A *Pneumocystis carinii* Group I Intron-Derived Ribozyme Utilizes an Endogenous Guanosine as the First Reaction Step Nucleophile in the Trans Excision−Splicing Reaction†

P. Patrick Dotson II, Joy Sinha, and Stephen M. Testa*

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

Received October 12, 2007; Revised Manuscript Received February 20, 2008

ABSTRACT: In the trans excision−splicing reaction, a *Pneumocystis carinii* group I intron-derived ribozyme binds an RNA substrate, excises a specific internal segment, and ligates the flanking regions back together. This reaction can occur both in vitro and in vivo. In this report, the first of the two reaction steps was analyzed to distinguish between two reaction mechanisms: ribozyme-mediated hydrolysis and nucleotide-dependent intramolecular transesterification. We found that the 3′-terminal nucleotide of the ribozyme is the first-reaction step nucleophile. In addition, the 3′-half of the RNA substrate becomes covalently attached to the 3′-terminal nucleotide of the ribozyme during the reaction, both in vitro and in vivo. Results also show that the identity of the 3′-terminal nucleotide influences the rate of the intramolecular transesterification reaction, with guanosine being more effective than adenosine. Finally, expected products of the hydrolysis mechanism do not form during the reaction. These results are consistent with only the intramolecular transesterification mechanism. Unexpectedly, we also found that ribozyme constructs become truncated in vivo, probably through intramolecular 3′-hydrolysis (self-activation), to create functional 3′-terminal nucleotides.

The trans excision−splicing (TES)[1] reaction, catalyzed by a *Pneumocystis carinii* group I intron-derived ribozyme, was developed (*1, 2*) to bind an exogenous RNA substrate, excise a specific internal segment, and ligate the flanking regions back together. Overall, the TES reaction is similar to the intron-catalyzed self-splicing reaction (see Figure 1), except that in self-splicing the intron is removed from the RNA substrate in an autocatalytic, intramolecular reaction (*3*). Furthermore, the reactions appear to utilize the same molecular recognition components for orienting their substrates and intermediates (*2, 3*) during each of the two reaction steps (historically called 5′-cleavage and exon ligation).

One major difference between the two reactions, however, is that the 5′-cleavage reaction in self-splicing requires an exogenous guanosine nucleophile (*3*) while that for the TES reaction does not (*1*). It was previously postulated (*1*) that TES might instead be utilizing a site-specific hydrolysis reaction (Figure 2A). An alternative explanation is that an endogenous nucleotide, for example, the one at the 3′-terminus of the ribozyme, is the nucleophile in an intramolecular transesterification reaction (Figure 2B). There is precedence for both types of reactions occurring in group I introns and their derived ribozymes (*4–17*). In particular, the 3′-terminal guanosine of a group I intron from *Tetrahymena thermophila* (*T. thermophila*) can act as a nucleophile in a natural cellular cyclization reaction (*3–5*). Group II introns are also known to undergo both site-specific hydrolysis and endogenous nucleotide-mediated reactions, in this case using an internal adenosine (*18, 19*).

We now report that the *P. carinii* ribozyme utilizes its 3′-terminal nucleotide to catalyze the first reaction step of the TES reaction (Figure 3). Several lines of evidence, including the size and identity of TES reaction intermediates in vitro and in vivo, support this conclusion. We additionally show that the identity of the 3′-terminal nucleotide is important for in vitro reactivity. Surprisingly, more than one reactive 3′-terminal guanosine can be created through truncation of the ribozyme transcript in vivo, most likely through ribozyme-mediated hydrolysis.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation. DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and used without further purification. RNA oligonucleotides were obtained from Dharmacon Research, Inc. (Lafayette, CO), and deprotected using the manufacturer's standard protocol. Oligonucleotide concentrations were calculated from UV absorption measurements using a Beckman DU 650 spectrophotometer (Beckman Coulter, Inc., Fullerton, CA) and the computer program StCalc (S. M. Testa, University of Kentucky). Oligonucleotides were radiolabeled at the 5′-end via phosphorylation

† This work was supported by grants from the Department of Defense Breast Cancer Research Program (DAMD17-03-1-0329), the Kentucky Lung Cancer Research Program, and the Kentucky Research Challenge Trust Fund.

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

[1] Abbreviations: TES, trans excision−splicing; TIS, trans insertion−splicing; RE1, recognition element 1; RE2, recognition element 2; RE3, recognition element 3; GBS, guanosine-binding site; GFP, green fluorescent protein; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; IPTG, isopropyl thiogalactoside; EGS, extended guide sequence.

10.1021/bi7020525 CCC: $40.75
Published on Web 03/26/2008

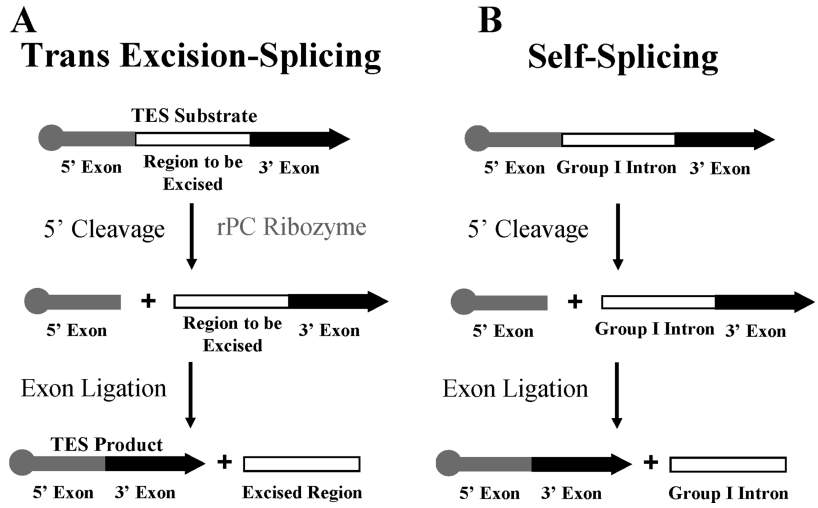

FIGURE 1: Trans excision−splicing and self-splicing reactions. (A) The trans excision−splicing substrate reacts with the rPC ribozyme through two consecutive transesterification reactions (5′-cleavage followed by exon ligation) to yield the TES product and the excised region. (B) The self-splicing reaction occurs through two consecutive transesterification reactions (5′-cleavage followed by exon ligation) to yield the ligated 5′- and 3′-exon sequences and the excised group I intron.

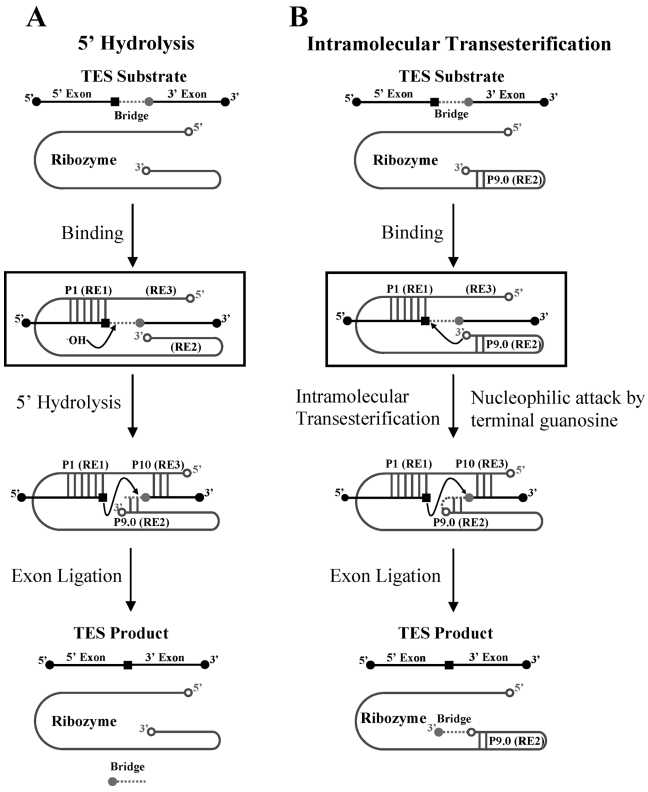

FIGURE 2: Comparison of potential TES reaction mechanisms. The *P. carinii* ribozyme is shown as a gray line, and the 5′- and 3′-exons are shown as black lines. The bridge region (the region to be excised) is shown as a dashed gray line. The black square within the 5′-exon represents a uridine, and the gray circle adjacent to the 3′-exon represents a guanosine. The white circle at the 3′-end of the ribozyme represents the 3′-terminal nucleotide of the ribozyme (ωG or ωA). The first step of the TES reaction (shown within boxes in the diagram) proceeds through either ribozyme-mediated 5′-hydrolysis (A) or ribozyme-mediated intramolecular transesterification (B).

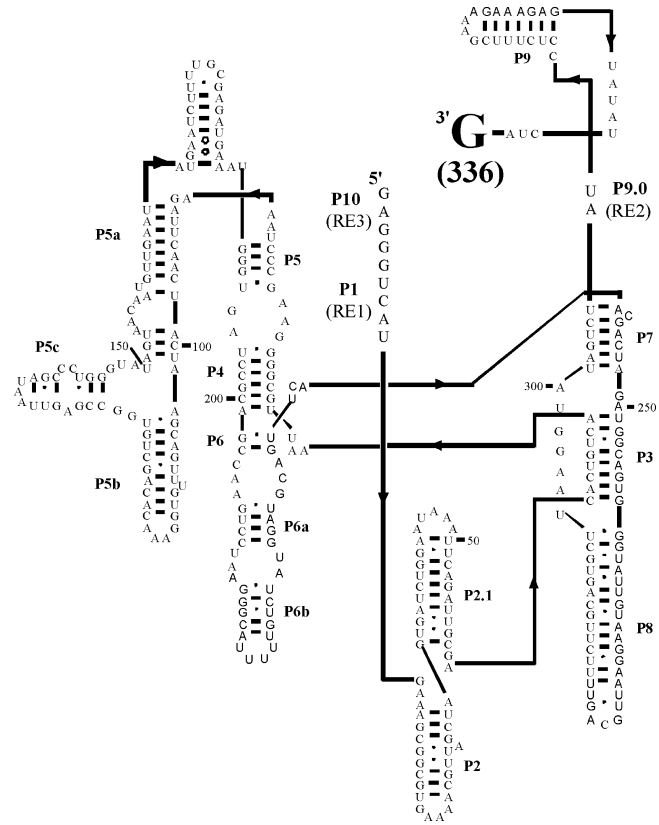

FIGURE 3: Sequence and proposed secondary structure of the *P. carinii* rPC ribozyme. Positions within the ribozyme that encompass RE1, RE2, and RE3 are shown. The 3′-terminal nucleotide (G336) of the rPC ribozyme is shown in bold.

of the 5′-terminal hydroxyl group with [γ-^{32}P]ATP (Amersham Pharmacia Biotech, Piscataway, NJ) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA), as previously described (*14*). Oligonucleotides were radiolabeled at the 3′-end via ligation of radiolabeled pCp (5′-*pCp) to the 3′-end of the oligonucleotide using T4 RNA ligase (New England Biolabs), as previously described (*17*). Note that this strategy results in 3′-end-radiolabeled substrates being one nucleotide longer than 5′-end-radiolabeled substrates.

Ribozyme Synthesis and Preparation. Four ribozymes were utilized in this study. Two DNA templates were prepared by simple linearization of the *P. carinii* ribozyme plasmid, PC. Linearization occurred in a 50 μL reaction mixture consisting of 16 μg of plasmid, 5 μL of 10× REACT 2 buffer (Invitrogen, Grand Island, NY), and 50 units of either *Xba*I (to form PC-X) or *Hin*dIII (to form PC-H) at 37 °C for 2 h.

The linearized plasmids were purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA). Two different DNA templates were prepared by PCR amplification of the PC plasmid using primers that would dictate the 3′-nucleotide of the resultant ribozyme. The forward primer for each construct was 5′CTCTAATACGACTCACTATAGAGGG3′, and the reverse primer (variable region underlined) was 5′TTAGATATACTCTTTCTTTCGAAAGAGG3′ for PC-ωA (*17*) and 5′CTAGATATACTCTTTCTTTCGAAAGAGG3′ for PC-ωG. The PCR mixture consisted of 100 ng of PC plasmid, 45 pmol of each primer, 1 unit of Pfu DNA polymerase (Stratagene, La Jolla, CA), 1× Pfu reaction buffer (Stratagene), and 0.5 μM dNTPs. The thermal cycler program was as follows: 95 °C for 30 s, followed by 40 cycles of 95 °C for 30 s, 55 °C for 45 s, and 68 °C for 120 s (360 s for the final cycle). PCR products were gel purified using a QIAquick gel extraction kit (Qiagen).

The run-off transcription mixture consisted of 1 μg of DNA template, 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, and 1 mM rNTP mix for 2 h in a total volume of 50 μL. The ribozymes were purified using a Qiagen plasmid midi kit (Qiagen), as previously described (*1*).

TES Reactions. TES reactions were conducted at 44 °C in H*X*Mg buffer, which consists of 50 mM HEPES (25 mM Na^+), 135 mM KCl, and *X* mM $MgCl_2$ (where *X* is either 0 or 10). Prior to each reaction, 200 nM ribozyme in 25 μL of H*X*Mg buffer was preannealed at 60 °C for 5 min and then slowly cooled to 44 °C. The reaction was initiated by adding 5 μL of a 8 nM solution of radiolabeled substrate, also in H*X*Mg buffer. The final concentration of the ribozyme was 166 nM, and that of the substrate was 1.3 nM. Time studies were conducted by the periodic removal of 3 μL aliquots, which were then quenched by adding an equal volume of 2× stop buffer (10 M urea, 3 mM EDTA, and 0.1× TBE). The products and substrates were denatured at 90 °C for 1 min and separated on a 12.5% denaturing polyacrylamide gel. Gels were dried under vacuum, and the bands were visualized on a Molecular Dynamics Storm 860 Phosphorimager. Observed rate constants, k_{obs}, were obtained from plots of the percent products over time (*14*). Because few time points could be taken, in some instances, before most reactions were complete, the k_{obs} values are most appropriately considered lower limits, and not precise values.

RT-PCR Amplification of TES Reaction Intermediates in Vitro. TES reactions were conducted as described above using the 20-mer substrate 5′AUGACUACUCUCGU*G*CUCUU3′. The reaction intermediates were RT-PCR amplified using an Access RT-PCR kit (Promega, Madison, WI) with primers 5′GAGGGTCATGAAAGCGGCGTG3′ and 5′GCACCGGTAAGAGCACGAGAG3′. Optimal amplifications occurred in a total reaction volume of 50 μL containing 2 μL of TES reaction mixture (166 nM ribozyme), 1 mM $MgSO_4$, 45 pmol of each primer, 0.2 mM dNTPs, 5 units of AMV reverse transcriptase, and 5 units of Tfl DNA polymerase. First-strand synthesis occurred at 45 °C for 45 min, followed by 2 min at 94 °C for deactivation of the AMV reverse transcriptase. The reaction mixtures were then subjected to 40 PCR cycles consisting of 94 °C for 30 s, 55 °C for 60 s, and 68 °C for 120 s (10 min for the final cycle). RT-PCR products were separated on a 2% agarose gel and excised from the gel matrix using the QIAquick gel extraction kit (Qiagen). The excised products were ligated into a pDrive cloning vector (Qiagen) using the manufacturer's recommended protocol and transformed into DH5α cells (Invitrogen). Colonies were picked and sequenced for identification (Davis Sequencing, Davis, CA).

In Vivo Characterization of TES Intermediates. Plasmids that contain both ribozyme and green fluorescent protein (GFP) templates were used to transform *Escherichia coli* strain JM109(DE3), which encodes an inducible T7 RNA polymerase, as described previously (*20*). Plasmids pQBI-Mut-GFP+p3x-RE3(5) (active ribozyme with mutant GFP) and pQBI-Mut-GFP+p3x-RE3(5)-ΔGBS (inactive ribozyme with mutant GFP) were used (*20*). After transformation and induction with IPTG (isopropyl thiogalactoside), total RNA (containing reaction intermediates) was isolated from 1.5 mL of cell culture using an Ambion RiboPure-Bacteria Kit (Ambion, Austin, TX). TES reaction intermediates were amplified via RT-PCR using the following primers: 5′GTGGAGAGGTAGAAAGCGGCGTG3′ and 5′CCATGCCATGTGTAATCCCAGCAGC3′, which amplify products containing both the ribozyme and 3′-end of the GFP transcript. RT-PCR, cloning, and sequencing were conducted as described above except that 1 μg of total RNA was the starting template.

RESULTS

TES Reaction Intermediates. If the substrate 5′AUGACU*G*CUC3′ is radiolabeled at the 3′-end (to form 5′AUGACU*G*CUCC3′), the size of the TES reaction intermediate will differ depending upon whether the first TES reaction step is hydrolysis [five nucleotides, 5′*G*CUCC3′ (Figure 2A)] or intramolecular transesterification [341 nucleotides (Figure 2B)]. Sequences shown in italics are excised in TES reactions, and those underlined are nucleotides added to the substrate during the 3′-end radiolabeling procedure. For the intramolecular transesterification mechanism, the intermediate is the result of the 3′-exon region of the substrate becoming covalently attached to the terminal nucleotide (G336) of the ribozyme (Figures 2 and 3).

TES reactions were conducted under previously optimized reaction conditions (*1*) using both 3′- and 5′-end-radiolabeled substrates for comparison (Figure 4, lanes G−J). The results (Figures 4 and 5) show that TES products are formed in approximately 70% yield (73 ± 1.8 and 69 ± 3.3%), regardless of which end of the substrate is radiolabeled. Because of the limitations of hand mixing relatively fast reactions, the only quantification that can be made of rate constants, k_{obs}, is a lower limit of >8 min^{-1} for the 3′-labeled substrate (Figure 5A) and >6 min^{-1} for the 5′-labeled substrate (Figure 5B) for the first step of the TES reaction. More importantly, however, when using a 3′-end-radiolabeled substrate, essentially all the radiolabeled intermediate is a high-molecular weight band (11 ± 1.6%) that corresponds to the predicted 341-nucleotide intermediate expected for the intramolecular transesterification mechanism. This is not consistent with the hydrolysis mechanism.

For confirmation, TES reactions were also conducted using the 3′-end-radiolabeled substrate 11-mer-dG, 5′AUGACU*d*-*G*CUCC3′, which is expected to prevent the second step of the TES reaction, resulting in the accumulation of the reaction

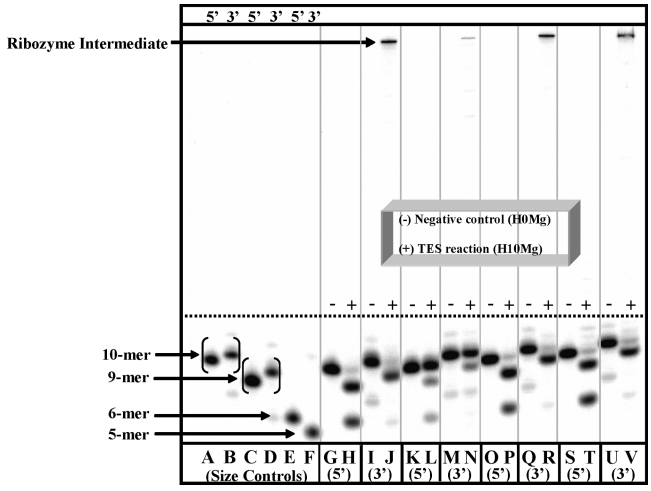

FIGURE 4: Polyacrylamide gel showing the substrates, intermediates, and products of TES reactions conducted with 166 nM ribozyme and 1.33 nM substrate at 44 °C for 15 min. Lanes A, C, and E contained 5′-end-radiolabeled 10-mer, 9-mer, and 6-mer size controls, respectively. Lanes B, D, and F contained 3′-end-radiolabeled 10-mer, 9-mer, and 5-mer size controls, respectively. Note that the 3′-end-radiolabeled size controls are one nucleotide larger than the 5′-end-radiolabeled size controls. Gel lanes that include a plus (+) sign show reactions conducted in H10Mg, and lanes that include a minus (−) sign show reactions conducted in H0Mg. Lanes G−J contained the normal rPC ribozyme with 5′-end-radiolabeled 10-mer (lanes G and H) or 3′-end-radiolabeled 10-mer (lanes I and J). Lanes K−N contained the rPC-ωA ribozyme with 5′-end-radiolabeled 10-mer (lanes K and L) or 3′-end-radiolabeled 10-mer (lanes M and N). Lanes O−R contained the rPC-ωG ribozyme with 5′-end-radiolabeled 10-mer (lanes O and P) or 3′-end-radiolabeled 10-mer (lanes Q and R). Lanes S−V contained the rPC-*Hin*dIII ribozyme with 5′-end-radiolabeled 10-mer (lanes S and T) or 3′-end-radiolabeled 10-mer (lanes U and V).

intermediate. Figure S1 (Supporting Information) shows that the high-molecular weight intermediate is now produced in yields of 87 ± 0.4% (Figure 5C). This is a further indication that the high-molecular weight band is the reaction intermediate and suggests that the TES reaction proceeds through the intramolecular transesterification mechanism. Again, the only quantification that can be made of the rate constant, k_{obs}, for the first step of this TES reaction is a lower limit of >6 min^{-1}.

For confirmation of the identity of TES reaction intermediates, RT-PCR was employed to amplify the high-molecular weight intermediates for subsequent sequencing (Figure 6). For these studies, a 20-mer substrate, $^{5'}$AUGACUACUCUCGU*G*CUCUU$^{3'}$, was used to facilitate RT-PCR amplification. Our results from sequencing five reaction products show that the TES reaction does produce the intermediate expected for the intramolecular transesterification reaction, whereby the 3′-end of the substrate ($^{5'}$ACUCUCGU*G*CUCUU$^{3'}$) attaches to the terminal nucleotide (G336) of the ribozyme (Figure S2 of the Supporting Information).

TES Reactions Using Ribozymes with 3′-Terminal Adenosines. A defining characteristic of the intramolecular transesterification mechanism is that the 3′-terminal guanosine of the ribozyme will act as the nucleophile. To test whether a guanosine is required at this position, a new ribozyme that contains a terminal adenosine was constructed (rPC-ωA) via run-off transcription of a PCR template. For comparison, this same strategy was used to make a ribozyme

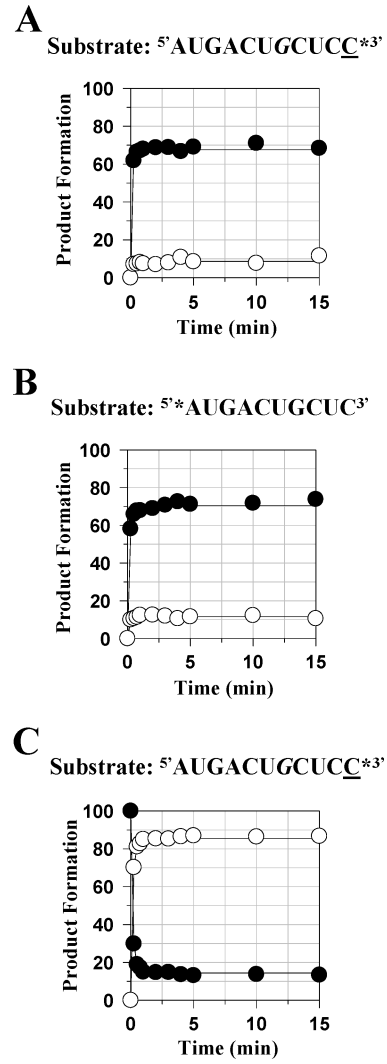

FIGURE 5: Graphical representations of TES reactions as a function of substrate. Reactions were conducted in duplicate with 166 nM rPC, 1.33 nM substrate, and H10Mg at 44 °C for 15 min. (A) TES reaction using 3′-end-radiolabeled substrate (AUGACU*G*CUCC). Data for the 10-mer TES product ($k_{obs} > 8$ min^{-1}) are represented by black circles, and data for the higher-molecular weight (ribozyme) intermediate are represented by white circles. (B) TES reaction using 5′-end-radiolabeled substrate (AUGACU*G*CUC). Data for the 9-mer TES product ($k_{obs} > 6$ min^{-1}) are represented by black circles, and data for the 6-mer intermediate are represented as white circles. (C) TES reaction using 3′-end-radiolabeled substrate (AUGACU*dG*CUCC). Data for the 11-mer-dG TES substrate are represented by black circles, and data for the higher-molecular weight (ribozyme) intermediate ($k_{obs} > 6$ min^{-1}) are represented by white circles. Each graph shows the average of two independent assays. Standard deviations for each time point are less than 15%.

construct that contains a terminal guanosine (rPC-ωG). The results show that the rPC-ωA ribozyme (Figure 4, lanes K−N) and the rPC-ωG ribozyme (Figure 4, lanes O−R) both produce approximately 80% TES product (80 ± 1.5% vs 76 ± 2.2%) (Figure 7). The terminal guanosine, however, results in a reaction that is faster ($k_{obs} = 2.6 \pm 0.09$ min^{-1}) than that with the terminal adenosine ($k_{obs} = 0.14 \pm 0.02$ min^{-1}). Nevertheless, the identity of the 3′-terminal end of the ribozyme does affect TES reaction kinetics, as expected only for the intramolecular transesterification mechanism.

That we obtain product using the rPC-ωA ribozyme can be explained with two scenarios. In the first scenario, a

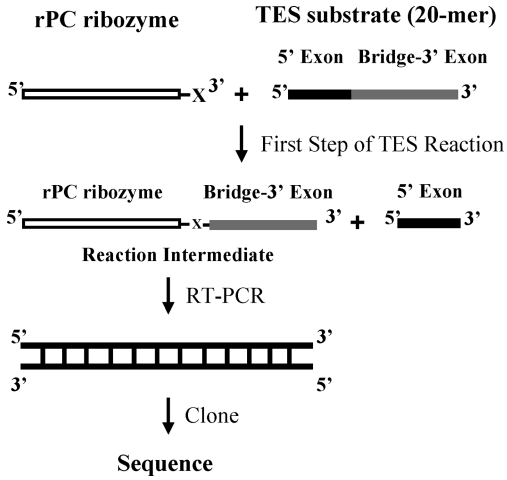

FIGURE 6: Experimental strategy for the amplification of TES reaction intermediates. TES reactions were conducted using the 20-mer substrate (5′AUGACUACUCUCGU*G*CUCUU3′). The TES reaction results in the covalent attachment of the bridge−3′-exon region (5′ACUCUCGU*G*CUCUU3′, denoted as a gray block) to the 3′-end of the ribozyme. The bridge−3′-exon region and the ribozyme are used as primer binding sites for RT-PCR amplification of the entire intermediate (but not the TES product, as the 3′-exon is not attached to the ribozyme). Isolated RT-PCR products were then cloned and sequenced.

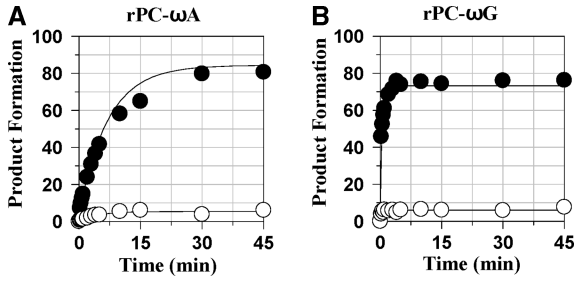

FIGURE 7: Graphical representations of TES reactions as a function of ribozyme construct. Reactions were conducted in duplicate with 166 nM ribozyme, 1.33 nM substrate, and H10Mg at 44 °C for 15 min. (A) TES reaction using the PCR-derived rPC-ωA ribozyme construct. Data for the 10-mer TES product (k_{obs} = 0.14 ± 0.02 min^{-1}) are represented by black circles, and data for the higher-molecular weight (ribozyme) intermediate are represented by white circles. (B) TES reaction using the PCR-derived rPC-ωG ribozyme construct. Data for the 10-mer TES product (k_{obs} = 2.6 ± 0.09 min^{-1}) are represented by black circles, and data for the higher-molecular weight (ribozyme) intermediate are represented by white circles. Each graph shows the average of two independent assays. Standard deviations for each time point are less than 15%.

functional guanosine is added to the 3′-end of the ribozyme during transcription (*21, 22*). In the second scenario, the adenosine itself is the nucleophile. There is a precedent for group I intron-derived ribozymes utilizing adenosines as intramolecular nucleophiles (*23*). To distinguish between these two scenarios, we employed the RT-PCR strategy described above to isolate and subsequently sequence the reaction intermediates that occur when using the rPC-ωA ribozyme. Results show that the terminal adenosine is the nucleophile in these reactions (Figure S2 of the Supporting Information). Apparently, the intramolecular transesterification reaction can be catalyzed by a terminal guanosine or adenosine. That guanosine is ~20-fold more effective (Figure 7) is likely related to the fact that guanosine is expected to have enhanced affinity for the guanosine binding site (GBS) of the ribozyme.

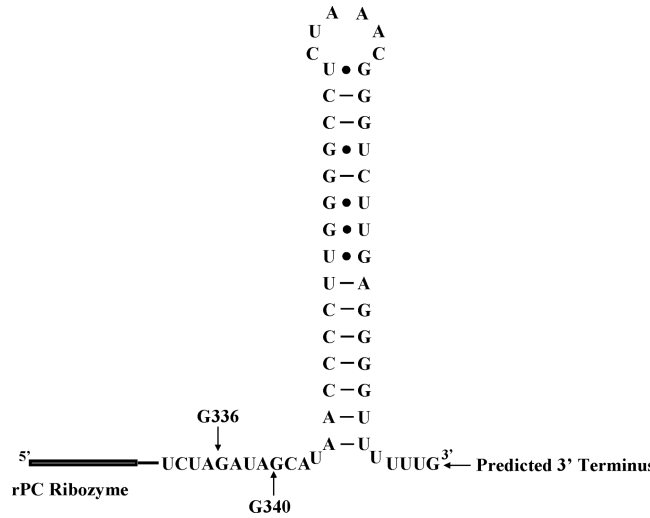

FIGURE 8: Sequence and predicted structure of the T7 terminator, which was appended onto the 3′-end of the rPC ribozyme for in vivo experiments. The guanosines (G336 and G340) found to act as ribozyme 3′-terminal guanosines are denoted.

Identification of TES Reaction Intermediates in Vivo. It was previously shown that a TES ribozyme could excise a central segment of a green fluorescent protein (GFP) transcript in *E. coli* (*20*). It was of interest, then, to determine if the intramolecular transesterification mechanism is occurring for TES reactions in vivo. Furthermore, because the GFP-specific ribozyme (GFP-rPC) is synthesized with a T7 RNA polymerase termination sequence (*20, 24, 25*) on its 3′-end (Figure 8), it is unknown which nucleotide is acting as the terminal nucleophile. Therefore, TES reactions were conducted in vivo (*20*), total RNA was isolated from the cells, and the TES reaction intermediates were amplified via RT-PCR and subsequently sequenced (scheme shown in Figure S3 of the Supporting Information).

Two plasmids were used in these studies, one that contains an active ribozyme construct and one purposely inactivated via elimination of the functionality of the GBS (*20*). As expected, no reaction intermediates could be RT-PCR amplified with the inactivated ribozyme construct. For the active ribozyme construct, 17 cloned RT-PCR products were isolated and sequenced. For all products, the 3′-end of the target substrate (the GFP transcript) was covalently attached to the 3′-end of the ribozyme, as expected for the intramolecular transesterification mechanism.

Although 11 products had the ribozyme covalently attached to the expected site on the target (U607; 5′-UACCUU$_{607}$-3′), six products had the ribozyme attached to an alternative site (U624; 5′-GCCCUU$_{624}$-3′). Apparently, the sequence specificity of this particular ribozyme construct reacting with this particular target region is not especially high.

Surprisingly, sequence information also showed that the T7 terminator sequence was missing from the 3′-end of the ribozyme in all 17 products, suggesting that it was somehow removed in vivo prior to the ribozyme becoming functional. In 16 sequences, the active ribozyme ended with the natural terminal nucleotide of the ribozyme (G336), and in one case, it ended with a guanosine four bases downstream (G340) (Figure 8). One possibility is that these terminal guanosines are being created through truncation of the ribozyme

transcript via ribozyme-mediated 3′-hydrolysis, which is well-known to occur in vitro (*6, 7, 9, 10*). If true, this would indicate that the ribozyme is catalyzing its own activation prior to catalyzing the TES reaction. Note that there were no sequences with a terminal adenosine, although the limited data set does not exclude the possibility.

3′-Hydrolysis Can Activate the Ribozyme in Vitro. To test whether TES ribozymes can generate 3′-terminal guanosines via 3′-hydrolysis, TES reactions were conducted in vitro using a ribozyme containing seven nucleotides added to its 3′-terminal end (downstream from G336). TES reactions were conducted using both 5′-end-radiolabeled ($^{5'}$AUGACU-*G*CUC$^{3'}$) and 3′-end-radiolabeled ($^{5'}$AUGACU*G*CUC$\underline{C}^{3'}$) substrates, and results show that the expected intermediates and products are formed (Figure 4, lanes S−V). Furthermore, sequence analysis of the TES ribozyme intermediates (five sequences) shows that the expected intramolecular transesterification intermediate is being formed. Apparently, the extra seven bases were removed in vitro, thus activating G336 for subsequent reactivity. The reaction is apparently autocatalytic, as no cofactors were added to the reaction mixture.

DISCUSSION

Intramolecular Transesterification Mechanism in Vitro. A series of experiments were conducted in vitro to determine the mechanism of the first reaction step of the TES reaction, focusing on whether the reaction is predominantly intramolecular transesterification or ribozyme-mediated hydrolysis (Figure 2). Several lines of evidence suggest that the primary mechanism is intramolecular transesterification (Figure 2B), and not hydrolysis (Figure 2A). First, 3′-end-radiolabeled substrates produce high-molecular weight TES intermediates (more than 300 nucleotides), which is expected for the intramolecular transesterification reaction (expect 341 nucleotides) and not the hydrolysis reaction (expect five nucleotides). Second, the 3′-half of the RNA substrate is covalently attached to the 3′-terminal nucleotide of the ribozyme during the intermediate step of the TES reaction. This indicates that the 3′-terminal nucleotide of the ribozyme is the nucleophile in the reaction. This intermediate is not expected to occur at any time during the hydrolysis mechanism. Third, it was shown that the catalytic rate of the reaction is affected by the identity of the 3′-terminal nucleotide, which is not expected for the hydrolysis mechanism.

An interesting consequence of these results is that they suggest the requirement for a conformational change within the ribozyme between the two steps of the TES reaction. Such a conformational change is well-documented in group I intron self-splicing reactions. These introns have a single guanosine binding site (GBS) binding in succession two different guanosines: first the exogenous guanosine cofactor and then the 3′-terminal guanosine of the intron (*26–28*). For TES, the 3′-terminal guanosine of the ribozyme would bind to the GBS first, followed by the guanosine at the 3′-end of the excised region of the substrate (Figure 2B).

Intramolecular Transesterification Mechanism in Vivo. To answer whether the intramolecular transesterification mechanism is also occurring in vivo, TES reactions were conducted in *E. coli* using GFP as the target substrate, and reaction intermediates were sequenced. Like that seen in vitro (the second line of evidence above), the 3′-half of the substrate was covalently attached to the 3′-terminal nucleotide of the ribozyme. Apparently, the intramolecular transesterification mechanism is occurring in vivo.

One hallmark of the intramolecular transesterification mechanism is that the 3′-terminal nucleotide of the ribozyme construct is the nucleophile, which presents a problem for in vivo reactions because of the need to include a transcription termination sequence at the 3′-end of the ribozyme (Figure 8). Our results show, however, that all or most of this termination sequence becomes removed from the ribozyme transcript before or during the TES reaction. This is fortuitous, as it results in the activation of a 3′-terminal nucleotide for subsequent nucleophilic attack, rather than burying it within the ribozyme transcript. One interesting possibility is that the ribozyme is folding in the presence of the termination sequence such that the junction between the ribozyme and the termination sequence is placed in the active site of the ribozyme, resulting in the cleavage of this phosphodiester bond through a well-characterized 3′-hydrolysis reaction (*6, 7, 9, 10*). If true, the ribozyme would be self-activating its 3′-terminal nucleophile, which greatly simplifies in vivo ribozyme design strategies.

Sequencing of TES reaction intermediates demonstrates that the ribozyme targets two similar sites on the GFP transcript, out of hundreds of possible sites, with the desired site being targeted 65% of the time. Previous studies using the *P. carinii* ribozyme have shown that stable base pairing between the substrate and the first four nucleotides of the ribozyme's RE1 region (to form a functional P1 helix) is primarily responsible for the molecular recognition components that bind the substrate to the ribozyme (*29*). It is not surprising, then, to find that both the major and minor target sites are able to form this four-nucleotide helix, which accounts for their similar reactivities. Nevertheless, this result does demonstrate the lack of high sequence specificity inherent in RNA catalytic reactions that rely on short base-paired regions to bind intermolecular substrates.

The 3′-Terminal Adenosine Ribozyme. Results demonstrate that the *P. carinii* ribozyme can utilize a 3′-terminal adenosine (A336) as the intramolecular transesterification nucleophile. Presumably, the terminal adenosine binds to the GBS of the ribozyme, which is known to preferentially bind guanosine (*26, 27*). Studies using a *T. thermophila* ribozyme, however, suggest that adenosine can bind the GBS with the help of non-GBS interactions (*23*). The most likely interaction is the formation of helix P9.0 (Figure 2B), which participates in the alignment of the 3′-splice site of group I introns into the GBS (*30–36*). It seems for the TES reaction, the role of P9.0 helix formation could be multifaceted in that it helps to align the terminal nucleotide of the ribozyme into the GBS prior to the intramolecular transesterification reaction and also aligns the substrate guanosine (the guanosine to be excised) into the GBS for the subsequent exon ligation reaction (Figure 2B).

Comparison with Previous Results. That the 3′-terminal guanosine of group I intron-derived ribozymes can be a nucleophile has previously been reported, although not in the context of the TES reaction. For example, the terminal guanosine has been implicated in ribonuclease, phosphotransferase, and phosphatase activities using a group I intron-derived ribozyme from *T. thermophila* (*7, 9*). It is interesting

to note that in these reactions the 3′-terminal guanosine of the ribozyme was obtained experimentally through site-specific 3′-hydrolysis, which consists of high-temperature ribozyme preincubations under alkaline conditions (*7, 9*). We found that production of a reactive 3′-terminal guanosine for the *P. carinii* ribozyme does not require this preincubation step. A 3′-terminal guanosine is also a nucleophile in a *Tetrahymena*-mediated reverse splicing reaction, in which the intron becomes embedded into cellular transcripts, as well as in in vivo cyclization reactions (*3–5, 37, 38*). Trans insertion−splicing ribozymes, which catalyze the insertion of one exogenous substrate into another exogenous substrate, also utilize a 3′-terminal guanosine (*17*). Other group I ribozymes have exploited their 3′-terminal guanosines for polymerization (*11*) and recombination (*16*) reactions. Apparently, 3′-terminal guanosines of group I intron-derived ribozymes can be involved in a whole host of natural and unnatural reactions.

Design Principles for the TES Reaction in Vivo. Synthesizing RNAs that contain specific terminal sequences in vivo, which is required for making active TES ribozyme constructs, is challenging because RNA transcripts typically end in 3′-termination sequences. Fortunately, our results indicate that, at least in *E. coli*, TES ribozymes appear to become truncated, which removes the transcription terminator, thus producing a functional ribozyme construct. The efficiency by which this 3′-truncation occurs, however, might be a limiting factor in the TES reaction. Therefore, enhancing the 3′ truncation reaction, thought to be a ribozyme-catalyzed hydrolysis reaction, could potentially enhance the yield of the TES reaction. To this end, it has been reported that P10 helix formation prior to the first step of self-splicing increases the yield of this type of 3′-hydrolysis reaction in vitro (*39*). Therefore, the yield of TES reactions in vivo could potentially be increased via addition of sequences after G336 that are complementary to RE3 of the ribozyme (see Figure 2).

It was also demonstrated that the TES ribozyme is not entirely specific for the intended site on the target. This lack of absolute sequence specificity also occurs with other group I intron-derived ribozyme reactions (*40*). Therefore, a variety of strategies have been devised to ameliorate this limitation, including the addition of an artificial substrate binding element [extended guide sequence (EGS)] to trans-splicing ribozymes, which could easily be adapted to TES ribozymes (*41–47*).

Implications for Native Group I Intron Function in Vivo. Native group I introns catalyze the self-splicing reaction. In this reaction, introns excise themselves (sometimes with the aid of proteins) from RNA transcripts. It has also been demonstrated that intact group I introns can be mobile elements by catalyzing a reverse-splicing reaction (*37, 38*). We now demonstrate that essentially intact excised group I introns can covalently attach themselves to central regions of RNA transcripts in trans. Moreover, these excised introns can further catalyze the removal of an internal segment from inside these RNA transcripts (through TES). Apparently, group I introns are multifaceted in vivo catalysts.

ACKNOWLEDGMENT

We thank the members of the Testa lab for helpful discussions and technical assistance, as well as two anonymous reviewers for insightful suggestions.

SUPPORTING INFORMATION AVAILABLE

Polyacrylamide gel showing the TES substrate, intermediate, and product of TES reactions (Figure S1), sequencing results for RT-PCR products isolated from TES reactions conducted with either the rPC or rPC-ωA ribozyme (Figure S2), and RT-PCR strategy for isolation of the TES intermediate in vivo (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bell, M. A., Johnson, A. J., and Testa, S. M. (2002) Ribozyme-catalyzed excision of target sequences from within RNAs. *Biochemistry 41*, 15327–15333.
2. Dotson, P. P., II, and Testa, S. M. (2006) Group I intron-derived ribozyme recombination reactions. *Recent Dev. Nucleic Acids Res. 2*, 307–324.
3. Cech, T. R. (1990) Self-splicing of group I introns. *Annu. Rev. Biochem. 59*, 543–568.
4. Brehm, S. L., and Cech, T. R. (1983) Fate of an intervening sequence ribonucleic acid: Excision and cyclization of the *Tetrahymena* ribosomal ribonucleic acid intervening sequence in vivo. *Biochemistry 22*, 2390–2397.
5. Zaug, A. J., Grabowski, P. J., and Cech, T. R. (1983) Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction. *Nature 301*, 578–583.
6. Zaug, A. J., Kent, J. R., and Cech, T. R. (1984) A labile phosphodiester bond at the ligation junction in a circular intervening sequence RNA. *Science 224*, 574–578.
7. Zaug, A. J., and Cech, T. R. (1986) The *Tetrahymena* intervening sequence ribonucleic acid enzyme is a phosphotransferase and an acid phosphatase. *Biochemistry 25*, 4478–4482.
8. Inoue, T., Sullivan, F. X., and Cech, T. R. (1986) New reactions of the ribosomal RNA precursor of *Tetrahymena* and the mechanism of self-splicing. *J. Mol. Biol. 189*, 143–165.
9. Zaug, A. J., and Cech, T. R. (1986) The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science 231*, 470–475.
10. Herschlag, D., and Cech, T. R. (1990) Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry 29*, 10159–10171.
11. Chowrira, B. M., Berzel-Herranz, A., and Burke, J. M. (1993) Novel RNA polymerization reaction catalyzed by a group I intron-derived ribozyme. *EMBO J. 12*, 3599–3605.
12. 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*, 14935–14947.
13. Mei, R., and Herschlag, D. (1996) Mechanistic investigations of a ribozyme derived from the *Tetrahymena* group I intron: Insights into catalysis and the second step of self-splicing. *Biochemistry 35*, 5796–5809.
14. 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*, 15303–15314.
15. Disney, M. D., Haidaris, C. G., and Turner, D. H. (2001) Recognition elements for 5′ exon substrate binding to the *Candida albicans* group I intron. *Biochemistry 40*, 6507–6519.
16. Riley, C. A., and Lehman, N. (2003) Generalized RNA-directed recombination of RNA. *Chem. Biol. 10*, 1233–1243.
17. Johnson, A. J., Sinha, J., and Testa, S. M. (2005) Trans insertion-splicing: Ribozyme-catalyzed insertion of targeted sequences into RNAs. *Biochemistry 44*, 10702–10710.
18. Daniels, D. L., Michels, W. J., and Pyle, A. M. (1996) Two competing pathways for self-splicing by group II introns: A quantitative analysis of *in vitro* reaction rates and products. *J. Mol. Biol. 256*, 31–49.
19. Fedorova, O., and Zingler, N. (2007) Group II introns: Structure, folding and splicing mechanism. *Biol. Chem. 388*, 665–678.
20. 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.
21. 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*, 8783–8798.

22. Tsai, C. H., and Dreher, T. W. (1993) *In vitro* transcription of RNAs with defined 3′ termini from PCR-generated templates. *BioTechniques 14*, 58–61.
23. Suh, E., and Waring, R. B. (1993) The conserved terminal guanosine of a group I intron can help prevent reopening of the ligated exons. *J. Mol. Biol. 232*, 375–385.
24. Studier, F. W., Rosenberg, A. H., and Dunn, J. J. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol. 185*, 60–89.
25. Fuerst, T. R., and Moss, B. (1989) Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells. *J. Mol. Biol. 206*, 333–348.
26. Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989) The guanosine binding site of the *Tetrahymena* ribozyme. *Nature 342*, 391–395.
27. Been, M. D., and Perrotta, A. T. (1991) Group I intron self-splicing with adenosine: Evidence for a single nucleoside-binding site. *Science 252*, 434–437.
28. Golden, B. L., and Cech, T. R. (1996) Conformational switches involved in orchestrating the successive steps of group I RNA splicing. *Biochemistry 35*, 3754–3763.
29. Johnson, A. K., Baum, D. A., Tye, J., Bell, M. A., and Testa, S. M. (2003) Molecular recognition properties of IGS-mediated reactions catalyzed by a *Pneumocystis carinii* group I intron. *Nucleic Acids Res. 31*, 1921–1934.
30. Burke, J. M. (1989) Selection of the 3′ splice site in group I introns. *FEBS Lett. 250*, 129–133.
31. Price, J. V., and Cech, T. R. (1988) Determinants the 3′ splice site for self-splicing of the *Tetrahymena* pre-rRNA. *Genes Dev. 2*, 1439–1447.
32. 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*, 80–82.
33. Michel, F., Netter, P., Xu, M. Q., and Shub, D. A. (1990) Mechanism of 3′ splice site selection by catalytic core of the *sun*Y intron of bacteriophage T4: The role of a novel base-pairing interaction in group I introns. *Genes Dev. 4*, 777–788.
34. Bevilacqua, P. C., Sugimoto, N., and Turner, D. H. (1996) A mechanistic framework for the second step of splicing catalyzed by the *Tetrahymena* ribozyme. *Biochemistry 35*, 648–658.
35. Russel, R., and Herschlag, D. (1999) Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme. *RNA 5*, 158–166.
36. Karbstein, K., and Herschlag, D. (2003) Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: Implications for RNA reorganization and function. *Proc. Natl. Acad. Sci. U.S.A. 100*, 2300–2305.
37. 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*, 335–345.
38. Roman, J., and Woodson, S. A. (1998) Integration of the *Tetrahymena* group I intron into bacterial rRNA by reverse splicing in vivo. *Proc. Natl. Acad. Sci. U.S.A. 95*, 2134–2139.
39. Ritchings, B. W., and Lewin, A. S. (1992) Mutational evidence for competition between the P1 and the P10 helices of a mitochondrial group I intron. *Nucleic Acids Res. 20*, 2349–2353.
40. Jones, J. T., Lee, S. W., and Sullenger, B. A. (1996) Tagging ribozyme reaction sites to follow trans-splicing in mammalian cells. *Nat. Med. 2*, 643–648.
41. Kohler, U., Ayre, B. G., Goodman, H. M., and Haseloff, J. (1999) Trans-splicing ribozymes for targeted gene delivery. *J. Mol. Biol. 285*, 1935–1950.
42. Watanabe, T., and Sullenger, B. A. (2000) Induction of wild-type p53 activity in human cancer cells by ribozymes that repair mutant p53 transcripts. *Proc. Natl. Acad. Sci. U.S.A. 97*, 8490–8494.
43. Ayre, B. G., Kohler, U., Turgeon, R., and Haseloff, J. (2002) Optimization of trans-splicing ribozyme efficiency and specificity by in vivo genetic selection. *Nucleic Acids Res. 30*, e141.
44. Byun, J., Lan, N., Long, M., and Sullenger, B. A. (2003) Efficient and specific repair of sickle β-globin RNA by trans-splicing ribozymes. *RNA 9*, 1254–1263.
45. Shin, K. S., Sullenger, B. A., and Lee, S. W. (2004) Ribozyme-mediated induction of apoptosis in human cancer cells by targeted repair of mutant p53 RNA. *Mol. Ther. 10*, 365–372.
46. Lundblad, E. W., Haugen, P., and Johansen, S. D. (2004) Trans-splicing of a mutated glycosylasparaginase mRNA sequence by a group I ribozyme. *FEBS J. 271*, 4932–4938.
47. Fiskaa, T., Lundblad, E. W., Henriksen, J. R., Johansen, S. D., and Einvik, C. (2006) RNA reprogramming of mannosidase mRNA sequences in vitro by myxomycete group IC1 and IE ribozymes. *FEBS J. 273*, 2789–2800.

BI7020525