5′ Transcript Replacement in Vitro Catalyzed by a Group I Intron-Derived Ribozyme†

Rashada C. Alexander, Dana A. Baum, and Stephen M. Testa*

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

Received December 27, 2004; Revised Manuscript Received March 15, 2005

ABSTRACT: Group I intron-derived ribozymes can perform a variety of catalytic reactions, including the replacement of the 3′ end of a mutant RNA transcript with a corrected version of the transcript [Sullenger, B. A., and Cech, T. R. (1994) *Nature 371,* 619−622]. We now demonstrate in vitro that a ribozyme, derived from a *Pneumocystis carinii* group I intron, can replace the 5′ end of a targeted exogenous RNA with an endogenous RNA. Our model system is a short synthetic mimic of a k-ras transcript, in which substitution mutations at codon 12 are implicated in a host of cancer types. In these experiments, yields of up to 70% were obtained. We analyzed the length dependence of two molecular contacts, P9.0 and P10, that occur between the ribozyme and the exogenous k-ras mimic, and determined that longer, and thus more stable, interactions result in higher product yields. Furthermore, the length of the loop region L1 can substantially influence the yield and the rate of the reaction. These results are a further demonstration that group I intron-derived ribozymes are quite malleable in terms of intermolecular recognition and catalysis, and that these properties can be exploited in developing potentially useful biochemical tools.

We have previously shown that group I intron-derived ribozymes can be engineered to excise sequence-specific central segments of RNA from within larger RNAs via the trans excision-splicing reaction (*1*, *2*). Through these studies, we have begun to explore the potential usefulness of exploiting the three forward reactions (5′ hydrolysis, G-addition, and exon ligation) and the three distinct molecular recognition interactions (P1, P9.0, and P10 helices) that group I ribozymes can utilize when interacting with exogenous substrates. The independence of the catalytic events, as well as the adaptability of the molecular recognition components, at least to some extent, can be controlled. This indicates that group I intron-derived ribozymes could be engineered to perform a much wider variety of feats than is known at present.

In fact, catalytic RNAs derived from group I introns have the ability to repair RNA transcripts by replacing the defective 3′ ends of transcripts with normal, "corrected" versions (trans splicing; shown in Figure 1A), thereby restoring the functionality of the resultant transcripts (*3*). Such mRNA repair activity can occur in vivo, in both prokaryotic and eukaryotic cells (*4*−*12*). Although 3′ replacement splicing was first reported in 1994 with a ribozyme derived from the *Tetrahymena* intron, an analogous ribozyme reaction that replaces the 5′ end of an RNA molecule has not yet been reported (outlined in Figure 1B).

In a 5′ replacement splicing scenario, RE1[1] would base pair with the new 5′ exon, which is covalently attached to the ribozyme, to form the P1 helix (Figure 1B). RE2 would base pair with the 5′ end of the exogenous substrate to form the P9.0 helix. RE3 would base pair with the 3′ end of the exogenous substrate to form the P10 helix. In addition, connecting the new 5′ exon and the remainder of the ribozyme is a hairpin loop (labeled L1). In group I intron-derived ribozymes developed so far, RE1 has been the primary molecular recognition interaction for binding exogenous transcripts. In our scheme, RE2 and RE3 would have to be exploited for binding exogenous transcripts. Previous work on TES ribozymes has revealed that P9.0 and P10 are normally weak interactions, yet they can be strengthened to enhance the binding of substrates (*2*), which gave us an indication that they might be similarly modifiable to serve as the primary target recognition factors.

While trans excision-splicing could aid in the removal of genetic insertion mutations (for example) at the RNA level, it is not always a viable strategy when dealing with substitution mutations. In these cases, one would prefer to replace the mutated region of an RNA transcript with a corrected version (in lieu of being able to site-specifically modify a single base in an RNA transcript). Since some genetic mutations are located near the 5′ end of the transcript, trans splicing (from now on termed 3′ replacement splicing) would not always be a practical strategy for mRNA repair. One particular example is the k-ras gene, which is mutated in approximately 25−50% of lung cancer cases (*13*). Codon 12, which is near the 5′ end of this transcript, is the most common mutation hotspot; its native glycine is required at this position to form a functional protein product (*14*−*17*). The location of this mutation renders 3′ replacement splicing a less than ideal choice for repair.

In this report, we describe a newly engineered group I intron-derived ribozyme that can effectively replace the 5′

† This work was supported by 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@uky.edu.

[1] Abbreviations: TES, trans excision-splicing; IGS, internal guide sequence; RE, recognition element.

10.1021/bi047284a CCC: $30.25
Published on Web 05/03/2005

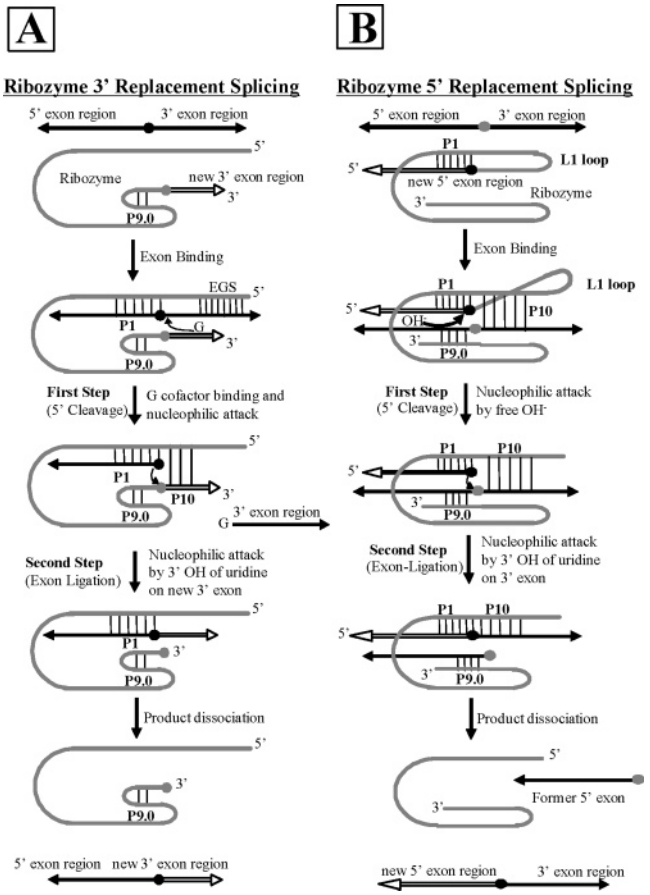

FIGURE 1: Schematic diagrams of (A) the 3′ replacement splicing (trans splicing) and (B) the 5′ replacement splicing reactions. The catalytic RNAs are represented by gray lines, the 5′ and 3′ exons within the exogenous substrates by black lines, and the replacement exons by white arrowheads. The black circle in the 5′ exon represents the uridine at the 5′ splice junction, and the gray circle adjacent to the 3′ exon represents a guanosine at the 3′ splice junction. P1, P9.0, and P10 are helices that result from recognition elements RE1, RE2, and RE3, respectively, of the ribozyme base pairing with the substrate. The 3′ replacement splicing strategies often utilize an extended guide sequence (EGS), which pairs with the 3′ exon region of the exogenous substrate. Note that the nucleophile in the first step of each reaction is not the same. Note that in the 5′ replacement splicing reaction, 5′ cleavage and then 5′ exon dissociation can occur prior to substrate binding, preventing a successful reaction.

end of an RNA with a "new" RNA in vitro, using as our model system a small synthetic k-ras mimic. In this model system under optimized conditions, 5′ replacement splicing generated up to 70% product. It was determined that P9.0 and P10 helix length, helices which form when the ribozyme binds to the exogenous RNA, is an important determinant of product yield. In addition, the length of L1 can substantially affect the rate and yield of the reaction. The implications of these results in terms of ribozyme engineering and potential applications are discussed.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation. DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). RNA oligonucleotides were purchased from Dharmacon, Inc. (Boulder, CO), and subsequently deprotected using the manufacturer's suggested protocol.

Designated RNA substrates were radiolabeled at the 3′ termini (*18*). First, cytidine monophosphate (Cp) was radiolabeled at the 5′ termini by incubating 10 μM Cp, 0.85 μM [γ-^{32}P]rATP (Amersham Pharmacia, Piscataway, NJ), 70 mM Tris-HCl (pH 7.5), 50 mM KCl, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mM EDTA, 0.1 mM rATP, 10% glycerol, and 10 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) in a reaction volume of 10 μL for 1.5 h at 37 °C. The kinase was then heat-inactivated by incubating the reaction mixture at 65 °C for 15 min. Next, 2 μM RNA substrate was radiolabeled with the newly created *pCp by incubating 0.28 μM *pCp, 50 mM Tris-HCl (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM rATP, 0.1 mM EDTA, 10% dimethyl sulfoxide (DMSO), and 20 units of T4 RNA ligase (New England Biolabs) at 4 °C for approximately 16 h. The products were gel purified as described previously (*19*).

Plasmid Construction and Synthesis. The *Pneumocystis carinii* ribozyme plasmid precursor, P-8/4x (*19*), served as our starting point for synthesizing the initial plasmid containing the 5′ replacement splicing ribozyme (designated PKA). A number of alterations were required to generate this plasmid (diagrammed in Figure 2). Briefly, an *Age*I site was engineered into the precursor plasmid in the P2 helix region of the ribozyme sequence. This unique restriction site was exploited, in conjunction with an existing *Sfi*I site upstream of the T7 promoter, to create an insertion site for a double-stranded oligonucleotide fragment containing a T7 promoter, the new 5′ exon, the L1 loop, and recognition elements RE1 and RE3. Ligation of this double-stranded oligonucleotide into the ribozyme precursor plasmid results in the creation of plasmid PKA. Finally, the RE2 region of the ribozyme sequence was modified specifically to target the k-ras mimic.

More specifically, a unique *Age*I site was created in the starting plasmid (P-8/4x) via site-directed mutagenesis using primers $^{5'}$CATGAAACCGGTGTGAAAACGTTAGGTAG-TG$^{3'}$ and $^{5'}$CACTACCTAACGTTTTCACACCGGTTTCA-TG$^{3'}$. The site-directed mutagenesis reaction was carried out as previously described (*1*). After initial denaturation for 30 s at 95 °C, the mixture underwent 15 cycles of 95 °C for 30 s, 55 °C for 2 min, and 68 °C for 6 min. The parental plasmid was then destroyed by digestion with 20 units of *Dpn*I (New England Biolabs) for 1 h at 37 °C. Transformation of DH5α competent cells (Invitrogen, Rockville, MD) was performed using 3 μL of this reaction mixture. The vector was purified using a QIAPrep miniprep kit (QIAGEN, Valencia, CA) and sequenced for confirmation.

The resulting plasmid (25 μg) was digested with *Age*I according to the manufacturer's recommended protocol (New England Biolabs) and purified using a QIAquick PCR purification kit (QIAGEN). The linear plasmid was then digested a second time with *Sfi*I according to the manufacturer's recommended protocol (New England Biolabs). The desired linear fragment was separated from undesired products on a 1.5% agarose gel, and the double-cut plasmid band was purified using a QIAquick gel extraction kit (QIAGEN). To prevent recircularization, the plasmid fragment was dephosphorylated according to the manufacturer's recommended protocol. Purification of the plasmid was conducted using a QIAquick PCR purification kit (QIAGEN).

Next, 300 pmol of each single-stranded oligonucleotide was phosphorylated according to the manufacturer's recom-

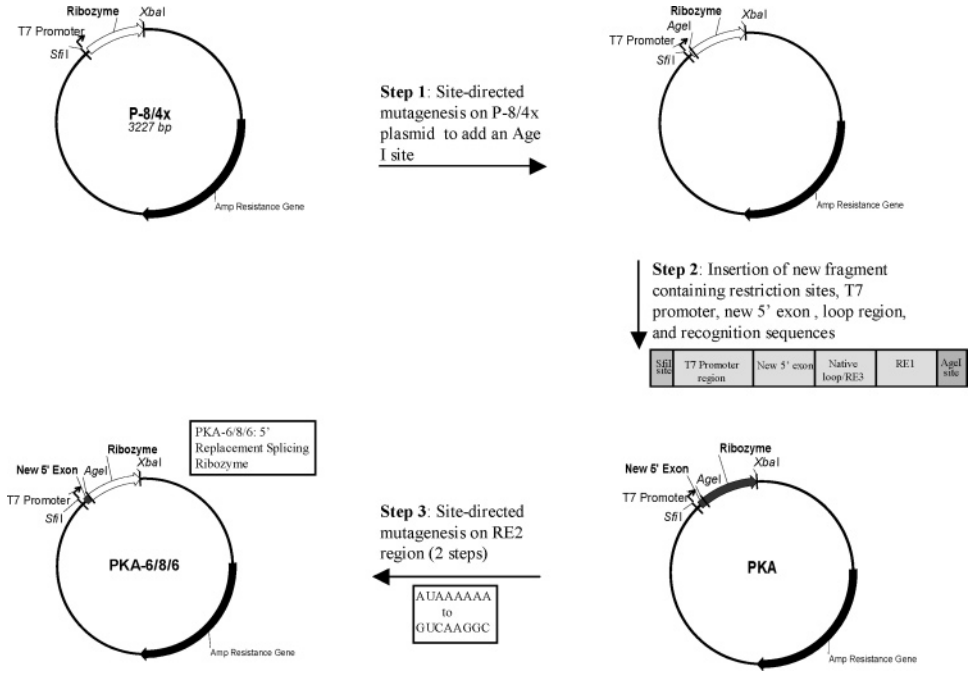

FIGURE 2: 5′ replacement splicing ribozyme plasmid construction. The starting plasmid containing the *P. carinii* ribozyme is designated P-8/4x (1), and the final plasmid containing the k-ras-specific 5′ replacement ribozyme is designated PKA-6/8/6. The digits following the base plasmid name are the lengths of the L1 loop, RE2 (to form P9.0), and RE3 (to form P10), respectively. Note that the relevant sequences are shown in Figure 4 and the construction methods are described in the text.

mended protocol (New England Biolabs). The resultant oligonucleotides were then purified using a QIAquick nucleotide removal kit (QIAGEN). Approximately 9 pmol of each oligonucleotide was mixed and heated at 50 °C for 5 min. The mix was then cooled to room temperature to allow the formation of the base-paired duplex. This duplex was then ligated into 4.3 pmol of the double-digested plasmid using 21 units of T4 DNA ligase (Promega, Madison, WI) in 2 μL of the manufacturer's supplied buffer. Ligation proceeded overnight at 22 °C, and 5 μL of the ligation mixture was used to transform DH5α cells. The resulting plasmids were purified using a QIAPrep miniprep kit (QIAGEN) and sequenced for confirmation.

Two more successive rounds of site-directed mutagenesis were performed with the following sets of primers to alter the RE2 region to target the k-ras mimic: $^{5'}$GCTTAAGGTATAGTCTGTCAAAAACCTCTTTCG$^{3'}$ and $^{5'}$CGAAAGAGGTTTTTGACAGACTATACCTTAAGC$^{3'}$, and $^{5'}$GGTATAGTCTGTCAAGGCCTCTTTCGAAAG$^{3'}$ and $^{5'}$CTTTCGAAAGAGGGCCTTGACAGACTATACC$^{3'}$. This new plasmid was designated PKA-6/8/6 to reflect the base plasmid (PKA) and the length of the L1 loop−RE2 region−RE3 region construct. The plasmid was sequenced by Davis Sequencing (Davis, CA) to confirm the changes. All further changes to L1, RE2, or RE3 were performed via site-directed mutagenesis as described above using primers specific for the desired changes, and were confirmed through sequencing.

Transcription. The 5′ replacement splicing ribozyme plasmids were linearized using *Xba*I (Invitrogen) and purified using a QIAquick PCR purification kit (QIAGEN) prior to runoff transcription. The transcription reaction was carried out as previously described (*19*). The purified products were 2-propanol precipitated, ethanol precipitated, and then reconstituted in sterile water and quantified using a Beckman DU-650 spectrophotometer.

5′ Replacement Splicing Reactions. Reactions were carried out in H*x*Mg buffer, composed of 50 mM HEPES (25 mM Na^+), 135 mM KCl, and *x* mM $MgCl_2$ at pH 7.5, where $x = 0-200$. Prior to each reaction, 100 nM ribozyme (0−400 nM ribozyme in concentration-dependent assays) was added to a solution of the appropriate buffer (total volume of 5 μL) and was preannealed at 60 °C for 5 min. In designated reactions, 0−300 μM guanosine monophosphate (GMP) was added to the ribozyme solution prior to preannealing. The ribozyme solution was allowed to cool to the reaction temperature listed in the figures. The reaction was initiated by adding a 1 μL aliquot of an 8 nM solution containing the radiolabeled substrate in the appropriate buffer to give a final volume of 6 μL. The reaction time was tested from 1 to 24 h. The reactions were terminated by adding an equal volume of stop buffer (10 M urea, 3 mM EDTA, and 0.1× TBE), followed by denaturation at 90 °C for 1 min. Products and reactants were separated on a 14% polyacrylamide−8 M urea gel. The gel was transferred to chromatography paper (Whatman 3MM CHR) and dried under vacuum. The bands were visualized and quantified on a Molecular Dynamics Storm 860 Phosphorimager. The results are the average of two independent assays. In all cases, the standard deviations are less than 10%. Observed rate constants were quantified by fitting the data from the time dependence studies to a simple exponential function.

To analyze whether any of the endogenous 5′ exon dissociates from the ribozyme before the second reaction step occurs, the 5′ replacement reaction was carried out as described above under optimized conditions, using the optimized ribozyme construct, except that 100 nM exogenous 5′ exon mimic, UUGACU, was added to the reaction mixture either before or after the preannealing step. Aliquots were periodically removed and added to an equal volume of stop buffer over the course of 5 h. GMP time studies were

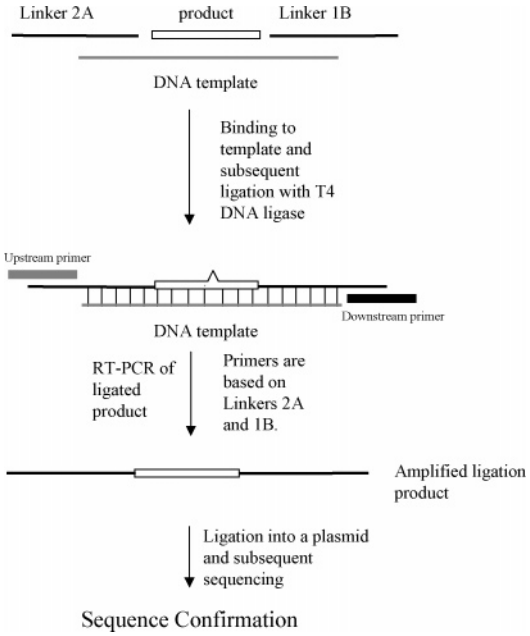

FIGURE 3: Sequencing protocol for confirmation of the 5′ replacement splicing product. The DNA linkers are colored black; the DNA template is colored gray, and the 5′ replacement splicing product is depicted as a white box. The system was designed so that the when the product binds to the template, a single-base bulge will occur at the new uridine in the product. In this way, we can be sure that our sequenced product is the result of 5′ replacement splicing. Once the linkers and 5′ replacement splicing product base pair with the template, they are ligated together. This ligated product is then amplified using RT-PCR, ligated into a receptive plasmid, cloned, and sequenced for product confirmation.

conducted under the optimized 5′ replacement splicing conditions, using the optimized ribozyme construct, except that 10 μM GMP was added to the ribozyme solution prior to preannealing or after 2.5 h for G-chase study. Aliquots were removed and added to an equal volume of stop buffer over the course of 5 h.

Sequencing of the 5′ Replacement Splicing Product. Sequencing was performed by modifying a previously reported method (*20*, *21*), as outlined in Figure 3. The 5′ replacement splicing reaction was carried out as described above, except that non-radiolabeled substrate was used. A 5 μL aliquot of this reaction mixture was added to 5 pmol of a template oligonucleotide with the sequence 5′GCACAGGGACCTGTATGTCAAGGCACTCGTCCCTTCCG3′, 10 pmol of linker 2A (5′TGCTCAGGCTCAAGGCTCGTCTAATCACAGTCGGAAGGGAC3′), and 3.3 pmol of linker 1B (5′GTCCCTGTGCTGACATGAATGACCGACTTGAGTGACCTGGCA3′). Note that linker 1B was previously phosphorylated using T4 polynucleotide kinase (New England Biolabs) according to the manufacturer's conditions. The oligonucleotides were ligated together using 6 units of T4 DNA ligase (Promega) in the manufacturer's recommended buffer for 18 h at 4 °C. The ligated products were purified from the reaction mixture using a QIAquick nucleotide removal kit (QIAGEN). Next, 5 μL of the reaction mixture was subjected to RT-PCR, using an Access RT-PCR kit (Promega). The RT-PCR primers were P1 (5′TGCCAGGTCACTCAAGTCGGTCATTCATGTCA3′) and P2 (5′TGCTCAGGCTCAAGGCTCGTCTAATCACAGT3′). RT-PCRs were conducted in the manufacturer's provided reaction buffer with 1 mM $MgCl_2$, 0.2 mM dNTPs, primers (1 μM each), 2.5 units of AMV reverse transcriptase, and 2.5 units of Tfl DNA polymerase in a total reaction volume of 25 μL. Reverse transcription proceeded at 45 °C for 45 min, and then PCR was carried out for 48 cycles of 94 °C for 30 s, 62.5 °C for 1 min, and 68 °C for 2 min. The amplified product was purified on a 1.5% agarose gel, and the DNA was isolated using a QIAquick gel extraction kit (QIAGEN). This DNA was then ligated into a pDrive cloning vector (QIAGEN) using the manufacturer's recommended conditions for 1.5 h at 4 °C. Finally, 5 μL of this ligation mixture was used to transform DH5α cells. The resultant vector was purified using a QIAPrep miniprep kit (QIAGEN), and the plasmid was sequenced by Davis Sequencing to provide confirmation of the 5′ replacement splicing product.

RESULTS

A *P. carinii* group I intron and its derived ribozymes have previously been shown to catalyze the self-splicing (*19*), suicide inhibition (*22*), and trans excision-splicing (TES) reactions (*1*, *2*). This intron served as the starting point for designing the 5′ replacement ribozyme, as it already catalyzes the necessary cleavage and ligation reactions.

Model System Design. The sequence of the 5′ replacement ribozyme and its substrate, which is an oligonucleotide mimic of a small portion of the k-ras gene, is shown in Figure 4. There were three main considerations in choosing the sequences for the model system. First, the site of the first catalytic step should occur at a wobble pair formed by a guanosine (gray) in the ribozyme and a uridine (blue) in the new 5′ replacement exon (Figure 4A). This interaction defines the first catalytic site and is denoted with an arrow. Second, a guanosine (red, also called ωG) in the transcript RNA defines the second catalytic site (also denoted with an arrow). ωG is defined by its positioning between helices P9.0 and P10, which form between the transcript RNA and ribozyme recognition elements RE2 and RE3, respectively. Third, the transcript guanosine chosen to be ωG will ultimately be replaced by a uridine in the reaction and so must yield a silent mutation in the codon when translated into the protein (Figure 4B). The first two considerations pertain to conserved and required elements of group I introns. The third ensures that the replacement exon will not alter the amino acid sequence when the exon is spliced into the transcript, thus maintaining proper coding function in the resultant transcript. The sequences of RE1, RE2, and RE3 are then chosen according to the transcript sequences flanking the two splice sites.

On the basis of visual inspection, the third nucleoside in codon 20 of the k-ras transcript was the most suitable guanosine to act as ωG. In the resultant 5′ replacement splicing reaction, the sequence upstream of (or 5′ to) codon 20 will be replaced, including codon 12, which is the mutation hotspot. The small model substrate used in these proof-of-principle studies is a short synthetic portion of the k-ras gene that is 15 nucleotides in length and contains codons 18−22 (Figure 4). In the ribozyme construct, the new 5′ exon is initially attached to the 5′ end of the ribozyme via a normal phosphodiester bond (to produce a 5′ exon−ribozyme conjugate) and contains the last base in codon 16 through the last base in codon 20 of the k-ras gene. Successful 5′ replacement splicing will result in a product

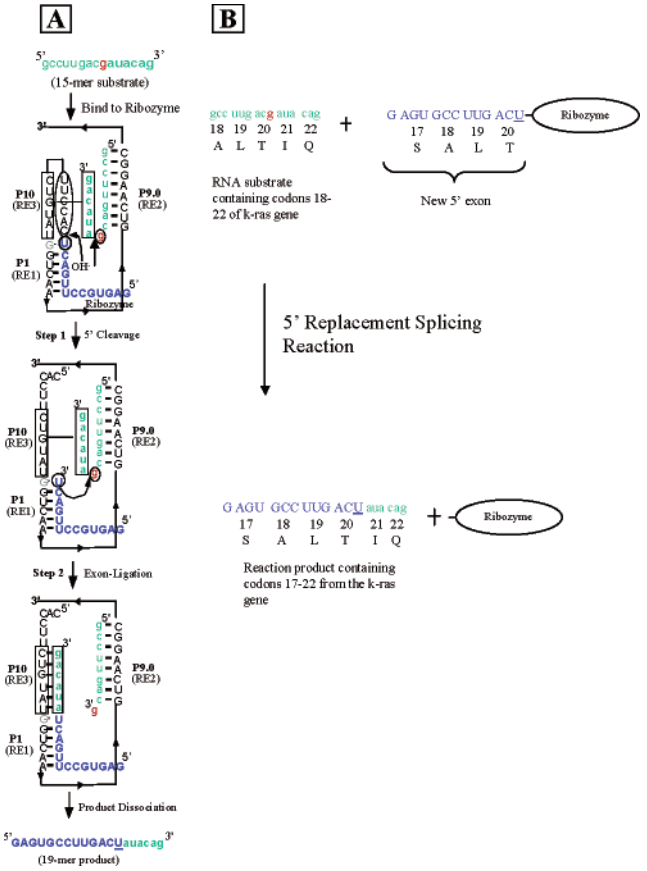

FIGURE 4: Schematic of the 5′ replacement splicing reaction. (A) The ribozyme is depicted in black uppercase lettering, the 15-mer k-ras mimic (the exogenous substrate) in green lowercase lettering, and the replacement exon in blue uppercase lettering. Regions to be ligated together are bold. The circled uridine in the replacement exon represents the 5′ splice site (for 5′ cleavage), and the red guanosine (ωG) in the substrate represents the 3′ splice site (for exon ligation). (B) A simplified version of the 5′ replacement splicing reaction showing the sequences and products grouped in codons. Note that codon 20 is "acg" in the substrate and "ACU" in the product. In addition, the product was designed to be four nucleotides longer than the substrate to aid in identification.

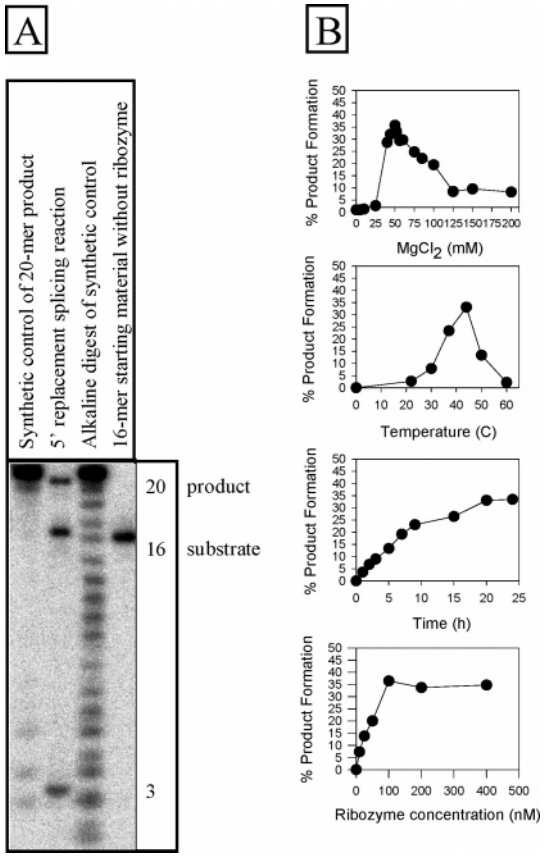

FIGURE 5: Representative gel and optimization data for the 5′ replacement splicing reaction using the initial ribozyme construct, rPKA-6/8/6. (A) Polyacrylamide gel showing the 5′ replacement splicing reaction. Reactions were carried out using 100 nM rPKA-6/8/6 ribozyme, 8 nM radiolabeled substrate, and 50 mM $MgCl_2$ at 44 °C for 20 h in H50Mg buffer composed of 50 mM HEPES (25 mM Na^+), 135 mM KCl, and 50 mM $MgCl_2$ (pH 7.5). The first lane contained the synthetic reaction product (20-mer), the second lane the 5′ replacement splicing reaction, the third lane an alkaline digest of the synthetic product, and the fourth lane the substrate (16-mer) under reaction conditions with no added ribozyme. Note that the 3′ radiolabeling procedure adds a cytosine to the 3′ end of both the substrate and the product, which results in gel bands one nucleotide longer than that shown in Figure 4. The 3-mer band results from cryptic cleavage of the last two bases of the 3′ exon plus the cytosine added by the radiolabeling procedure. (B) Optimization of the 5′ replacement splicing reaction using the rPKA-6/8/6 ribozyme. The reactions were carried out under the optimized reaction conditions described above except for the listed changing variable. Each data point is the average of two independently run assays, and the standard deviation for every point is less than 10%.

19 nucleotides in length. In these studies, we used 3′ radiolabeled substrates to view and quantify both the substrate and product. The 3′ radiolabeling process adds an extra cytosine at the 3′ end of the substrate and product, to give 16- and 20-mers, respectively.

5′ Replacement Splicing Reaction. Our initial ribozyme construct contains the new 5′ exon, an L1 loop of six bases, a P1 helix of six bases, a P9.0 helix of eight bases, and a P10 helix of six bases (diagrammed in Figure 4A). The lengths of the recognition elements were designed on the basis of optimized lengths of these regions for the TES reaction (*2*). As shown in Figure 5A, the 5′ replacement splicing reaction occurs, giving an optimized percentage yield of 33 ± 2 (the average of four independently run assays) and an observed rate constant of 1.58×10^{-3} min^{-1}. Figure 5B shows the yield of the reaction as a function of $MgCl_2$ concentration, reaction temperature, reaction time, and 5′ exon−ribozyme conjugate concentration under otherwise optimized reaction conditions. The final optimized conditions are 50 mM $MgCl_2$, 44 °C, 20 h, and 100 nM ribozyme.

Note the appearance of a three-nucleotide side product (confirmed by running size controls, data not shown), which consists of the 3′ end of the substrate (as per its radiolabel), and which apparently results from misfolding of the ribozyme to form a cryptic 5′ cleavage site between the ribozyme and the substrate (or product). That the yield of 5′ replacement splicing product remains stable over time suggests that the majority of this side product originates from the substrate and not the 5′ replacement splicing product. Nevertheless, in all cases, the yield of this side product was taken into account when calculating the yield of the replacement splicing product.

Length Dependence of P9.0. The effect of P9.0 length on reaction yield was investigated under the optimized conditions described above. RE2 length was decreased in increments of two, resulting in P9.0 helices of 2, 4, 6, and 8 bp. The results (Figure 6A) demonstrate that product yield

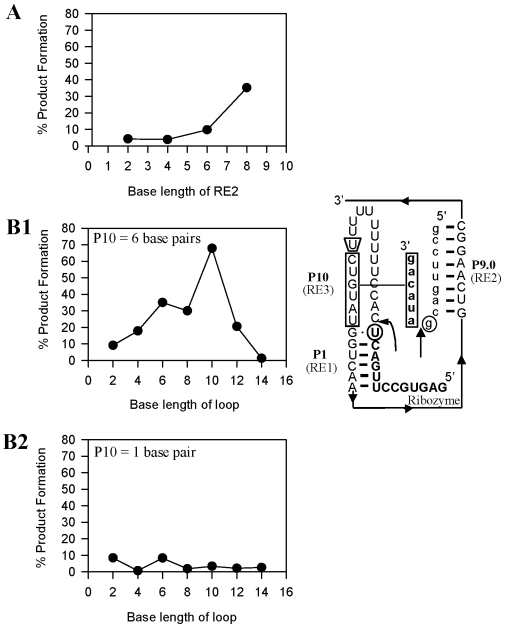

FIGURE 6: 5′ replacement splicing reaction as a function of RE2, RE3, and L1 loop length. The reactions were carried out with 100 nM ribozyme, 8 nM substrate, and 50 mM $MgCl_2$ at 44 °C for 20 h. All data points are the average of at least two independent assays, and the standard deviations for all points are less than 10%. Circles represent percent product formation in the graphs. (A) Effect of P9.0 length on product formation using ribozyme rPKA-6/*x*/6, where *x* ranges from 2 to 8 in increments of 2. The longest length tested (8 bases) is shown in the diagram at the right, and the shorter P9.0 helices have 2 bases removed, sequentially, from the 3′ end of RE2. (B1) Effect of L1 loop length (when RE3 is 6 bases long) on product formation using ribozyme rPKA-*x*/8/6, where *x* ranges from 2 to 14 in increments of 2. The longest length tested (14 bases) is shown in the diagram at the right, and the shorter L1 loops have two bases removed, sequentially, starting with boxed uridine in the loop. (B2) Effect of L1 loop length (when RE3 is 1 base long) on product formation using ribozyme rPKA-*x*/8/1, where *x* ranges from 2 to 14 in increments of 2. The longest length tested (14 bases) is shown in the diagram at the right, and the shorter L1 loops have two bases removed, sequentially, starting with boxed uridine in the loop. In this reaction, the five 5′ most bases of RE3 in the ribozyme were deleted, which is expected to give a continuous P10 helix of only 1 bp. Note that these experiments were designed such that simple base pair shifting within the ribozyme is not expected to be able to reconstitute stable lengths of P9.0, P10, and L1 other than that designed.

decreases as P9.0 length decreases. Therefore, just as in the TES reaction (*2*), elongating P9.0 is beneficial for the 5′ replacement splicing reaction. Since P9.0 forms between the ribozyme and the exogenous transcript, elongating this helix likely stabilizes this intermolecular interaction. Even though P9.0 could not be made longer than 8 bp in this model system, it is likely that even longer lengths would be even more beneficial.

Length Dependence of P10 and L1. The effect of P10 length on reaction yield was also investigated under the optimized conditions. RE3 and L1, however, are immediately adjacent in the ribozyme, so length changes in one could alter the length of the other. Using the system diagrammed in Figure 4A as an example, changing RE3 from six bases ($^{5'}$CUGUAU$^{3'}$) to four bases ($^{5'}$GUAU$^{3'}$) would still likely result in a 6 bp P10 helix, as the uridines in the loop would be available for base pairing. The only other length of P10 (other than 6 bp) that could possibly not form such a slipped P10 helix was 1 bp, although it is possible that P10 does not even form in this case. Nevertheless, using a ribozyme construct that does not have a stable P10 helix, regardless of whether it is 1 or 0 bp, can still give us information regarding the importance of stable P10 formation. Therefore, we investigated the effect of L1 length using P10 helix lengths of 1 (potentially) and 6 bp (Figure 6B) using loop lengths of 2, 4, 6, 8, 10, 12, and 14 nucleosides. Uridines were utilized to increase loop length because they are least likely to form stable intermolecular and intramolecular structures.

When the results in Figures 6B1 and 6B2 are compared, it is apparent that a longer and more stable P10 helix increases product yield. In the case with the unstable P10 helix (Figure 6B2), reaction percentages do not exceed 10%, regardless of loop length. It appears that any effect loop length could have on product formation is likely undetectable due to the detrimental effects of having such an unstable P10 helix. Apparently, stable P10 formation is an essential ingredient of the 5′ replacement splicing reaction.

Once P10 is increased to 6 bp, however, the yield of product changes, sometimes considerably, as a function of loop length. The general trend is that yields increase as P10 goes from 2 to 10 bp (with one exception), and then the yield decreases sharply thereafter. Evidently, a 10 bp L1 loop, combined with a 6 bp P10, provides an optimal level of product formation in this model system.

Optimized 5′ Replacement Splicing Construct. We analyzed whether the "optimized" ribozyme construct (with a 10 bp L1, an 8 bp P9.0, and a 6 bp P10) reacted differently to changes in $MgCl_2$ concentration, temperature, time, and ribozyme concentration in comparison to the initial ribozyme construct (with a 6 bp L1, an 8 bp P9.0, and a 6 bp P10). The reaction yield with this optimized construct is 73.6 ± 0.1% product (average of four independently run assays), which is approximately twice as much product as seen with the initial ribozyme construct (Figure 7). Perhaps not surprisingly, the optimized reaction conditions for the two ribozyme reactions were the same. Surprisingly, the reaction for the optimized construct is almost 6-fold faster (k_{obs} = 9.17×10^{-3} min^{-1}) than the initial construct (1.58×10^{-3} min^{-1}). As the only difference in the two constructs is the length of loop L1, it appears that the length of L1 is an important factor in defining the dynamics of the 5′ replacement splicing reaction. Note that the yield of the three-nucleotide side product is greatly reduced when using the optimized ribozyme construct in the reaction (compare Figures 5A and 7A).

The 5′ Exon Can Dissociate before the Substrate Binds. It is entirely possible that the first step of 5′ replacement splicing can occur without the ribozyme binding to its transcript target. This has only a negative consequence on the reaction, however, if the endogenous 5′ exon then dissociates from the ribozyme. To test whether such dissociation occurs, the 5′ replacement splicing reaction was carried out as described above except that an equimolar (relative to the ribozyme concentration) amount of non-radiolabeled 5′ exon competitor was added. The mimic is six nucleotides in length, and the endogenous 5′ exon is 13 nucleotides in length. If the endogenous 5′ exon does dissociate between the two reaction steps, then the competitor could bind and form a P1 helix, and subsequently be used in the 5′ replacement splicing reaction. The competition

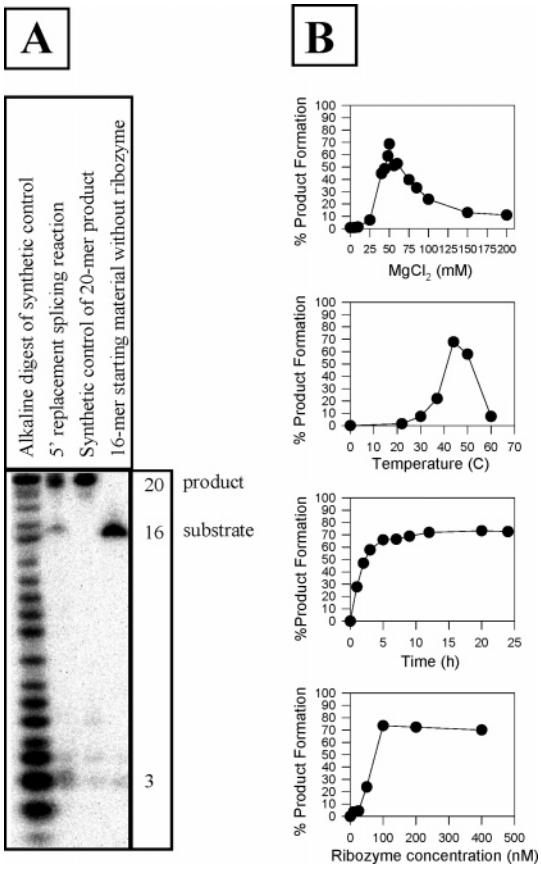

FIGURE 7: Representative gel and optimization data for the 5′ replacement splicing reaction using the optimized ribozyme construct, rPKA-10/8/6. (A) Polyacrylamide gel showing the 5′ replacement splicing reaction. Reactions were carried out using 100 nM rPKA-10/8/6 ribozyme, 8 nM radiolabeled substrate, and 50 mM $MgCl_2$ at 44 °C for 20 h. The first lane shows an alkaline digest of the synthetic product, the second lane the 5′ replacement splicing reaction, the third lane the synthetic reaction product (20-mer), and the fourth lane the substrate (16-mer) under reaction conditions with no added ribozyme. Note that the 3′ radiolabeling procedure adds a cytosine to the 3′ end of both the substrate and the product, which results in gel bands one nucleotide longer than those shown in Figure 4. The 3-mer band results from cryptic cleavage of the last two bases of the 3′ exon plus the cytosine added by the radiolabeling procedure. (B) Optimization of the 5′ replacement splicing reaction using the rPKA-10/8/6 ribozyme. The reactions were run under the optimized reaction conditions described above except for the listed changing variable. Each data point is the average of two independently run assays, and the standard deviation for every point is less than 10%.

product (12 nucleotides) could then be distinguished from the product (19 nucleotides) on the basis of size. The results show that 25% of the substrate reacts to form this competition product, indicating that the endogenous 5′ exons can dissociate before the substrate binds to the ribozyme. Furthermore, the amount of the replacement splicing product decreases by approximately 20% upon addition of the competitor, which indicates that as much as 20% of the replacement splicing product forms when the dissociated 5′ exon rebinds the ribozyme. Note that the results are the same regardless of whether the competitor is added prior to or after ribozyme annealing.

Sequencing Confirmation of the 5′ Replacement Splicing Product. A sequencing protocol used for small RNAs was adapted (*20*, *21*) to confirm the sequence of the 5′ replacement splicing product. In this protocol (Figure 3), linker

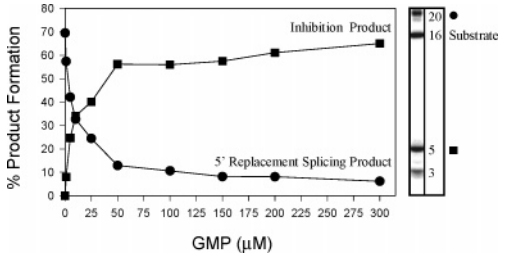

FIGURE 8: Dependence of the 5′ replacement splicing reaction on GMP concentration. The graph displays 5′ replacement splicing product formation (●) and inhibition product formation (■) with the optimized 5′ replacement splicing construct, rPKA-10/8/6, when increasing amounts of GMP are added to the reaction. The reactions were carried out with 100 nM ribozyme, 8 nM substrate, and 50 mM $MgCl_2$ at 44 °C for 20 h. Each data point is the average of two independent assays, and standard deviations for each point are less than 10%. The representative gel lane of the reaction run with 10 μM GMP added during preannealing shows the 20-mer replacement splicing product, the 16-mer substrate, the 5-mer inhibition product, and the 3-mer cryptic hydrolysis product.

oligonucleotides are ligated to the reaction product using a partially complementary template. The template does not contain the complement to the new uridine that is expected in the reaction product. This ensures that when the resultant ligation reaction mixture is RT-PCR amplified, any products containing the new uridine must have originated from the expected reaction. After amplification, the gel-purified product was ligated into a plasmid, cloned, and sequenced for confirmation. The sequence that was obtained was that expected for the 5′ replacement splicing reaction product, thus confirming the 5′ replacement splicing reaction product.

GMP Dependence. The 5′ replacement splicing reaction was conducted using the optimized reaction conditions and the optimized ribozyme construct in the presence of varying concentrations of GMP. The results of this experiment were of interest as guanosine phosphates are required cofactors in the 3′ replacement splicing reaction, but not in the 5′ replacement splicing reaction. As seen in Figure 8, concentrations as low as 50 μM GMP result in a loss of the expected 5′ replacement splicing product, while simultaneously resulting in the formation of a new inhibition product. This five-nucleotide inhibition product (determined using size controls, data not shown) contains the 3′ end of the substrate (as per its radiolabel), and is likely the result of a 5′ cleavage reaction (i.e., the G-addition reaction) occurring at a cryptic site on the 3′ end of the substrate (or product). Both types of 5′ cleavage reactions (G-addition and ribozyme-mediated hydrolysis) are well-known to occur at cryptic 5′ splice sites (*23*−*27*). That the yield of 5′ replacement splicing product remains stable over time when 10 μM GMP is added during annealing or after 2.5 h (data not shown) suggests that the majority of this inhibition product originates from the substrate and not the 5′ replacement splicing product. Apparently, the presence of GMP inhibits the 5′ replacement splicing reaction because it results in a competing side reaction.

DISCUSSION

In this report, we describe a newly engineered *P. carinii* group I intron-derived ribozyme that can effectively (70% yield) replace the 5′ end of a RNA with a different sequence in vitro, using as our model system a small synthetic k-ras

mimic. This is the first report of a group I intron-derived ribozyme specifically replacing the 5′ end of an RNA with a new RNA.

In general, elongating RE2 and RE3 of the ribozyme, which form helices P9.0 and P10 when bound to the exogenous RNA substrate, is beneficial to the 5′ replacement splicing reaction. In fact, both appear to be required for the reaction as a decrease in length of either one of these interactions severely reduces product yield. This is perhaps not surprising as both elements are involved in target identification and define the 3′ splice site. The effect of elongating these molecular recognition elements could encompass thermodynamic enhancement of the binding between the ribozyme and its substrate and intermediates, as well as to aid in the preorganization of the catalytic intermediates, as suggested for P9.0 and P10 with a trans excision-splicing reaction (*2*). Surprisingly, the length of loop L1 is also an important determinant for the 5′ replacement splicing reaction. The initial and optimized ribozyme constructs used in this study differ at only the L1 loop length (6 and 10 nucleotides, respectively), yet this change enhances the product yield by 2-fold and the rate of the reaction by 5-fold. One possible explanation is that the 10 base loop region provides ample "maneuvering room" for P1 and P10 formation to occur. A larger loop, however, might sterically deter the correct, most active folded state of the ribozyme−substrate complex. Similar results regarding P10 helix length (*5*) and L1 loop length (*28*) have been demonstrated in vivo using a *Tetrahymena* ribozyme in a 3′ replacement splicing reaction.

Note that the observed rate constant of the 5′ replacement splicing reaction ($k_{obs} = 0.009$ min^{-1}) is considerably lower than those of other *P. carinii* ribozyme-catalyzed reactions. For example, under individually optimized reaction conditions, the rate of self-splicing is 0.9 min^{-1} (*19*), the rate of suicide inhibition is 0.081 min^{-1} (*22*), and the rate of trans excision-splicing is 0.05 min^{-1} (*1*). The reasons for this are unknown; however, it is expected that enhancing this rate will be important for conducting this reaction in the in vivo environment.

In hindsight, there have been at least four barriers toward the development of a 5′ replacement splicing ribozyme. First, native length P10 and P9.0 helices do not result in sufficiently stable binding of the exogenous substrate to the ribozyme for the reaction to occur. Second, while GMP is usually an essential cofactor in group I ribozyme reactions, relatively low concentrations of GMP inhibit 5′ replacement splicing product formation. Third, at least using our model system, the conditions for running this reaction are far from that normally associated with ribozyme-catalyzed reactions. Fourth, 5′ exon dissociation prior to substrate binding could inhibit product formation. For these reasons, in its current incarnation, it is not expected that the 5′ replacement splicing reaction would work very well in vivo. Therefore, substantial modifications to the ribozyme target system may be required. This could include further elongation of P9.0 and P10 for stabilization of the interactions between the substrate and the ribozyme, developing ribozymes with a weakened tendency to catalyze cryptic reactions (especially the G-addition reaction) which could be toxic to the cell, "evolving" low-salt reactive ribozyme constructs, and elongating P1 to inhibit 5′ exon dissociation. Nevertheless, this proof-of-principle study shows, for the first time, that it is possible for group I intron-derived ribozymes to catalyze the 5′ replacement splicing reaction.

Comparison with 3′ Replacement Splicing. In 3′ replacement splicing, the 3′ end of an RNA transcript is replaced with an RNA that is covalently attached to a ribozyme (*3*). In cases where the desired region to be replaced is at the 5′ end of a transcript, however, this strategy might not be ideal. For example, being able to directly target mutations that are proximal to the 5′ end of a transcript could be less problematic and more efficient than replacing the majority of the transcript with a 3′ replacement ribozyme. Conversely, we have shown that the first reaction step of 5′ replacement splicing can occur before the ribozyme construct binds the target, resulting in some 5′ exon dissociation. This potential limitation might have to be addressed, perhaps by enhancing the stability of the P1 helix, for effective reactivity in vivo. This is not a limitation of the 3′ replacement splicing reaction, however, as the first reaction step occurs only when the exogenous substrate binds.

Fundamentally, the 5′ replacement splicing and 3′ replacement splicing reactions are similar, with a few important differences. While 3′ replacement splicing ribozymes utilize P1 and to a lesser extent the EGS (extended guide sequence) to bind an exogenous transcript (Figure 1A), 5′ replacement splicing ribozymes utilize P9.0 and P10 (Figure 1B). In addition, each of these reactions has been studied using different ribozyme systems: 3′ replacement splicing utilizes a ribozyme derived from *Tetrahymena*, and 5′ replacement splicing utilizes a ribozyme derived from *P. carinii*. We have previously reported many fundamental differences between how these two ribozymes recognize their substrates and catalyze their reactions (*1*, *2*, *19*, *22*, *29*), and so many of the differences in this report could be ascribed to this fact. For example, the *P. carinii* ribozyme appears to be much better able to catalyze the 5′ hydrolysis reaction than the *Tetrahymena* ribozyme, and so the *P. carinii* ribozyme does not require a GMP cofactor. In addition, under the conditions that were analyzed, 3′ replacement splicing has a substantially faster rate of reaction in vitro (0.17 min^{-1} vs 0.009 min^{-1}) than 5′ replacement splicing (*30*).

Comparison with SmaRT. Another strategy for modifying transcripts is spliceosome-mediated RNA trans splicing (SmaRT) (*31*−*34*). This strategy has been exploited for conducting both 3′ and 5′ replacement splicing reactions in mammalian cells. The catalyst in SmaRT is the spliceosome, which is a protein−RNA nuclear complex. SmaRT also utilizes a pre-trans splicing molecule (PTM), which contains the new replacement exon. Because this method relies on spliceosomes, the splice sites are limited to intron−exon junctions. In comparison, 5′ replacement splicing does not require mammalian intron−exon junctions as splice sites, and does not require a host that contains a spliceosome.

Implications. These results are a further demonstration that group I intron-derived ribozymes are quite malleable in terms of intermolecular recognition and catalysis. These studies demonstrate that the recognition sequences responsible for defining the 3′ splice site (RE2 and RE3) can be used as the primary elements of target recognition, indicating greater flexibility in molecular recognition than previously realized. As a biochemical tool, 5′ replacement splicing is a corollary to 3′ replacement splicing, which has already proven to be

effective in repairing some genetically damaged RNA transcripts (*3*−*12*). In many specific cases, the act of repairing mutated transcripts, via either a replacement splicing method, SmaRT, or some other method, is an attractive strategy for disease amelioration relative to the wholesale degradation of such transcripts (via antisense or siRNA strategies), as it retains the function of the corrected transcript. Repairing the common k-ras mutation at codon 12 that leads to a predisposition for lung cancer is one such relevant example. Furthermore, such catalytic RNAs may also have utility as biochemical tools, for example, by the targeted in vivo induction of transcript recombination. As the array of catalytic RNA tools for transcript manipulation grows, and the knowledge of their in vitro and in vivo activity improves, the potential for exploiting such ribozymes to solve real world problems is becoming more and more accessible.

ACKNOWLEDGMENT

We thank members of the Testa lab for helpful discussions and Joy Sinha and Patrick Dotson for technical assistance.

REFERENCES

1. Bell, M. A., Johnson, A. K., and Testa, S. M. (2002) Ribozyme-Catalyzed Excision of Targeted Sequences from within RNAs, *Biochemistry 41*, 15327−15333.
2. 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*, 4323−4331.
3. Sullenger, B. A., and Cech, T. R. (1994) Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing, *Nature 371*, 619−622.
4. Lan, N., Howrey, R. P., Lee, S. W., Smith, C. A., and Sullenger, B. A. (1998) Ribozyme-mediated repair of sickle β-globin mRNAs in erythrocyte precursors, *Science 280*, 1593−1596.
5. 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.
6. 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*, 8490−8495.
7. Samarsky, D., Ferbeyre, G., and Bertrand, E. (2000) Expressing active ribozymes in cells, *Curr. Issues Mol. Biol. 2*, 87−93.
8. Ayre, B. G., Kohler, U., Goodman, H. M., and Haseloff, J. (1999) Design of highly specific cytotoxins by using trans-splicing ribozymes, *Proc. Natl. Acad. Sci. U.S.A. 96*, 3507−3512.
9. 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* (24), e141.
10. 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.
11. Kastanos, E., Hjiantoniou, E., and Phylactou, L. A. (2004) Restoration of protein synthesis in pancreatic cancer cells by trans-splicing ribozymes, *Biochem. Biophys. Res. Commun. 322*, 930−934.
12. Shin, K., Sullenger, B. A., and Lee, S. (2004) Ribozyme-Mediated Induction of Apoptosis in Human Cancer Cells by Targeted Repair of Mutant p53 RNA, *Mol. Ther. 10*, 365−372.
13. Jackson, E. L., Willis, N., Mercer, K., Bronson, R. T., Crowley, D., Montoya, R., Jachs, T., and Tuveson, D. A. (2001) Analysis of lung tumor initiation and progression using conditional expression of oncogenic k-ras, *Genes Dev. 15*, 3243−3248.
14. Chipperfield, R. G., Jones, S. S., Lo, K. M., and Weinberg, R. A. (1985) Activation of Ha-ras p21 Substitution, Deletion, and Insertion Mutations, *Mol. Cell. Biol. 5*, 1809−1813.
15. Barbacid, M. (1987) ras Genes, *Annu. Rev. Biochem. 56*, 779−827.
16. Ellis, C. A., and Clark, G. (2000) The importance of being k-ras, *Cell. Signalling 12*, 425−434.
17. Ahrendt, S. A., et al. (2001) Cigarette Smoking is Strongly Associated with Mutation of the K-ras Gene in Patients with Primary Adenocarcinoma of the Lung, *Cancer 92*, 1525−1530.
18. England, T. G., and Uhlenbeck, O. C. (1978) 3′-Terminal labeling of RNA with T4 RNA ligase, *Nature 275*, 560−561.
19. 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.
20. Djikeng, A., Shi, H., Tschudi, C., and Ullu, E. (2001) RNA interference in *Trypanosoma brucei*: Cloning of small interfering RNAs provides evidence for retroposon-derived 24−26-nucleotide RNAs, *RNA 7*, 1522−1530.
21. Virlon, B., Cheval, L., Buhler, J., Billon, E., Doucet, A., and Elalouf, J. (1999) Serial microanalysis of renal transcriptomes, *Proc. Natl. Acad. Sci. U.S.A. 96*, 15286−15291.
22. Testa, S. M., Gryaznov, S. M., and Turner, D. H. (1999) *In vitro* suicide inhibition of self-splicing of a group I intron from *Pneumocystis carinii* by an N3′ → P5′ phosphoramidate hexanucleotide, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2734−2739.
23. Doudna, J. A., Cormack, B. P., and Szostak, J. W. (1989) RNA structure, not sequence, determines the 5′ splice-site specificity of a group I intron, *Proc. Natl. Acad. Sci. U.S.A. 86*, 7402−7406.
24. Barfod, E. T., and Cech, T. R. (1989) The conserved U-G pair in the 5′ splice site duplex of a group I intron is required in the first but not the second step of self-splicing, *Mol. Cell. Biol. 9*, 3657−3666.
25. Herschlag, D. (1992) Evidence of processivity and two-step binding of the RNA substrate from studies of J1/2 mutants of the *Tetrahymena* ribozyme, *Biochemistry 31*, 1386−1399.
26. Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994) Dissection of the role of the conserved G·U pair in group I RNA self-splicing, *Biochemistry 33*, 13864−13879.
27. 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*, 1067−1077.
28. Hasegawa, S., Jackson, W. C., Tsien, R. Y., and Rao, J. (2003) Imaging *Tetrahymena* ribozyme splicing activity in single live mammalian cells, *Proc. Natl. Acad. Sci. U.S.A. 100*, 14892−14896.
29. Johnson, A. K., Baum, D. A., Tye, J., Bell, M., 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. Zarrinkar, P. P., and Sullenger, B. A. (1999) Optimizing the substrate specificity of a group I intron ribozyme, *Biochemistry 38*, 3426−3432.
31. Mansfield, S. G., Clark, R. H., Puttaraju, M., Kole, J., Cohn, J. A., Mitchell, L. G., and Garcia-Blanco, M. A. (2003) 5′ Exon replacement and repair by spliceosome-mediated RNA trans-splicing, *RNA 9*, 1290−1297.
32. Puttanjaru, M., Jamison, S. F., Mansfield, S. G., Garcia-Blanco, M. A., and Mitchell, L. G. (1999) Spliceosome-mediated RNA trans-splicing as a tool for gene therapy, *Nat. Biotechnol. 17*, 246−252.
33. Chao, H., Mansfield, S. G., Bartel, R. C., Hiriyanna, S., Mitchell, L. G., Garcia-Blanco, M. A., and Walsh, C. E. (2003) Phenotype correction of hemophilia A mice by spliceosome-mediated RNA trans-splicing, *Nat. Med. 9*, 1015−1016.
34. Garcia-Blanco, M. (2003) Mending the message, *Nat. Biotechnol. 21*, 1448−1449.

BI047284A