′-half of the 12mer starting material). This reaction produces the expected 18mer TIS product (Figure 8, lane M, 29% 18mer, average of two reactions). Since the 18mer product forms with only 6mer added, the product sequence must be coming from the ribozyme, as proposed. The control reaction of 5′-end radiolabeled 6mer and rPC (the standard ribozyme) (Figure 8, lane O), shows no 18mer product. In addition, running the reaction with rPC-4 and the 6mer with its ribo-U (on its 3′-end) changed to a 2′-deoxy-U (AUGACdU called 6mer-dU) greatly reduces the amount of TIS product (Figure 8, lane N, 3% 18mer, average of two reactions), indicating that the 3′-U on the 5′-half of the starting material is the nucleophile for the third reaction step.

Finally, the third step of the reaction can be inhibited by changing the ωG in the rPC-3 ribozyme to an ωA (rPC-3ωA in Table 1). The reaction utilizing 12mer starting material and rPC-3ωA initiates at the second step of TIS (Figure 7, 3a), but the ωA in the ribozyme should inhibit the TIS reaction at the third step (Figure 7, 4a). The results show less 18mer product forms (Figure 8, lane G), as expected, compared to the reaction with rPC-3 (Figure 8, lane F). This result suggests that the ωG in the ribozyme is not only required for the first step, but is important for the third step of the TIS reaction as well.

P9.0 Formation Inhibits the TIS Reaction. The TIS reaction was also run with the native *P. carinii* intron sequence (called rPC-1). In reactions with rPC-1 (data not shown), 18mer product yields were 6-fold lower than with rPC (11.0 ± 0.8%; the percentage is the average value from four independent TIS reactions, and the error is the standard deviation). The only difference between rPC-1 and rPC is three nucleotides at the 3′-end of the ribozymes (Figure 9). This difference allows only rPC-1 to form a P9.0 helix. Apparently, P9.0 formation inhibits the TIS reaction, perhaps because it inhibits the proposed conformational shift that occurs between the first and second steps, essentially halting the reaction at the second step.

DISCUSSION

TIS Mechanism. All evidence suggests there are three reaction steps in forming the 18mer TIS product (Figure 7). In step 1, the 9mer binds the IGS of the ribozyme forming a P1i helix, and the ωG on the ribozyme then attacks at the newly formed 5′-splice site, giving GCU + rPC-ωGCUCGUG + AUGACUAAACAU (Figure 7, 2a). This reaction step appears to be the same as the reverse of the second step of group I intron self-splicing (*15*), and has been seen in other ribozyme reactions (*1*, *6*−*8*). In step 2, the 12mer starting material displaces the insert fragment and forms a second P1 helix. The ωGi from the 9mer insert (attached to the 3′-end of the ribozyme) attacks at the G-U wobble pair at the newly formed second 5′-splice site, giving AUGACU + rPC-ωGCUCGUGAAACAU (Figure 7, 3a). This step is similar to the first reaction step (described above). In step 3, the free 3′-OH on the 3′-U of AUGACU can then attack at the ωG of the ribozyme intermediate (rPC-ωGCUCGUGAAA-CAU), releasing the 18mer TIS product AUGACUCUC-GUGAAACAU (Figure 7, 4a). This reaction step appears to be the same as the second step of group I intron self-splicing (*15*). No experiments contradicted our proposed mechanism. A competing reaction occurs, however, when the 12mer starting material forms the P1 helix first (before the insert forms the P1i helix). This reaction results in the formation of 6mer (AUGACU) and the 3′-end of the 12mer (AAACAU) attached to the 3′-end of the ribozyme (Figure 7, 2b). This ribozyme intermediate has no ωG and therefore cannot perform the TIS reaction.

The TIS reaction, along with other ribozyme reactions (*1*, *6*−*8*), is dependent on the ribozyme ending in G (which is called ωG). There is evidence that the T7 RNA polymerase used to synthesize RNA frequently adds one or two template-independent nucleotides onto the 3′-end of RNA transcripts (*16*). There is no simple way to ensure that extra nucleotides are not being added to our ribozymes. Nevertheless, when we synthesize the ribozyme ending in ωA, it substantially reduces reactivity (Figure 3, lane C). Apparently, although template-independent nucleotides may be added to the TIS ribozymes, it appears that they are not prevalent and that they are not predominantly Gs. This agrees with other studies that show the extra nucleotides added by T7 RNA polymerase tend to be mostly As and Cs (*16*, *17*).

Factors that Influence the TIS Reaction. The TIS reaction was optimized for both yield and rate. The optimum reaction conditions are 10 mM $MgCl_2$, 200 nM rPC ribozyme, and 1 μM 9mer insert, using 1 nM radiolabeled starting material (Figure 5). The relative proportions of these reaction components are fairly specific (200:1000:1 ribozyme:insert: starting material ratio), as decreasing or increasing the concentrations of insert or ribozyme gives decreased TIS product yields. At ribozyme concentrations higher than the optimum 200 nM, the amount of 18mer product begins to decrease (Figure 5). This decrease could be due to the insert and starting material not binding the same ribozyme. Another possibility is that higher concentrations of ribozyme may lead to TIS product breakdown. TIS reactions run with insert concentrations higher than the optimum (1 μM) also show a decrease in the level of 18mer product formation. This could be due to the 12mer starting material not competing as well for binding the IGS in the presence of increased 9mer insert

concentrations. Also, the optimum $MgCl_2$ concentration range is surprisingly narrow, with a concentration of 10−14 mM $MgCl_2$ giving the greatest TIS yields (Figure 5) and rate constants (Figure 6). The optimum reaction time for TIS is 2 h, and no degradation of the TIS product is seen at extended reaction times. This indicates that the TIS product is stable once it forms.

It appears that P9.0 formation inhibits the TIS reaction. The TIS reaction mechanism requires a conformational shift between the first and second nucleophilic attacks to disrupt the P1i helix and form the second P1 helix. In group I introns, the P9.0 helix forms with the two nucleotides preceding ωG of the intron and can help position the ωG into the GBS (*18*−*20*). In addition, the GBS in the ribozyme binds both the exogenous G and ωG for the first and second nucleophilic attacks, respectively (*21*). The Sullenger group has shown that a *Tetrahymena* group I intron-derived ribozyme has different affinities for binding either the exogenous G or the ωG in the GBS between the two steps of trans splicing (*22*). We suspect that a similar change in affinity for binding ωG or ωGi in the GBS is important for the TIS reaction. In our proposed mechanism, ωG on the ribozyme interacts with the GBS in the first step of the reaction, although no P9.0 forms with this rPC ribozyme (Figure 9). After the first step of the TIS reaction, the insert is attached to the ribozyme and the formation of a P9.0 helix at this point could help position the ωGi of the insert in the GBS for the second step of the TIS reaction (see Figure 7, structure 3a). If the ribozyme forms a P9.0 in the first reaction step, however, this could inhibit the second step of the TIS reaction by preventing the ωGi from interacting with the GBS. The 6-fold decrease in product yields for the TIS reaction with the rPC-1 ribozyme (compared to the rPC ribozyme with no P9.0) suggests that P9.0 formation in the ribozyme does indeed inhibit the TIS reaction.

Comparison with Previous Studies. Other ribozymes have been shown to insert large intron sequences into RNA transcripts. For example, reverse splicing occurs when a group I intron splices back into an RNA transcript (*23*). Group II introns can act as mobile genetic elements by splicing out of an RNA transcript, and then reverse splicing into DNA (reviewed in ref *24*). These reactions differ from the TIS reaction in that the intron itself, and not an exogenous substrate, is being inserted.

A related insertion reaction using a twin hairpin ribozyme (*25*) differs from the TIS reaction by the type of ribozyme used and also the reaction mechanism. The twin ribozyme reaction is essentially a replacement reaction, using an engineered ribozyme to exchange one section of RNA with a larger section. This reaction is less efficient than TIS, taking 30 h and producing 30% product compared to the TIS reaction that is complete in 2 h with product yields of greater than 60%.

Other group I intron-derived ribozyme reactions have creatively exploited the functionality of ωG. A recombination reaction performed by Lehman (*8*) and a polymerization reaction by Burke (*6*, *7*) have an ωG in the ribozyme catalyzing a nucleophilic attack similar to the first step of the TIS reaction. In the recombination reaction, substrates AB and CD form CB and AD (*8*), but do not continue polymerizing because there is no ωG on the resultant products. In the polymerization reaction, substrate AB can form ABB, ABBB, and $A(B)_n$ (*3*) because the 3′-G on the end of B allows it to mimic the 3′-end of the ribozyme. The TIS reaction occurs in part because one substrate has a 3′-G (ωGi on the insert) and one does not (starting material). This setup allows for insertion, but not a repeating polymerization reaction.

Implications for the Mechanism. The TIS reaction is a new group I intron-derived ribozyme reaction, and it demonstrates unexpected molecular recognition interactions between the *P. carinii* group I intron-derived ribozyme and its RNA substrates. There is an added complexity in the TIS reaction compared to other group I intron-derived ribozyme reactions, since the TIS mechanism is dependent on the ribozyme binding two different substrates and forming the P1 helix twice (called P1 and P1i in this report). Therefore, a crucial conformational shift must occur between the first and second reaction steps (Figure 7, between structures 2a and 3a) to disrupt the first P1i helix (formed with the 9mer insert) and allow formation of the second P1 helix (formed with the 12mer starting material). Each substrate can base pair with six nucleotides of the ribozyme's IGS, suggesting that the relative binding strength of the substrates is critical for allowing the correct mechanism to occur. Another added complexity to this reaction is that the GBS appears to be used three different times in the reaction: binding ωG from the ribozyme for the first step, ωGi from the insert for the second step, and then ωG from the ribozyme again for the third step.

Implications for Potential Applications. The TIS reaction could potentially be useful as a biochemical tool, for example, to insert a sequence, perhaps even one containing a modified nucleotide or a marker, into a large RNA transcript. Large RNA transcripts are typically synthesized in vitro by T7 runoff transcription, so adding site specific modifications, while possible (*26*), is arduous. Using the TIS reaction, a small RNA insert could be synthesized (perhaps with a desired modification), targeted to an exact location, and inserted into a large transcript. The TIS reaction could also be used as an RNA repair agent. Deletion and frame shift mutations could potentially be repaired by insertion of an RNA sequence at a specific location. Figure 10 shows the sequence requirements for designing a TIS system. As made clear in the diagram, potentially any oligonucleotide sequence ending in guanosine can be inserted into potentially any target immediately downstream of a uridine.

The need for the ribozyme to bind two separate substrates, however, is an anticipated complication for the TIS reaction. This can easily be overcome, fortunately, by using the TIS ribozyme with the first intermediate already attached, as seen in structure 3a of Figure 7. This particular reaction does work (Figure 8, lane F), requires only one substrate (instead of two), and proceeds through two reaction steps (instead of three).

Implications for Cellular Function. That group I intron-derived ribozymes can catalyze the TIS reaction also has implications for the potential function of the intron once it splices out of its transcript in the cell. We have previously shown that *P. carinii* ribozymes can excise sections of RNA from transcripts, both in vitro (*9*) and artificially in vivo (*27*). Now we know that these same ribozymes can insert sections of RNA into transcripts in vitro (this report). Since both TES and TIS ribozymes are largely unaltered from spliced introns,

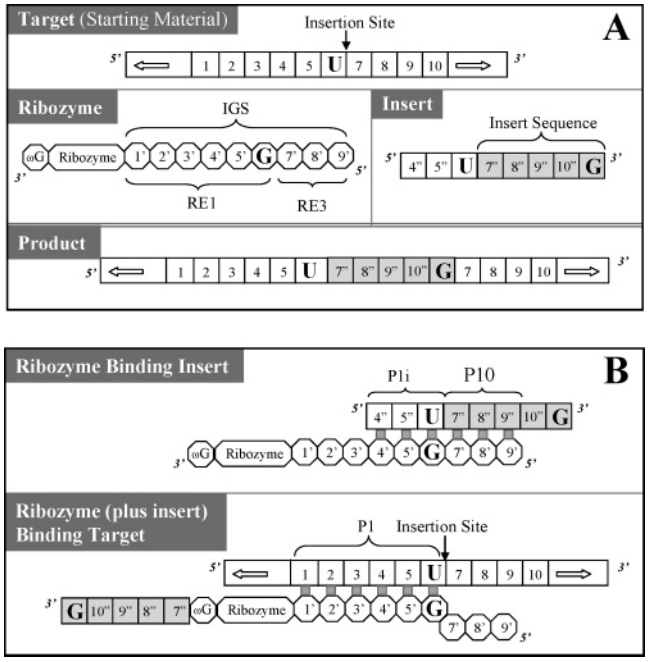

FIGURE 10: Generic diagram of (A) TIS substrates, ribozyme, and product and (B) ribozyme-bound substrates. Numbers are shown instead of nucleotides; the nucleotide positions of the IGS are shown as primed numbers, and the positions of the insert are shown as double-primed numbers. Single-nucleotide abbreviations are used where specific nucleotides are required for effective reactivity. In the ribozyme construct, only the IGS and ωG positions are specifically delineated. Base pairs are represented by broad vertical lines. The top and bottom structures in panel B are analogous to structures 2a and 3a in Figure 7, respectively. Figure 7 also shows formation of helix P9.0 via insert bases 9″ and 10″. While this pairing was possible in our model system, it is not expected to be required, and so it was omitted in this diagram.

this opens up the possibility that at one point in evolution (and perhaps even today), introns could be shuffling RNA information by the TIS and TES reactions.

ACKNOWLEDGMENT

We thank Byron Stapleton for technical support.

REFERENCES

1. Zaug, A. J., and Cech, T. R. (1986) The intervening sequence RNA of *Tetrahymena* is an enzyme, *Science 231*, 470−5.
2. Zaug, A. J., Davila-Aponte, J. A., and Cech, T. R. (1994) Catalysis of RNA Cleavage by a Ribozyme Derived from the Group I Intron of *Anabaena* Pre-tRNA^{Leu}, *Biochemistry 33*, 14935−47.
3. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) A *Pneumocystis carinii* Group I Intron Ribozyme That Does Not Require 2′ OH Groups on Its 5′ Exon Mimic for Binding to the Catalytic Core, *Biochemistry 36*, 15303−14.
4. Sullenger, B. A., and Cech, T. R. (1994) Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing, *Nature 371*, 619−22.
5. Alexander, R. C., Baum, D. A., and Testa, S. M. (2005) 5′ transcript replacement in vitro catalyzed by a group I intron-derived ribozyme, *Biochemistry 44*, 7796−804.
6. Chowrira, B. M., Berzal-Herranz, A., and Burke, J. M. (1993) Novel RNA polymerization reaction catalyzed by a group I ribozyme, *EMBO J. 12*, 3599−605.
7. Chowrira, B. M., Berzal-Herranz, A., and Burke, J. M. (1995) 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, *Nucleic Acids Res. 23*, 849−55.
8. Riley, C. A., and Lehman, N. (2003) Generalized RNA-directed recombination of RNA, *Chem. Biol. 10*, 1233−43.
9. Bell, M. A., Johnson, A. K., and Testa, S. M. (2002) Ribozyme-catalyzed excision of targeted sequences from within RNAs, *Biochemistry 41*, 15327−33.
10. Bell, M. A., Sinha, J., Johnson, A. K., and Testa, S. M. (2004) Enhancing the second step of the trans excision-splicing reaction of a group I ribozyme by exploiting P9.0 and P10 for intermolecular recognition, *Biochemistry 43*, 4323−31.
11. Baum, D. A., Sinha, J., and Testa, S. M. (2005) Molecular Recognition in a Trans Excision-Splicing Ribozyme: Non-Watson−Crick Base Pairs at the 5′ Splice Site and ωG at the 3′ Splice Site Can Play a Role in Determining the Binding Register of Reaction Substrates, *Biochemistry 44*, 1067−77.
12. Golden, B. L., and Cech, T. R. (1996) Conformational switches involved in orchestrating the successive steps of group I RNA splicing, *Biochemistry 35*, 3754−63.
13. Bass, B. L., and Cech, T. R. (1986) Ribozyme inhibitors: Deoxyguanosine and dideoxyguanosine are competitive inhibitors of self-splicing of the *Tetrahymena* ribosomal ribonucleic acid precursor, *Biochemistry 25*, 4473−7.
14. Shan, S. O., and Herschlag, D. (1999) Probing the role of metal ions in RNA catalysis: Kinetic and thermodynamic characterization of a metal ion interaction with the 2′-moiety of the guanosine nucleophile in the *Tetrahymena* group I ribozyme, *Biochemistry 38*, 10958−75.
15. Cech, T. R. (1990) Self-splicing of group I introns, *Annu. Rev. Biochem. 59*, 543−68.
16. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res. 15*, 8783−98.
17. Turnbull-Ross, A. D., Else, A. J., and Eperon, I. C. (1988) The dependence of splicing efficiency on the length of 3′ exon, *Nucleic Acids Res. 16*, 395−411.
18. Burke, J. M., Esherick, J. S., Burfeind, W. R., and King, J. L. (1990) A 3′ splice site-binding sequence in the catalytic core of a group I intron, *Nature 344*, 80−2.
19. Watanabe, S., Kawai, G., Muto, Y., Watanabe, K., Inoue, T., and Yokoyama, S. (1996) An RNA fragment consisting of the P7 and P9.0 stems and the 3′-terminal guanosine of the *Tetrahymena* group I intron, *Nucleic Acids Res. 24*, 1337−44.
20. Michel, F., Netter, P., Xu, M. Q., and Shub, D. A. (1990) Mechanism of 3′ splice site selection by the catalytic core of the sunY intron of bacteriophage T4: The role of a novel base-pairing interaction in group I introns, *Genes Dev. 4*, 777−88.
21. Been, M. D., and Perrotta, A. T. (1991) Group I intron self-splicing with adenosine: Evidence for a single nucleoside-binding site, *Science 252*, 434−7.
22. Zarrinkar, P. P., and Sullenger, B. A. (1998) Probing the Interplay between the Two Steps of Group I Intron Splicing: Competition of Exogenous Guanosine with ωG, *Biochemistry 37*, 18056−63.
23. Woodson, S. A., and Cech, T. R. (1989) Reverse Self-Splicing of the *Tetrahymena* Group I Intron: Implication for the Directionality of Splicing and for Intron Transposition, *Cell 57*, 335−45.
24. Lambowitz, A. M., and Zimmerly, S. (2004) Mobile Group II Introns, *Annu. Rev. Genet. 38*, 1−35.
25. Welz, R., Bossmann, K., Klug, C., Schmidt, C., Fritz, H. J., and Muller, S. (2003) Site-directed alteration of RNA sequence mediated by an engineered twin ribozyme, *Angew. Chem., Int. Ed. 42*, 2424−7.
26. Moore, M. J., and Sharp, P. A. (1992) Site-specific modification of pre-mRNA: The 2′-hydroxyl groups at the splice sites, *Science 256*, 992−7.
27. Baum, D. A., and Testa, S. M. (2005) In vivo excision of a single targeted nucleotide from an mRNA by a trans excision-splicing ribozyme, *RNA 11*, 897−905.

BI0504815