Ribozyme-Catalyzed Excision of Targeted Sequences from within RNAs†

Michael A. Bell, Ashley K. Johnson, and Stephen M. Testa*

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

Received August 27, 2002; Revised Manuscript Received October 23, 2002

ABSTRACT: We demonstrate that a group I intron-derived ribozyme from the opportunistic pathogen *Pneumocystis carinii* can bind an RNA in trans and excise from within it an internal segment, resulting in the splicing of the remaining ends of the RNA back together (the trans excision−splicing reaction). The reaction is intramolecular with regard to substrate. The ribozyme targets its substrate by base pairing with two or three noncontiguous regions on the substrate, and the reaction occurs through a nucleotide cofactor independent mechanism. The excised segment can be as long as 28 nucleotides, or more, and as little as one nucleotide. The potential usefulness of this reaction is demonstrated by engineering a ribozyme that excises the triplet-repeat expansion region from a truncated myotonic dystrophy protein kinase transcript mimic in vitro.

Since the discovery of catalytic RNA (*1*, *2*), numerous RNA catalysts have been developed for a wide variety of applications (*3*). For example, catalytic RNAs are being developed for detection protocols (*4*, *5*), for therapeutic intervention of diseases (*6*−*8*), and for use as biochemical tools (*9*, *10*). As we continue to exploit the steadily increasing knowledge base of RNA structure, folding, and catalysis, designing and applying novel and effective RNA catalysts is becoming more and more tractable.

Group I introns are catalytic RNAs (*1*) that can splice themselves out of RNA transcripts (Figure 1A). To position the 5′ and 3′ exons for catalysis, the intron encoded internal guide sequence (IGS;[1] RE1 and RE3 in Figure 1A) base pairs with the 5′ exon (to form the P1 helix) and the 3′ exon (to form the P10 helix). Physically eliminating the 5′ and/or 3′ exons from group I introns produces catalytic RNAs (ribozymes) that can base pair with exogenous exon mimics and catalyze reactions analogous to the individual steps of the self-splicing reaction (*12*−*15*). In fact, group I intron ribozymes that lack both the exons and the IGS can catalyze a complete splicing-like reaction in trans upon binding, through tertiary interactions, pseudoknot-shaped substrates that contain 5′ exon, IGS, and 3′ exon sequences (*16*). Note that in the present study the intron and ribozyme sequences that base pair with the exogenous substrate will be referred to as recognition elements (RE). Hence, RE1 base pairs with the 5′ exon to form the P1 helix, RE2 base pairs with the bridge region to form the P9.0 helix, and RE3 base pairs with the 3′ exon to form the P10 helix (Figure 1).

We theorized that the above ribozyme activities could be combined for the development of ribozymes that base pair with specific RNA sequences in trans and excise specific segments from within them, resulting in the splicing of the remaining ends of the RNA back together (Figure 1B). This trans excision−splicing (TES) reaction could potentially be useful as a biochemical tool and as a therapeutic strategy. For example, disease-causing insertion mutations that lead to altered mRNA reading frames could be removed, thus restoring the correct codon reading frame. We show here that a ribozyme, rP-8/4x (Figure 2A), from the opportunistic pathogen *Pneumocystis carinii* (*15*, *17*−*19*) can catalyze the TES reaction. This occurs without the aid of a nucleotide cofactor, which is customarily required for the first step of the self-splicing reaction (*4*). We show that the intermediates that form the complete reaction products do not dissociate from the ribozyme. Thus, the reaction is intramolecular with regard to substrate. In addition, as little as a single nucleotide can be excised, indicating a nonessential role of the P9.0 helix in this reaction.

For the TES reaction to be generally applicable, the ribozyme recognition elements have to be mutable such that these ribozymes can be directed to target non-native sequences. There were previous indications that this might be possible, as altering the sequence of portions of the IGS from similar group I intron ribozymes results in ribozymes that bind and react primarily with substrates that are complementary to the new IGS sequences (*6*, *13*, *16*, *20*−*25*). To test this, we re-engineered the recognition elements of the rP-8/4x ribozyme to direct it to excise the RNA triplet repeat region from a truncated myotonic dystrophy protein kinase (DMPK) transcript mimic in vitro. The RNA expansion (or elongation) of this triplet repeat region (above 15 repeats) directly correlates with the severity of this common form of adult-onset muscular dystrophy (*26*−*28*). We show that, in a sequence specific manner, this ribozyme does indeed excise the triplet expansion region from this DMPK model system. The implications of group I intron ribozymes being able to catalyze the TES reaction are discussed.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Preparation. DNA oligonucleotides were purchased from Integrated DNA Technolo-

† This work was aided by Grant 85-001-13-IRG from the American Cancer Society, by a Research Innovation Award from the Research Corporation, by the Kentucky Research Challenge Trust Fund, and by the Muscular Dystrophy Association.

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

[1] Abbreviations: TES, trans excision−splicing; IGS, internal guide sequence; DMPK, myotonic dystrophy protein kinase.

10.1021/bi0267386 CCC: $22.00
Published on Web 11/27/2002

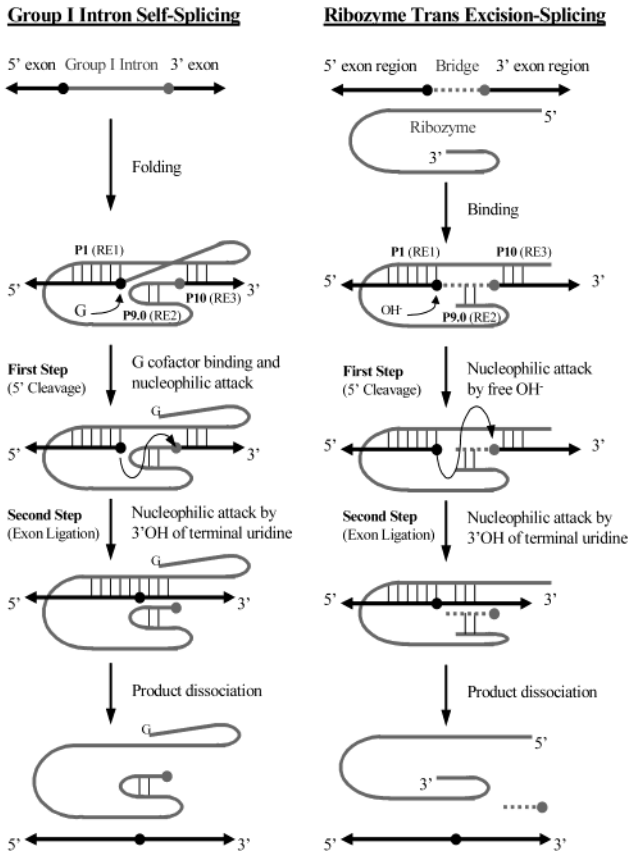

FIGURE 1: Group I intron self-splicing (A: left) and TES (B: right) reactions. The catalytic RNAs are represented by gray lines, the 5′ and 3′ exons and mimics are black lines, and the bridge (excised region) is a dotted line. The circle in the 5′ exon represents a uridine, and the circle adjacent to the 3′ exon represents a guanosine. Note that the intron inherently contains all the required activities for the TES reaction. RE1, RE2, and RE3 are the three recognition elements that the catalytic RNAs use to base pair with their substrates (RE1 + RE3 makes up the IGS). P1, P9.0, and P10 are helices that result from the recognition elements RE1, RE2, and RE3 (respectively) base pairing with the substrate. Note that these are very simplistic models. For example, it is possible that RE3 (and even RE2) does not bind the substrate until after the first reaction step (*11*).

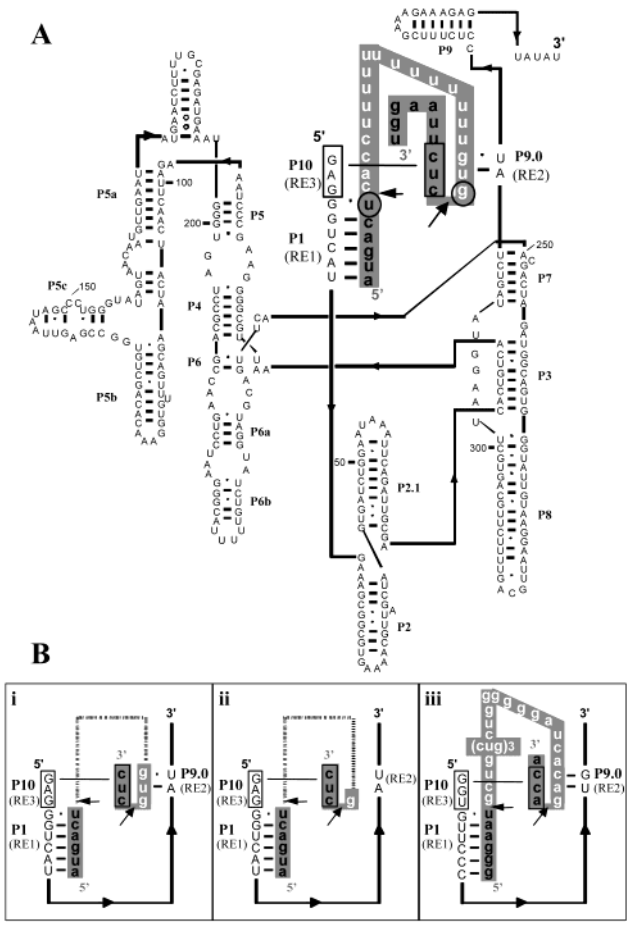

FIGURE 2: rP-8/4x and rP-8/4x-MD ribozymes base pairing with various substrates. (A) The *P. carinii* ribozyme, rP-8/4x (uppercase lettering), binding to the 36-mer substrate (lowercase lettering with a gray background). In the TES reaction, the bridging region (white lettering) is excised, and the 5′ and 3′ regions of the substrate (black lettering) are subsequently spliced together. Note that P1, P9.0, and P10 are helices that result from the recognition elements RE1, RE2, and RE3 base pairing with the substrate. The large bold arrows indicate the sites of catalysis for the first (left) and second (right) steps of the TES reaction. The 5′ uridine and 3′ guanosine that may be required are circled. The ribozyme bases are numbered according to that for the *P. carinii* intron (*15*). (B) Simplified diagrams of various substrates base pairing with various ribozymes. Only the recognition element sequences are shown for the ribozymes. The dashed line indicates a single normal phosphodiester bond between the adjoining sequences. (i) The 12-mer substrate binding to rP-8/4x. (ii) The 10-mer substrate binding to rP-8/4x. (iii) The 38-mer DMPK mimic binding to the rP8/4x-MD ribozyme.

gies (Coralville, IA) and were used without further purification. RNA oligonucleotides were purchased from Dharmacon Research Inc. (Boulder, CO) and deprotected following the manufacturer's protocol. Select RNAs were 5′ end radiolabeled, and the RNAs were either gel or TLC purified as previously described (*15*).

Plasmid Construction and Synthesis. The *P. carinii* ribozyme plasmid precursor, P-8/4x, was generated as previously described (*15*). The myotonic dystrophy-specific ribozyme plasmid precursor, P-8/4x-MD, was derived from the P-8/4x plasmid by site-directed mutagenesis. Briefly, three successive rounds of mutagenesis were performed to modify each of the three recognition elements using the following pairs of mutagenic primers (underlined bases represent altered recognition elements as compared to P-8/4x): $^{5'}$CACGCCGCTTTCGGGAACCTCTATAGTGAGTCG$^{3'}$ and $^{5'}$CGACTCACTATAGAGGTTCCCGAAAGCGGCGTG$^{3'}$ for RE1 formation, $^{5'}$GGTATAGTCTTGCCTCTTTCGAAAG$^{3'}$ and $^{5'}$CTTTCGAAAGAGGCAAGACTATACC$^{3'}$ for RE2 formation, and then $^{5'}$CGACTCACTATAGGTGTTCCCGAAAGCGGC$^{3'}$ and $^{5'}$GCCGCTTTCGGGAACACCTATAGTGAGTCG$^{3'}$ for RE3 formation. Each set of primers (15 pmol for each primer) was used in an amplification reaction comprising 25 ng P-8/4x parental plasmid, 2.5 units Pfu DNA polymerase (Stratagene; La Jolla, CA), and 0.5 μM dNTPs in a buffer consisting of 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$, 0.1% Triton X-100, and 0.1 mg/mL BSA (total reaction volume is 50 μL). After an initial denaturation for 30 s at 95 °C, the mixture was subjected to 15 cycles of 95 °C for 30 s, 50 °C for 2 min, and 68 °C for 6 min. Parental plasmid was then digested with 20 units *Dpn*I (Gibco BRL; Rockville, MD) in 5.7 μL of the manufacturer's supplied buffer for 2 h at 37 °C. A 3-μL aliquot of this mixture was then used to transform *Escherichia coli* DH5α competent cells (Gibco BRL). The vectors were purified using a QIAprep Spin

Miniprep Kit (QIAGEN; Valencia, CA). The resultant final plasmid, P-8/4x-MD, was sequenced for confirmation (ACGT, Inc; Northbrook, IL). The plasmids were linearized with *Xba*I, phenol/chloroform extracted, and ethanol precipitated in preparation for runoff transcription.

Transcription. Both rP-8/4x and rP-8/4x-MD were transcribed from their appropriate plasmid precursors as previously described for rP-8/4x (*15*). The resultant RNA was purified using QIAGEN-tip 100 anion-exchange columns. First, each column was equilibrated with 4.0 mL of buffer I (750 mM NaCl, 50 mM MOPS (pH 7.0), 15% ethanol, and 0.15% Triton X-100). Second, the transcription reactions were loaded onto the column, and the column was washed with 7.0 mL of buffer I. Third, the transcripts were eluted using 4.0 mL of buffer II (1.0 M NaCl, 50 mM MOPS (pH 7.0), and 15% ethanol). Following a 2-propanol and then an ethanol precipitation, the samples were dissolved in water and quantified using a Beckman UV−vis DU-650 spectrophotometer.

TES Reactions. Reactions were conducted in HxMg buffer consisting of 50 mM Hepes (25 mM Na^+), 135 mM KCl, and *x* mM $MgCl_2$ (listed in the figures) at pH 7.5. The TES reactions were optimized for the rP-8/4x and rP-8/4x-MD ribozymes over a $MgCl_2$ concentration range of 0−15 mM at 30, 37, and 44 °C. Prior to each reaction, 200 nM ribozyme (from 0.2−200 nM in the concentration-dependent assays) in 5.0 μL of the appropriate buffer was preannealed at 60 °C for 5 min and then allowed to slow-cool to the appropriate temperature. Reactions were initiated by adding 1.0 μL of an 8 nM solution of radiolabeled substrate. The substrate and ribozyme sequences and how they base pair are shown in Figure 2. In each case, the substrates were incubated in the appropriate buffer (listed in the figures). After 1 h (from 15 s to 90 min in the time-dependent assays), the reactions were terminated by adding an equal volume of stop buffer (10 M urea, 3 mM EDTA, and 0.1X TBE). The products and reactants were denatured for 1 min at 90 °C and then separated on a 12% polyacrylamide/8 M urea gel. The gel was transferred to chromatography paper (Whatman 3MM CHR) and dried under vacuum. The bands were visualized and quantified on a Molecular Dynamics Storm 860 Phosphorimager.

The observed rate constant, k_{obs}, for the first (5′ cleavage) and second (exon ligation) step of each TES reaction was quantified (*15*, *29*). The first step was obtained from a plot of the percent intermediate plus percent product formed versus time, and the second step was obtained from a plot of percent product formed versus time. Note that these observed rate constants reflect the rate of chemistry and any requisite conformation changes that occur.

Products obtained from the TES reactions were gel purified and sequenced by partial nuclease digestion, along with chemically synthesized versions of the expected TES products for comparison. Oligonucleotides were enzymatically sequenced using T1 (Epicentre; Madison, WI), U2 (Research Unlimited; Wellington, New Zealand), and Cl-3 (Research Unlimited) RNA nucleases. T1 reactions used 1.0×10^{-4} units T1 in 200 mM Tris-HCl (pH 7.5), Cl-3 reactions used 0.33 units Cl-3 in 200 mM Tris-HCl (pH 7.5), and U2 reactions used 0.33 units U2 in 200 mM Tris-HCl (pH 3.5). Sequencing reactions utilized approximately 50 fmol RNA and were incubated for 10 min at 55 °C. Immediately after adding an equal volume of stop buffer to each reaction, aliquots were loaded on a 13.5% polyacrylamide/8 M urea gel.

To determine if the 3′ product of the first reaction step (5′ cleavage) (Figure 1B) is dissociating and then rebinding the same (or different) ribozyme before the second reaction step (exon ligation), TES reactions were conducted for 1 h in 7 mM $MgCl_2$, 166 nM rP-8/4x ribozyme, 1.33 nM radiolabeled 36-mer substrate, and either 66.5 nM (50×) or 665 nM (500×) 3′ exon mimic competitor r(GUGCUCU). The values reported are the average of six independent assays. Likewise, to determine if the 5′ product of the first reaction step (5′ cleavage) (Figure 1B) is dissociating and then rebinding the ribozyme before the second step (exon ligation), TES reactions were conducted for 1 h in 7 mM $MgCl_2$, 166 nM rP-8/4x ribozyme, 1.33 nM 36-mer nonradiolabeled substrate, and 1.33 nM radiolabeled 5′ exon mimic competitor r(AUGACU). In each case the competitors were added simultaneously with the substrates.

RESULTS AND DISCUSSION

The P. carinii Group I Intron Ribozyme Catalyzes the TES Reaction. To test whether a ribozyme derived from a group I intron (and specifically one from *P. carinii*) catalyzes the TES reaction, a 36 nucleotide substrate was designed to bind the rP-8/4x ribozyme's native recognition element sequences (Figure 2A). These exon mimic sequences were separated by a bridge consisting of the first four bases of the intron and 13 uridines. Uridines were chosen because of their relatively poor ability to form intramolecular and intermolecular structures, and 13 was chosen arbitrarily. Typical results at the optimized $MgCl_2$ concentration (7 mM) and temperature (44 °C) at 1 h using 125-fold excess ribozyme over 5′ radiolabeled substrate are shown in Figure 3A. We expected a product band at 16 nucleotides in length, which was obtained in a yield of 25 ± 5% (for six independently run assays). The product band was extracted from the gel matrix, purified, and subjected to enzymatic sequencing, along with a chemically synthesized version of the expected product. The sequence and banding patterns obtained (Figure 4) show that the expected TES product is being generated. The TES reaction also produces a radiolabeled band at six nucleotides, which is a product of the first step of the reaction (5′ cleavage). Apparently, a majority of the 36-mer substrate undergoes only the first step of the reaction. The dependence of the TES reaction on $MgCl_2$ concentration, time, and rP-8/4x concentration is shown in Figure 3B. The k_{obs} for the first and second step of the reaction are 1.69 and 0.05 min^{-1}, respectively. The reaction is complete after 40 min and only requires 20 nM ribozyme for maximal activity (with 1.33 nM substrate). These results show that the rP-8/4x ribozyme, in the absence of a nucleotide cofactor, inherently contains the ability to bind a substrate in trans and catalyze the TES reaction.

Even though there are 18 other uridines that could be sites of 5′ cleavage and four other guanosines that could be sites of 3′ splicing for the 36-mer substrate, only the expected TES product is generated. Apparently, the ribozyme recognition elements that define the individual catalytic steps are sequence specific. That the ribozyme can catalyze this reaction at all is perhaps surprising in that the relatively long

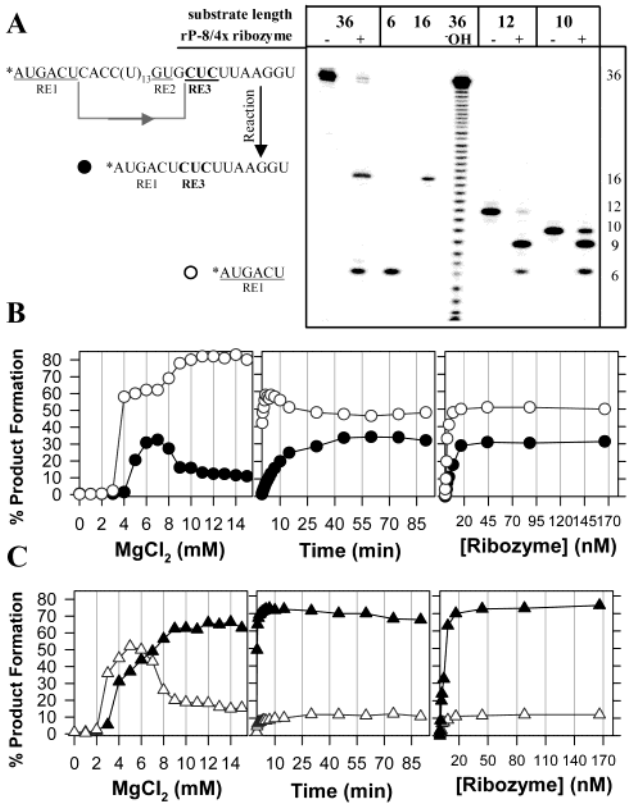

FIGURE 3: TES reaction using the *P. carinii* system. (A) Polyacrylamide gel showing substrates and products of the TES reaction using 166 nM rP-8/4x ribozyme and 1.33 nM substrate at 7 mM $MgCl_2$ (36-mer) and 10 mM $MgCl_2$ (10-mer and 12-mer) at 44 °C. The reaction using the 36-mer substrate is diagrammed on the left. The regions of the substrate that bind to the ribozyme's recognition elements (labeled RE1, RE2, and RE3) are underlined. All reactions in the presence (+) and absence (−) of the rP-8/4x ribozyme were subjected to the same incubation conditions. TES reactions were conducted using a 36-mer substrate (to give a 16-mer product), a 12-mer substrate (9-mer product), and a 10-mer substrate (9-mer product). See Figure 2 for the sequence of these substrates. The 6-mer lane shows a synthetic control for the 5′ cleavage products, the 16-mer lane shows a synthetic control for the 16-mer TES product, and the $^{-}$OH lane shows an alkaline digest of the 36-mer starting material. (B) Graphs of TES reactions using the 36-mer substrate. All reactions were run as above except for the changing variable. The TES product is represented by filled circles, and the 5′ cleavage product is represented by open circles. (C) Graphs of TES reactions using the 10-mer substrate. All reactions were run as above except for the changing variable. The TES product is represented by filled triangles, and the 5′ cleavage product is represented by open triangles. Each graph shows the average of two independent assays.

bridging region could be expected to sterically hinder the binding of the substrate to the catalytic core of the ribozyme or at least significantly hinder the required conformational rearrangement between the two catalytic steps (*11*, *30*). One or both of these might account for the majority (>50%) of the 36-mer substrate only undergoing the first catalytic step, in contrast to less than 10% only undergoing the first step for the self-splicing reaction in vitro (*15*, *17*). The lack of product breakdown seen in the time-dependence studies, however, indicates that the TES products themselves are not substrates for further reactions, although guanosine-dependent 5′ cleavage or ribozyme-mediated hydrolysis of the products could be a factor in vivo.

Previous studies utilizing the rP-8/4x ribozyme (*15*) show that the 5′ exon mimic r(AUGACU) binds to the rP-8/4x

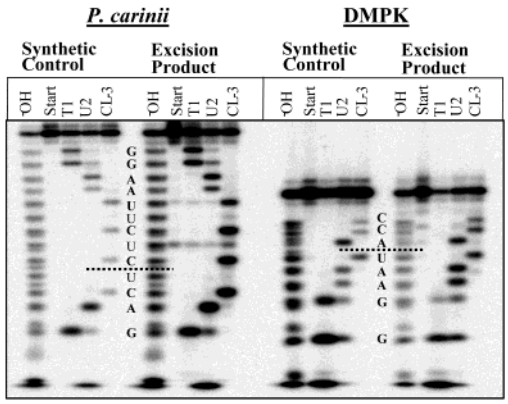

FIGURE 4: Sequencing of 5′ end radiolabeled TES products. For each system, a chemically synthesized version of the expected product was sequenced and run adjacent to the isolated excision product. Left: the *P. carinii* 16-mer product that results from treating the 36-mer substrate with the rP-8/4x ribozyme. Right: the 10-mer product that results from treating the 38-mer DMPK mimic with the rP-8/4x-MD ribozyme. The dotted line represents the newly created splice junction between the 5′ and the 3′ ends of the substrate. Nuclease T1 is specific for guanosine, U2 for adenosine, and CL-3 primarily for cytidine. The $^{-}$OH lane shows an alkaline digest of the starting material, and the lanes labeled Start show the starting material.

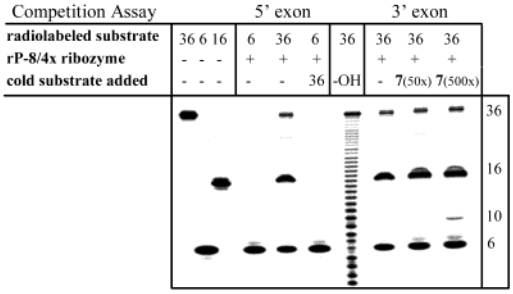

FIGURE 5: Competition TES reactions. Polyacrylamide gel showing substrates and products of TES reactions using 166 nM rP-8/4x and 7 mM $MgCl_2$ at 44 °C. Radiolabeled substrate refers to the length of the radiolabeled substrate at 1.33 nM final concentration, and cold substrate refers to the length of the nonradiolabeled substrate. The final concentrations of the cold substrates are 1.33 nM for the 36-mer and 66.5 nM (50×) or 665 nM (500×) for the 7-mer, which is r(GUGCUCU). The two substrates for each reaction were added simultaneously. Lanes designated in the first box are length controls, the second box shows the 5′ exon competition assay, the third box shows an alkaline digest of the 36-mer starting material, and the fourth box shows the 3′ exon competition assay.

ribozyme (K_d = 5.2 nM at 37 °C) 3 orders of magnitude more tightly than the 3′ exon mimic r(GUGCUCU) ($K_d \approx$ 20 μM at 37 °C). Interestingly, maximum TES product formation occurs with as little as 20 nM ribozyme (at 44 °C), indicating that for final product formation the 5′ and 3′ exon intermediates produced during the 5′ cleavage step might not dissociate and then rebind the ribozyme before the exon ligation step. To test for 5′ exon dissociation and rebinding between the two steps, TES reactions were conducted with 166 nM rP-8/4x, 1.33 nM nonradiolableled 36-mer, and 1.33 nM radiolabeled 5′ exon, r(AUGACU). In this case, if the 5′ exon intermediate dissociates from the ribozyme, the radiolabeled 5′ exon is just as likely to then bind the ribozyme and form the 16-mer product as the nonradiolabeled 5′ exon intermediate. As seen in Figure 5, no radiolabeled TES products are observed, indicating that the 5′ exon intermediate does not dissociate from the

ribozyme between the two steps (for those 5′ exon intermediates that undergo the complete reaction).

Likewise, to test for 3′ exon intermediate dissociation and rebinding between the two reaction steps, TES reactions were conducted with 166 nM rP-8/4x, 1.33 nM radiolabeled 36-mer, and a 50 (66.5 nM) or 500 (665 nM) fold excess of a nonradiolabeled 3′ exon mimic competitor, r(GUGCUCU), which would form a 10-mer competition product. At equal molar concentrations, if the 3′ exon intermediate dissociates from the ribozyme, the 7-mer competitor is 2.5 times more likely to bind the ribozyme and be a substrate in the second reaction step than the 30-mer 3′ exon intermediate (data not shown). The results (Figure 5) show that a 500-fold excess of cold competitor over substrate does not significantly reduce the amount of 16-mer product formed (19.4 ± 2.3% vs 22.8 ± 3%, respectively). The small amount of 10-mer product that is observed at 500-fold excess competitor over substrate (but not 50-fold excess) is not actually competing with the TES reaction. In these cases, the ribozymes that have bound radiolabeled 5′ exon regions, and for which the 3′ exon region has dissociated, are binding and reacting with a small amount of the huge excess of 3′ exon competitor. Therefore, the vast majority of substrates that undergo the complete TES reaction do not have 3′ exon intermediate dissociation and rebinding occurring between the two steps of the reaction. Apparently, substrates that undergo only the first reaction step do so because of nearly irreversible 5′ or 3′ exon intermediate dissociation. It follows that since intermediates to the complete TES reaction do not dissociate from the ribozyme, the TES reaction is intramolecular with regard to substrate.

Note that other group I intron-derived ribozymes might be able to catalyze the TES reaction. If so, these ribozymes might need the guanosine cofactor usually required for the first catalytic step. This could be undesirable, however, as the cofactor is expected to increase the reversibility of the first reaction step (*31*, *32*), potentially leading to less TES product being formed.

The P. carinii Ribozyme Can Excise as Little as One Nucleotide. To determine if a lower limit exists to the length of the excised region, we tested the TES reaction using the rP-8/4x ribozyme with two new substrates. One substrate is a 12-mer, r(AUGACUGUGCUC) and was designed to contain the minimum length bridging sequence that could utilize the 2 base pair RE2 interaction (to form the P9.0 helix) and the 3′ guanosine thought to be required for self-splicing (Figure 2B, panel i). The other substrate is a 10-mer, r(AUGACUGCUC), which cannot utilize the RE2 interaction and from which only one nucleotide would be excised (Figure 2B, panel ii). The results (Figure 3A) show that, under the optimal conditions of 10 mM $MgCl_2$ and 44 °C, both the 12-mer and the 10-mer reactions lead to the formation of the expected 9-mer products, as confirmed by enzymatic sequencing (data not shown). The optimized reactions produce 72 ± 3.9% product for the 12-mer substrate and 69.3 ± 4.4% product for the 10-mer substrate (for six independently run assays). Apparently, the rP-8/4x ribozyme can excise as little as a single nucleotide. That we get the same approximate yield using the 12-mer and 10-mer substrates suggests that the role of forming the P9.0 helix is not large in this case. Therefore, the RE2 interaction, although perhaps beneficial, is not required for sequence specific TES reactions. In addition, the 12-mer and 10-mer substrates lead to more than twice the product as compared with the 36-mer substrate, implicating the longer bridging region (which includes the four 5′ bases of the intron) as being detrimental for this reaction. As the amount of substrate that undergoes at least the first reaction step is similar for all of the different substrates, 3′ exon intermediate dissociation for the 36-mer likely accounts for the difference in extent of final product formation. The dependence of the 10-mer substrate reaction on $MgCl_2$ concentration, time, and rP-8/4x concentration is shown in Figure 3C. The k_{obs} for the first and second step of the reaction are 4.12 and 2.89 min^{-1}, respectively. In contrast to that for the 36-mer, the reaction with the 10-mer substrate is more favorable at $MgCl_2$ concentrations greater than 7 mM, and the second reaction step occurs approximately 50-fold faster. The origin of this effect is unknown but could be due to the reduced steric hindrance of the smaller bridge on the required conformational rearrangement between the two reaction steps. This could reflect an increased affinity or accessibility of the 3′ guanosine of the bridge for the G-binding site of the ribozyme. Indeed, previous reports suggest that the ability of the G-binding site to bind this endogenous guanosine drives the second step of the reaction (*29*, *31*).

Sequence Specificity of the P. carinii Ribozyme Can Be Altered. For the TES reaction to be useful, the recognition elements must be mutable to target predetermined sequences. Therefore, a truncated DMPK mimic model system was developed (Figure 2B, panel iii) for analyzing the ability of a re-engineered ribozyme to excise the RNA triplet repeat region whose expansion is the causative agent of the effects of the genetic disease myotonic dystrophy (*26*−*28*, *33*, *34*). We designed a ribozyme to target the uridine immediately upstream of the repeats. Because there are five successive guanosines immediately following the repeats (which could lead to a mixture of TES products), we targeted the first guanosine downstream from the five successive guanosines. The substrate bridge contained five CUG repeats, as this is the smallest number of repeats thought to form the hairpin structure similar to that seen with the expanded transcripts (*35*). Note that myotonic dystrophy patients actually have greater than 15 CUG repeats. Nevertheless, the ribozyme targets the flanking regions and would be the same regardless of the number of triplet repeats. We termed this ribozyme the rP-8/4x-MD ribozyme and this substrate the 38-mer DMPK mimic. The results at the optimum $MgCl_2$ concentration (13 mM) and temperature (44 °C) are shown in Figure 6A. A 10 nucleotide product was obtained, as expected, in a yield of 61.1 ± 4.6% (for six independently run assays). Besides unreacted 6-mer generated from the first step of the reaction, no other products are produced to any significant amount, indicating a reasonably specific reaction. The 10-mer product was extracted from the gel and enzymatically sequenced, along with a chemically synthesized version of the expected product. The sequence and banding patterns obtained (Figure 4) show that the expected TES product is being generated. Apparently, the ribozyme can be modified (and simply; at just RE1, RE2, and RE3) to target non-native substrates. The dependence of this reaction on $MgCl_2$ concentration, time, and rP-8/4x-MD concentration is shown in Figure 6B. The k_{obs} for the first and second step of the reaction are 0.41 and 0.44 min^{-1}, respectively (note that these

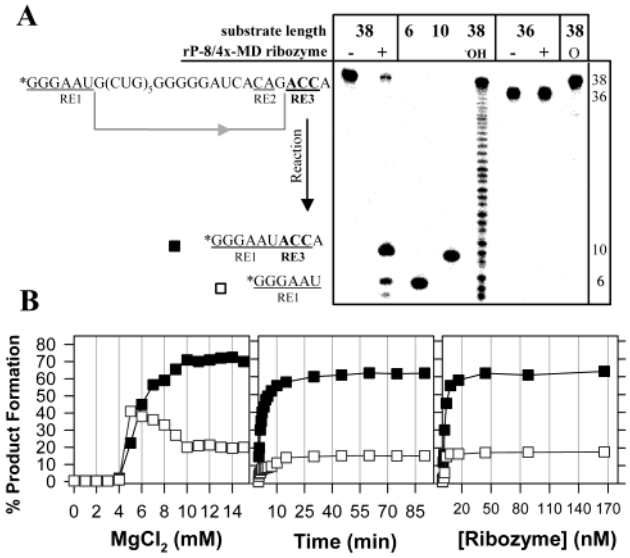

FIGURE 6: TES reaction using the DMPK model system. (A) Polyacrylamide gel showing substrates and products of the TES reaction using 166 nM rP-8/4x-MD ribozyme, 1.33 nM substrate, and 13 mM $MgCl_2$ at 44 °C. The reaction using the 38-mer DMPK mimic is diagrammed on the left. The regions of the substrate that bind to the ribozyme's recognition elements are labeled RE1, RE2, and RE3. All reactions in the presence (+) and absence (−) of the rP-8/4x-MD ribozyme were subjected to the same incubation conditions. TES reactions were conducted using the 38-mer substrate (to give a 10-mer product). The 6-mer lane shows a synthetic control for the 5′ cleavage products, the 10-mer lane shows a synthetic control for the 10-mer TES product, and the $^{-}$OH lane shows an alkaline digest of the 38-mer starting material. The lanes labeled 36 are TES reactions using the rP-8/4x-MD ribozyme with the 36-mer *P. carinii* substrate (at 13 mM $MgCl_2$), and the lane labeled 38 (lane O) is a reaction using the rP-8/4x ribozyme with the 38-mer DMPK mimic (at 7 mM $MgCl_2$). In these cases, no reaction occurs. (B) Graphs of TES reactions using the 38-mer substrate and rP-8/4x-MD. All reactions were run as above except for the changing variable. The TES product is represented by filled squares, and the 5′ cleavage product is represented by open squares.

values are within experimental error and that the rate of the second step is likely limited by the rate of the first step). Interestingly, the excision of this 28-mer bridge, which could include a triplet repeat hairpin structure (*35*), is substantially more effective than the 36-mer rP-8/4x system, which excises a 20-mer unstructured bridge. Apparently, targets with large bridge regions are not necessarily poor reaction substrates.

To determine whether each ribozyme has specificity for its intended target, we ran TES reactions using the rP-8/4x-MD ribozyme with the 36-mer *P. carinii* substrate at 13 mM $MgCl_2$ and the rP-8/4x ribozyme with the 38-mer DMPK mimic at 7 mM $MgCl_2$, each at 44 °C (Figure 6A). In addition, we ran the reaction using the rP-8/4x-MD ribozyme at 7 mM $MgCl_2$ and the reaction using the rP-8/4x ribozyme at 13 mM $MgCl_2$ (data not shown). In these reactions, not even 5′ cleavage products are observed, indicating that the ribozymes have some specificity for their intended target substrates. Note that although the $MgCl_2$ concentration required for maximum activity differs for all the reactions, they all occur at or below that required (15 mM) for maximum activity for the *P. carinii* group I intron self-splicing reaction in vitro (*15*).

While it was known that the sequence of RE1 could be modified in reactions mimicking the first step of the self-splicing reaction (*14*, *20*, *21*), it was previously not known that all three recognition elements were modifiable. That these recognition elements completely specify binding and reactivity indicates that they are the primary determinants of specificity between the ribozyme and its substrate. In addition, bridging regions of different sequences and lengths (1, 3, 20, and 28 nucleotides) have been excised, indicating that perhaps a 3′ G in the bridging region might be the only sequence requirement for the excised segment.

Comparison with Previous Results. Sargueil and Tanner (*16*) previously reported that, at high $MgCl_2$ concentration (10−100 mM) and temperature (55−65 °C), *Tetrahymena* group I intron ribozymes that lack both of the exons and the IGS can catalyze a guanosine cofactor-mediated TES-like reaction upon binding pseudoknot structured substrates, which creates the in trans equivalent of the P1 and P10 helices (P9.0 helix formation was not required). The ribozymes target and bind these pseudoknot structures entirely through tertiary interactions. The ribozymes used in this report, in contrast, contain a modifiable IGS (RE1 and RE3) and RE2, which allows the ribozyme to target designated substrates at the level of simple base pairing. Furthermore, the TES reactions using the *P. carinii* ribozymes do not require a guanosine cofactor and optimally occur at a lower $MgCl_2$ concentration (7−13 mM) and temperature (44 °C).

Sullenger and Cech (*22*) previously reported that *Tetrahymena* group I intron ribozymes that lack a 5′ exon, but contain an endogenous non-native 3′ exon, can catalyze the covalent attachment of the endogenous 3′ exon to mutant transcripts in such a way as to replace the 3′ end of mutant transcripts with normal corrected versions. While trans-splicing can repair RNA, it exploits only the RE1 molecular interaction, the ribozymes must be covalently attached to the repaired half of the transcript, it is single turnover, and repairing mutations distant from the 3′ end of long transcripts could be problematic. In contrast, the TES reaction exploits multiple molecular interactions (RE1, RE3, and perhaps RE2), TES ribozymes excise an internal segment from within RNA substrates, the reaction is potentially multiple turnover, and the position of the mutations within the transcript is not a limiting factor. Note, however, that under the conditions used in this report, little (if any) turnover is occurring. Nevertheless, these two complementary reactions share many similarities, and so the wealth of knowledge reported for the trans-splicing reaction (*6*, *22*−*24*) will likely be applicable to the TES reaction.

Implications. Group I introns can splice themselves out of RNA transcripts, and through reverse-splicing, can embed themselves into intronless transcripts (*36*, *37*). Perhaps in combination with host cell activities such as reverse transcription and recombination, these activities could account for the integration of introns into the genome (*38*−*40*). That group I intron ribozymes can inherently excise specific regions from within RNAs suggests that ribozymes could have been used for altering transcripts without becoming embedded within them, for example, as trans splicing (or trans alternative splicing) agents. This would be even more compelling if such ribozyme activities could occur at the DNA level, which is not implausible, as endonuclease encoding group I introns have been shown to act as mobile genetic elements at the DNA level (*41*). Similarly, it would be of interest if group II introns, which encode homing

proteins that result in their insertion into DNA (*41*, *42*), could catalyze the TES reaction.

Unfortunately, nature has afforded a huge number of mutations that occur at the RNA level that predispose individuals to disease. New therapeutic strategies to combat these diseases are needed. That the recognition elements of the rP-8/4x ribozyme and the size and sequence of the excised region are all modifiable suggests that TES ribozymes may be generally useful as potential therapeutics against a variety of diseases. Therapeutic strategies include removing premature stop codons, restoring altered reading frames, and removing insertion mutations that affect transcription or the regulation of translation. Note, however, that group I intron-derived ribozymes appear to have two inherent sequence requirements: the first is a uridine (*13*, *43*) at the 5′ cleavage site (although a cytidine might also work (*44*, *45*)), and the second is a guanosine (*46*) at the 3′ splice site (Figure 2A). In addition, both steps of the TES reaction must occur to achieve the therapeutic goal. Also note that the general strategy of targeting mutations at the RNA level should be considered a potential treatment rather than a potential cure, as mutant RNA transcripts will continue to be produced. Furthermore, large RNAs have inherent limitations as in vivo therapeutic agents, including their apparent low sequence specificity (*25*, *37*, *47*−*52*). It will also be important to determine whether the TES reaction is permitted using actual myotonic dystrophy transcripts, which likely contain a number of stable secondary structure elements, as well as a larger number of CUG repeats. Nevertheless, TES ribozymes represent a new way to approach these genetic problems.

ACKNOWLEDGMENT

The authors thank Rashada Alexander for technical assistance and Dana Baum for critical reading of the manuscript.

REFERENCES

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) *Cell 31*, 147−157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) *Cell 35*, 849−857.
3. Landweber, L. F., Simon, P. J., and Wagner, T. A. (1998) *BioScience 48*, 94−103.
4. Seetharaman, S., Zivarts, M., Sudarsan, N., and Breaker, R. R. (2001) *Nat. Biotechnol. 19*, 336−341.
5. Frauendorf, C., and Jaschke, A. (2001) *Bioorg. Med. Chem. 9*, 2521−2524.
6. Watanabe, T., and Sullenger, B. A. (2000) *Adv. Drug Delivery Rev. 44*, 109−118.
7. Lewin, A. S., and Hauswirth, W. W. (2001) *Trends Mol. Med. 7*, 221−228.
8. Guo, H., Karberg, M., Long, M., Jones, J. P., Sullenger, B. A., and Lambowitz, A. M. (2000) *Science 289*, 452−457.
9. Barroso-Del Jesus, A., and Berzal-Herranz, A. (2001) *EMBO Rep. 2*, 1112−1118.
10. Jarvis, T. C., Bouhana, K. S., Lesch, M. E., Brown, S. A., Parry, T. J., Schrier, D. J., Hunt, S. W., III, Pavco, P. A., and Flory, C. M. (2000) *J. Immunol. 165*, 493−498.
11. Jaeger, L., Michel, F., and Westhof, E. (1996) in *Catalytic RNA* (Eckstein, F., and Lilley, D., Eds.) Vol. 10, pp 33−51, Springer-Verlag, Heidelberg, Germany.
12. Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543−568.
13. Zaug, A. J., Been, M. D., and Cech, T. R. (1986) *Nature 324*, 429−433.
14. Zaug, A. J., and Cech, T. R. (1986) *Science 231*, 470−475.
15. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) *Biochemistry 36*, 15303−15314.
16. Sargueil, B., and Tanner, N. K. (1993) *J. Mol. Biol. 233*, 629−643.
17. Testa, S. M., Gryaznov, S. M., and Turner, D. H. (1998) *Biochemistry 37*, 9379−9385.
18. Testa, S. M., Gryaznov, S. M., and Turner, D. H. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 2734−2739.
19. Testa, S. M., Disney, M., Turner, D. H., and Kierzek, R. (1999) *Biochemistry 38*, 16655−16662.
20. Been, M. D., and Cech, T. R. (1986) *Cell 47*, 207−216.
21. Murphy, F. L., and Cech, T. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9218−9222.
22. Sullenger, B. A., and Cech, T. R. (1994) *Nature 371*, 619−622.
23. Lan, N., Howrey, R. P., Lee, S. W., Smith, C. A., and Sullenger, B. A. (1998) *Science 280*, 1593−1596.
24. Phylactou, L. A., Darrah, C., and Wood, M. J. (1998) *Nat. Genet. 18*, 378−381.
25. Roman, J., Rubin, M. N., and Woodson, S. A. (1999) *RNA 5*, 1−13.
26. McLaughlin, B. A, Spencer, C., and Eberwine, J. (1996) *Am. J. Hum. Genet. 59*, 561−569.
27. Korade-Mirnics, Z., Babitzke, P., and Hoffman, E. (1998) *Nucleic Acids Res. 26*, 1363−1368.
28. Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000) *Science 289*, 1769−1772.
29. Zarrinkar, P. P., and Sullenger, B. A. (1998) *Biochemistry 37*, 18056−18063.
30. Cech, T. R., Herschlag, D., Piccirilli, J. A., and Pyle, A. M. (1992) *J. Biol. Chem. 267*, 17479−17482.
31. Golden, B. L., and Cech, T. R. (1996) *Biochemistry 35*, 3754−3763.
32. Tanner, M. A., and Cech, T. R. (1996) *RNA 2*, 74−83.
33. Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., et al. (1992) *Science 255*, 1253−1255.
34. Bush, E. W., Taft, C. S., Meixell, G. E., and Perryman, B. (1996) *J. Biol. Chem. 271*, 548−552.
35. Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, C. A. (2000) *RNA 6*, 79−87.
36. Woodson, S. A., and Cech, T. R. (1989) *Cell 57*, 335−345.
37. Roman, J., and Woodson, S. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2134−2139.
38. Cech, T. R. (1985) *Int. Rev. Cyt. 93*, 3−22.
39. Sharp, P. A. (1985) *Cell 42*, 397−400.
40. Dujon, B. (1989) *Gene 82*, 91−114.
41. Belfort, M., and Perlman, P. S. (1995) *J. Biol. Chem. 270*, 30237−30240.
42. Cousineau, B., Smith, D., Lawrence-Cavenagh, S., Mueller, J. E., Yang, J., Mills, D., Manias, D., Dunny, G., Lambowitz, A. M., and Belfort, M. (1988) *Cell 94*, 451−462.
43. Waring, R. B., Towner, P., Minter, S. J., and Davies, R. W. (1986) *Nature 321*, 133−139.
44. Hur, M., and Waring, R. B. (1995) *Nucleic Acids Res. 23*, 4466−4470.
45. Disney, M. D., Testa, S. M., and Turner, D. H. (2000) *Biochemistry 39*, 6991−7000.
46. Price, J. V., and Cech, T. R. (1988) *Genes Dev. 2*, 1439−1447.
47. Watanabe, T., and Sullenger, B. A. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 8490−8494.
48. Zaug, A. J., Grosshan, C. A., and Cech, T. R. (1988) *Biochemistry 27*, 8924−8931.
49. Herschlag, D., and Cech, T. R. (1990) *Biochemistry 29*, 10172−10180.
50. Roman, J., and Woodson, S. A. (1995) *RNA 1*, 478−490.
51. Jones, J. T., Lee, S.-W., and Sullenger, B. A. (1996) *Nat. Med. 2*, 643−648.
52. Campbell, T. B., and Cech, T. R. (1996) *Biochemistry 35*, 11493−11502.

BI0267386