Enhancing the Second Step of the Trans Excision-Splicing Reaction of a Group I Ribozyme by Exploiting P9.0 and P10 for Intermolecular Recognition†

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

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

Received October 19, 2003; Revised Manuscript Received February 6, 2004

ABSTRACT: We previously reported that a group I intron-derived ribozyme can catalyze the excision of targeted sequences from within RNAs in vitro and that dissociation of the bridge-3′ exon intermediate between the two reaction steps is a significant contributing factor to low product yields. We now analyze the effects of increasing the length, and thus the strength, of helices P9.0 and P10, which occur between the ribozyme and the bridge-3′ exon region of the substrate, on this trans excision-splicing reaction. Using substrates where lengthy targeted regions are excised, these modifications can significantly increase product yields, specifically by enhancing the second reaction step. A threshold for product formation is obtained, however, at around five base pairs for P10 and eight base pairs for P9.0. Nevertheless, elongating P9.0 appears to be the more effective strategy, as both substrate binding and the rate of the second reaction step increase. In addition, P10 is required when P9.0 is not elongated. Also, a strong P9.0 helix cannot replace a weaker P10 helix, indicating that P9.0 and P10 play somewhat distinct roles in the reaction. We also show that second-step inhibition stems from the formation of an extended P1 helix (P1ex), consisting of as little as a single Watson−Crick base pair, as well as the mere presence of substrate nucleosides immediately downstream from P10. Both of these inhibitory components can be overcome by utilizing P9.0 and P10 elongated ribozymes. This work sets forth an initial framework for rationally designing more effective trans excision-splicing ribozymes.

We have previously demonstrated that a ribozyme derived from a *Pneumocystis carinii* group I intron can catalyze the excision of targeted sequences from within RNAs (*1*). This trans excision-splicing (TES)[1] reaction consists of two catalytic steps: 5′ cleavage followed by exon ligation (see Figure 1). In the 5′ cleavage step, a nucleophile from the aqueous solvent cleaves the phosphodiester backbone of a substrate to generate 5′ exon and bridge-3′ exon intermediates. In the exon-ligation step, the nucleophilic 5′ exon intermediate attacks a specific guanosine within the bridge-3′ exon intermediate. The result is the ligation of the exon intermediates and the concurrent excision of an internal, bridging segment of the RNA substrate. We have shown that such ribozymes can be modified to target a variety of sequences, including those where the excised segment is a single nucleotide (*1*).

A potential limitation of the TES reaction is that the bridge-3′ exon intermediate can dissociate from the ribozyme before the second reaction step (*1*). This appears to occur because of the relatively poor binding of the bridge-3′ exon intermediate to the ribozyme in comparison with the 5′ exon intermediate (*1*, *2*). For example, in previous studies, we utilized a TES substrate that mimics the native 5′ and 3′ exon sequences of the *P. carinii* group I intron, with the addition of a uridine-rich region bridging the two exons (shown in Figure 1). This 36-mer substrate, although able to effectively undergo the first reaction step (>80%), was only moderately successful in the second reaction step (30%), with the inhibition markedly stronger at relatively high magnesium concentrations (>8 mM). The origin of this second-step inhibition was unknown but did not appear to correspond with the particular sequences at the molecular interface between the ribozyme and substrate, as short mimics of the 36-mer substrate that contain the same molecular contacts (the P1, P9.0, and P10 helices) work well (70%) (*1*). Likewise, the length of the excised fragment did not appear to be problematic, as a ribozyme directed to target a mimic of a myotonic dystrophy RNA transcript (*1*), which has a similar bridge length, works well (70%).

Group I intron-derived ribozymes, including TES ribozymes, can utilize three molecular recognition elements to bind the bridge-3′ exon intermediate (*3*−*8*): RE2 to form P9.0, RE3 to form P10, and the last guanosine of the intron, referred to as ωG (shown in Figure 1). We analyzed whether increasing the length (and thus the strength) of P9.0 and P10, via the addition of base pairs, can aid in the stabilization of the bridge-3′ exon intermediate, thus leading to more effective TES ribozymes. Our results show that elongating P10 from 3 to 5 to 10 base pairs enhances product yields (from 30 to 50%). In addition, incrementally elongating P9.0 from 2 to 14 base pairs incrementally enhances product yields (from 30 to 70%), with the resultant P9.0 helix stability gain increasing more than linearly with the addition of successive

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

* 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; RE1, recognition element #1; RE2, recognition element #2; RE3, recognition element #3.

10.1021/bi035874n CCC: $27.50
Published on Web 03/16/2004

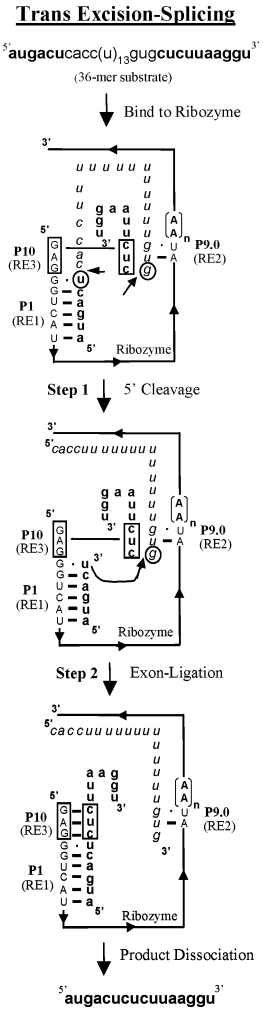

FIGURE 1: Schematic of the two-step TES reaction. The rP-8/4x ribozyme is in uppercase lettering, the 36-mer substrate is in lowercase lettering, the substrate bridge is in italic, and the exon regions to be ligated are in bold. The ribozyme recognition elements RE1, RE2, and RE3 base pair with the substrate to form the P1, P9.0, and P10 helices, respectively. The catalytic sites for the 5′ cleavage and exon-ligation steps are represented by arrows in the uppermost diagram. The P10 helix is boxed. The 5′ uridine and 3′ guanosine that help define the catalytic sites are circled. Note that the bracketed region above P9.0 is non-native. In designated studies, adenosine doublets were inserted at this position ($n = 0-6$), which can base pair with the substrate bridge to form an extended P9.0 helix.

base pairs. A strong P9.0 helix, however, cannot completely make up for the loss of the apparently weaker P10 helix, indicating that these two molecular recognition components do not play purely redundant roles. We also show that enhancing the stability of P9.0 reduces the rate of the first reaction step of the TES reaction, while it increases the rate of the second reaction step.

We investigated why TES reactions that use the 36-mer substrate give relatively poor second-step yields. Two molecular factors are largely responsible. First, if the 5′ base of the region to be excised forms a Watson−Crick base pair with the 3′ base of RE3, the resultant base pair, after the first reaction step, partially inhibits the formation of the P10 helix. We show that P10 formation is critical for the reaction to occur. Second, the mere presence of substrate nucleosides immediately downstream from P10 also inhibits the reaction. Finally, we show that elongating P9.0 and P10 in TES ribozymes can reduce or negate these inhibitory substrate−ribozyme interactions. The implications of these results regarding group I intron-derived ribozyme activity, as well as for rationally designing more effective TES ribozymes, are discussed.

MATERIALS AND METHODS

Nucleic Acid Synthesis and Preparation. DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and used without further purification. RNA oligonucleotides were obtained from Dharmacon Research Inc. (Boulder, CO) and deprotected following the manufacturer's protocol. Select RNAs were 5′-end radiolabeled and gel purified as previously described (*1*). The *P. carinii* ribozyme, rP-8/4x, was generated from the plasmid precursor P-8/4x as previously described (*1*, *2*).

The synthesis of ribozymes with modified P1 and P10 helices was accomplished by altering the IGS of plasmid P-8/4x using site-directed mutagenesis. Mutagenesis, transcription, and ribozyme purification were performed as previously described (*1*, *2*). Primers CGACTCACTATAGGTGGGTTCATGAAAG and CTTTCATGAACCCACCTATAGTGAGTCG were used to generate the plasmid P-8/4x-P1ex; primers CGACTCACTATAGAGAGGGTCATGAAAG and CTTTCATGACCCTCTCTATAGTGAGTCG were used to generate the plasmid P-8/4x-P10.5; primers CGACTCACTATAGCCTTAAGAGGGTCATGAAAGCGGC and GCCGCTTTCATGACCCTCTTAAGGCTATAGTGAGTCG were used to generate the plasmid P-8/4x-P10.10; and primers CGACTCACTATAGGTCATGAAAGCGGC and GCCGCTTTCATGACCTATAGTGAGTCG were used to generate the plasmids P-8/4x-noP10 and P-8/4x-noP10-10A (underlined bases indicate the modified RE3 regions).

The synthesis of ribozymes with modified P9.0 helices was accomplished similar to that for P1 and P10 above. In these cases, however, successive site-directed mutagenesis reactions were conducted such that two adenosines were added to the native P9.0 helix per reaction (six reactions were run in total). The resultant plasmid precursors were named P-8/4x-Y, where Y is the number and type of nucleosides added. The primers used were GGTATAGTCTAT(AA)$_n$CCTCTTTCG and CGAAAGAGG(TT)$_n$ATAGACTATACC, where n is from 1 to 6.

TES Reactions. Reactions were conducted in H*x*Mg buffer consisting of 50 mM Hepes (25 mM Na^+), 135 mM KCl, and *x* mM $MgCl_2$ at pH 7.5, where *x* refers to the amount of $MgCl_2$ in mM in the buffer (listed in the text and figures). Prior to each reaction, 200 nM ribozyme (from 0.2 to 200 nM in the concentration-dependent assay) was preannealed in 5 μL of the appropriate buffer at 60 °C for 5 min and then slowly cooled to 44 °C. Reactions were initiated by adding 1.0 μL of an 8 nM solution of 5′-radiolabeled RNA substrate (previously incubated in the appropriate buffer at 44 °C for 5 min). Final ribozyme and substrate concentrations were 166 and 1.3 nM, respectively (except for the ribozyme concentration-dependent assay). After 1 h (from 15 s to 90 min in the time-dependent assay), the reactions were terminated by the addition of an equal volume of stop buffer (10 M urea, 3 mM EDTA, and 0.1 × TBE). All reactions were confirmed to be complete at 1 h (time-dependent graphs

of representative TES reactions are available in the Supporting Information). Products and reactants were denatured for 1 min at 90 °C and separated on a 12% polyacrylamide, 8 M urea denaturing gel. Gels were dried under vacuum, and the bands were quantified with a Molecular Dynamics Phosphorimager Storm 860. Standard deviations for all graphed points are less than 10%.

The observed rate constant, k_{obs}, for the first (5′ cleavage) and second (exon ligation) steps of the TES reaction using the rP-8/4x-10A ribozyme was quantified as previously described (*1*, *2*). The first step was obtained from a plot of the 5′ cleavage intermediate (percent) plus the TES product (percent) versus time. The second step was obtained from a plot of the TES product (percent) versus time. Note that these observed rate constants reflect both the rate of chemistry and any requisite conformation changes that occur.

Exon-Ligation Reactions. Prior to each reaction, 233 nM ribozyme was preannealed in 5 μL H*x*Mg buffer at 60 °C for 5 min and then slowly cooled to 44 °C. At this point, 1.0 μL of an 8 nM solution of 5′-radiolabeled 5′ exon mimic r(*AUGACU*) in H*x*Mg buffer previously heated to 44 °C was added, and the solution was equilibrated for 15 min. Reactions were initiated by adding 1.0 μL of a 140 μM solution of either r(GUG*CUCU*) or r(CACC-$(U)_{13}$GUG*CUCUUAAGGU*) in H*x*Mg buffer at 44 °C. After 1 h, the reactions were terminated by the addition of an equal volume of stop buffer. Products were separated and quantified as outlined above.

Determination of Dissociation Constants. The dissociation constant, K_d, of the bridging segment of the 36-mer substrate, r(CACC$(U)_{13}$GUG), binding various ribozymes was approximated using band-shift gel electrophoresis assays (*2*, *9*, *10*). The final ribozyme concentrations used in these assays ranged from 0.1 nM to 6 μM. The ribozymes were first preannealed at 60 °C for 5 min in a 5 μL solution of H14Mg buffer and 4.5% glycerol (v/v) and then slowly cooled to 22 °C. Binding assays were initiated by adding 2.5 μL solutions of 1.3 nM 5′-end-radiolabeled bridge, which was previously heated to 22 °C in H14Mg buffer. The reactions equilibrated for 1 h. The fraction of the bridging region bound to the ribozyme was separated from the unbound on a 22 °C, 10% native polyacrylamide gel (made with H14Mg buffer). Gels were dried, and bands were quantified on a Molecular Dynamics Phosphorimager. Note that the ribozyme rP-8/4x was also analyzed in H7Mg buffer, but no measurable binding was detected. Dissociation constants were calculated as previously described (*2*). The data reported are the average of at least two independent assays. Quantification of dissociation constants at 44 °C, which is the usual temperature at which TES reactions are conducted, was unsuccessful because of weak binding.

RESULTS

To analyze each individual ribozyme−substrate permutation, we conducted TES reactions as a function of magnesium concentration. This was done because previous studies have shown that TES reactions using different substrates/ribozyme combinations can be optimal at different magnesium concentrations (*1*). Such assays allow for the comparison of the individual permutations as a function of the maximum yield that each reaction can attain, independent of whether the optimal magnesium concentration changes between the permutations. The reactions were all conducted for 1 h, because that is an endpoint for each reaction. In fact, almost all reactions (both the first and second steps) were complete within 30 min, with no measurable product breakdown occurring. The reactions were conducted under ribozyme excess conditions (relative to the substrate) because preliminary indications are that TES ribozymes, in their current incarnation, are not acting in a multiple-turnover manner.

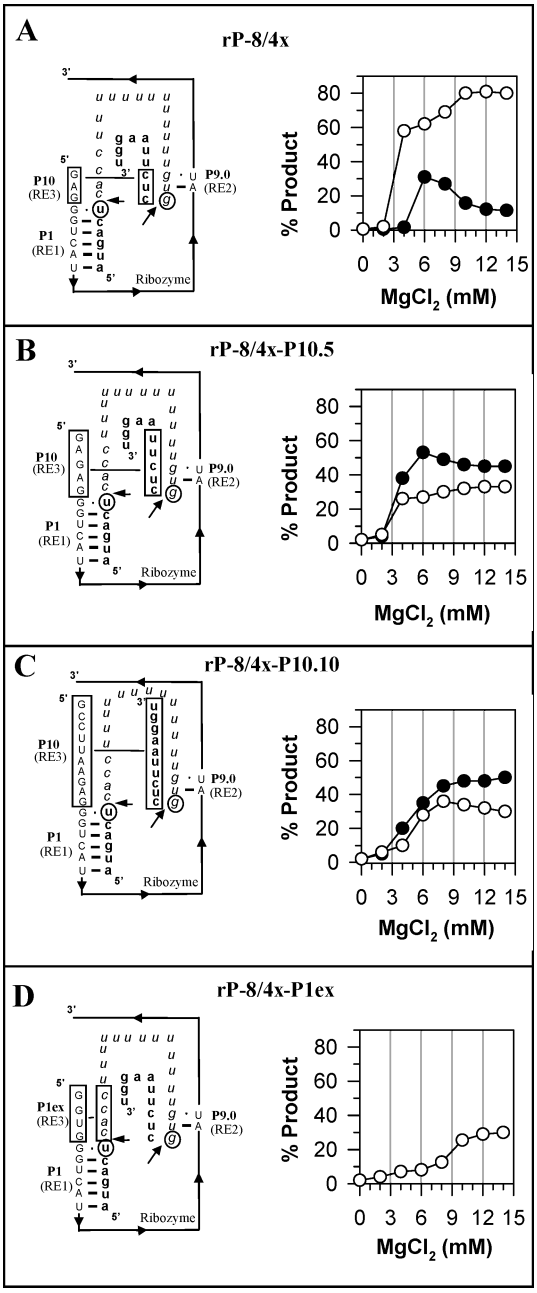

FIGURE 2: Magnesium dependence of TES reactions using RE3-modified ribozymes with the 36-mer substrate. See the caption of Figure 1 for a description of the various diagram components. Reactions were conducted at 44 °C for 1 h using 166 nM ribozyme and 1.3 nM 5′-end-radiolabeled substrate at the indicated magnesium concentration. The open circles in the graphs represent the 5′ exon intermediate, and the closed circles represent the TES product. The ribozymes used were rP-8/4x (A), rP-8/4x-P10.5 (B), rP-8/4x-P10.10 (C), and rP-8/4x-P1ex (D). Note that in D no TES product is formed. The data points are the average of two independent assays. A representative polyacrylamide gel showing TES reactions is shown in Figure 4A.

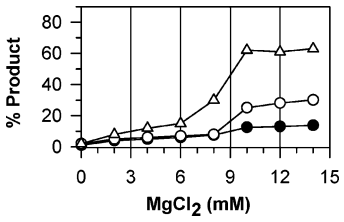

FIGURE 3: Magnesium-dependent graphs of TES reactions using the 36-mer substrate with ribozymes lacking RE3. Reactions were conducted at 44 °C for 1 h using 166 nM ribozyme and 1.3 nM 5′-end-radiolabeled substrate at the indicated magnesium concentration. Open triangles represent the 5′ exon intermediate, which forms when using the rP-8/4x-noP10 ribozyme. rP-8/4x-noP10 is the same as rP-8/4x but lacks RE3. Note that only the first reaction step occurs. Open circles represent the 5′ exon intermediate, and the closed circles represent the TES product that forms when using the rP-8/4x-noP10-10A ribozyme. rP-8/4x-noP10-10A is the same as rP-8/4x-noP10 but has 10 adenosines added to RE2. All assays represent the average of two independent assays.

This, then, allows us to maximize substrate binding and hence product yield.

Elongating P10 Can Enhance the Second Step of the TES Reaction. The ribozymes rP-8/4x, rP-8/4x-P10.5, and rP-8/4x-P10.10 are expected to form 3, 5, and 10 base pair P10 helices, respectively, with the 36-mer substrate. With the P10-elongated ribozymes, approximately 50% TES product is formed (parts B and C of Figure 2), compared with a maximum of 30% for the native ribozyme sequence (Figure 2A). Therefore, elongating P10 does enhance the TES reaction. Note that P10 appears to have a reactivity threshold of around five base pairs.

While this result shows that P10 formation plays a role in second-step reactivity, its importance in the TES reaction has not been analyzed. Therefore, we synthesized ribozyme rP-8/4x-noP10, which does not contain a RE3 and so cannot form P10 (the 5′ GAG is removed, see Figure 1). Using the 36-mer substrate, the results (Figure 3) show that rP-8/4x-noP10 can readily undergo the first step of the TES reaction, but the second step is completely prohibited. Evidently, P10 formation is critical for this TES reaction.

Elongating P9.0 Can Enhance the Second Step of the TES Reaction. We synthesized a series of P9.0-elongated ribozymes by inserting adenosines, in groups of two, into RE2 of the ribozyme such that the adenosines can base pair with the uridine-rich region of the 36-mer-substrate bridge (see Figure 1). The results (Figure 4) show that these ribozymes are functional and generate the expected TES product. While the extent of reaction for the first reaction step is independent of P9.0 length (Figure 4C), the extent of reaction for the second step increases with increasing P9.0 length. Therefore, elongating P9.0 enhances the amount of TES product formed, specifically by enhancing the second reaction step.

We confirmed the above result by isolating the second reaction step in exon-ligation reactions with P9.0-altered ribozymes. In these assays, the ribozymes ligate the 5′ exon mimic r(*AUGACU*) to one of two bridge-3′ exon mimics (Figure 5). The first bridge-3′ exon mimic is a 7-mer, r(GUG*CUCU*), which can form only a two base P9.0, irrespective of the amount of adenosines added to RE2 of the ribozyme. This system was used as a control for ribozyme activity. The results (Figure 5A) show that each of the ribozymes are effective at catalyzing the second step of the TES reaction. Note that all reactions work well in the absence

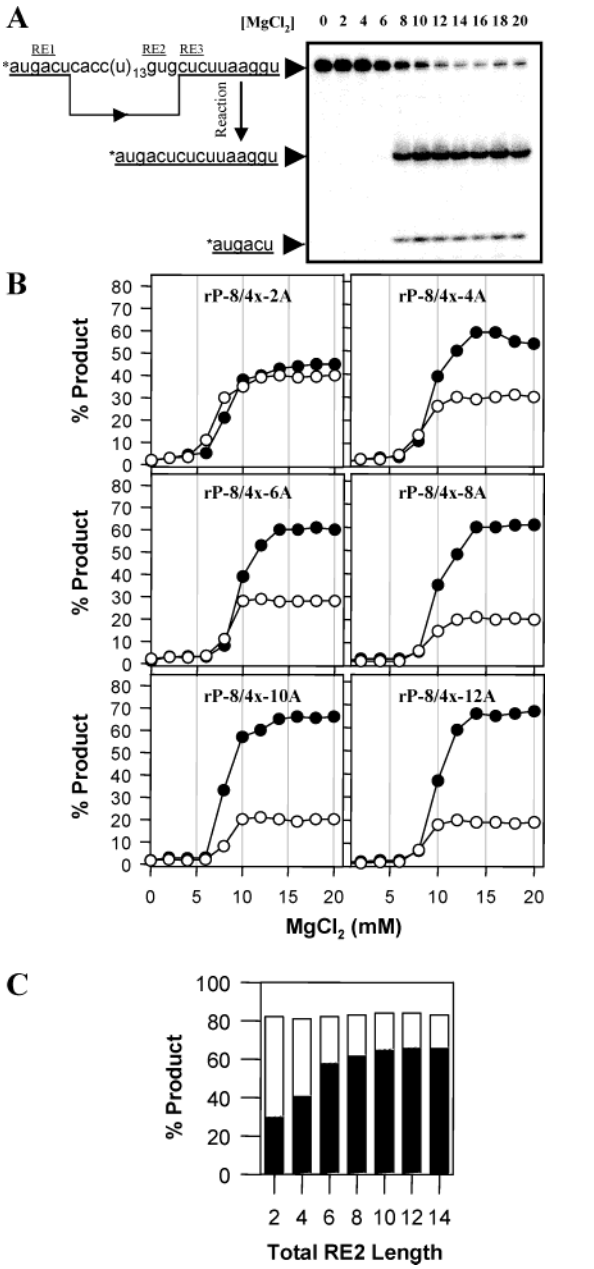

FIGURE 4: Magnesium dependence of TES reactions using RE2-modified ribozymes with the 36-mer substrate. Reactions were conducted at 44 °C for 1 h using 166 nM ribozyme and 1.3 nM 5′-end-radiolabeled substrate (marked with an asterisk) at the indicated magnesium concentration. The open circles in the graphs represent the 5′ exon intermediate, and the closed circles represent the TES product. (A) A representative polyacrylamide gel of the TES reaction using the ribozyme rP-8/4x-10A. The regions of the substrate that bind the ribozyme's recognition elements RE1, RE2, and RE3 are labeled as such. (B) Magnesium-dependent graphs of the TES reaction using designated P9.0 elongated ribozymes. (C) A summary of the data in B at 14 mM $MgCl_2$ (except for rP-8/4x, which is from Figure 2 and is at its optimum at 7 mM $MgCl_2$). Total RE2 length equals the native RE2 length plus the number of inserted adenosines. Black bars represent the final TES product, and white bars represent the 5′ exon intermediate. Note that the TES product necessarily has undergone the 5′ cleavage reaction. The data points are the average of two independent assays.

of the long bridging segment, which is characteristic of TES reactions where short bridging segments (from one to three nucleosides) are excised (*1*). The second bridge-3′ exon mimic analyzed is a 30-mer, which can potentially base pair

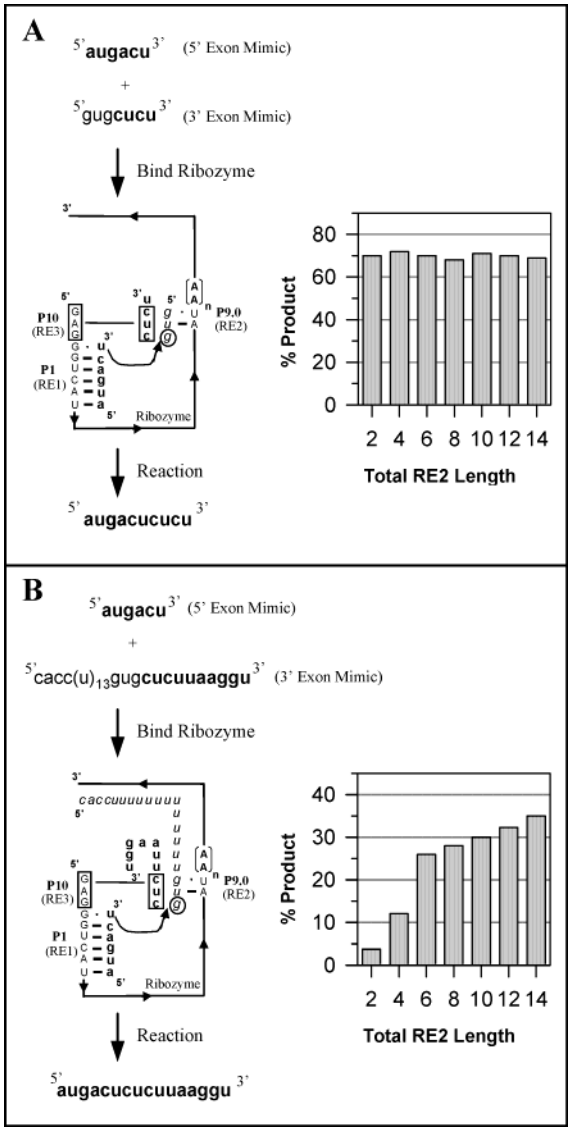

FIGURE 5: Diagrams and graphs of exon-ligation reactions as a function of RE2 length. See the caption of Figure 1 for a description of the various diagram components. Reactions were conducted at 44 °C for 1 h using 166 nM ribozyme, 1.3 nM 5′-end-radiolabeled substrate, and 20 μM of a bridge-3′ exon mimic at 14 mM $MgCl_2$ (7 mM $MgCl_2$ for rP-8/4x, where the total RE2 length is two). The diagrams show the 5′ exon mimic r(AUGACU) reacting with either the 7-mer bridge-3′ exon mimic r(GUG*CUCU*) (A) or the 30-mer bridge-3′ exon mimic r(CACC(U)$_{13}$GUG*CUCUUAAGGU*) (B). Solid bars in the graph represent the exon-ligation product. Note that, in these reactions, $n = 0-6$ in terms of the adenosine doublets added to RE2. Total RE2 length equals the native RE2 length plus the number of inserted adenosines. For clarity, the scale for graph A is twice that of graph B. The data points are the average of two independent assays.

with the adenosines added to the RE2 of the modified ribozymes. As seen in Figure 5B, increasing the length of RE2 can increase the yield of exon-ligation reactions in a manner that is analogous to the TES reaction (Figure 4C). Therefore, lengthening RE2 does enhance the second step of the TES reaction, probably by increasing the P9.0 length.

To determine whether P9.0-enhanced ribozymes do have an increased affinity for the bridging region of the substrate, direct gel-shift electrophoresis binding assays were conducted using a mimic of the bridging region, r(CACC(U)$_{13}$GUG). Assays were performed at 22 °C instead of 44 °C because

Table 1: Dissociation Constants of the Substrate Bridge Binding RE2-Elongated Ribozymes[a]

ribozymes	total RE2 length (nucleosides)	K_d (nM)
rP-8/4x-12A	14	7.2 ± 0.9
rP-8/4x-10A	12	88 ± 2.8
rP-8/4x-8A	10	470 ± 140
rP-8/4x-6A	8	1276 ± 220
rP-8/4x-4A	6	>2000
rP-8/4x-2A	4	>2000
rP-8/4x	2	>2000

[a] Dissociation constants were determined at 22 °C in 14 mM $MgCl_2$.

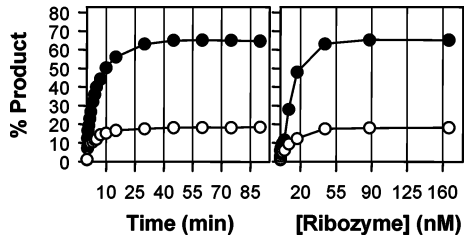

FIGURE 6: Characterization of the rP-8/4x-10A ribozyme. The graphs show the ability of rP-8/4x-10A to catalyze the TES reaction as a function of time and ribozyme concentration using the 36-mer substrate. Except for the changing variable, reactions were conducted at 44 °C for 1 h using 166 nM ribozyme and 1.3 nM 5′-end-radiolabeled substrate at 14 mM $MgCl_2$. The open circles in the graphs represent the 5′ exon intermediate, and the closed circles represent the TES product. All assays represent the average of two independent assays.

binding in this type of experiment, in the absence of the 5′ exon region, was not observed at the higher temperatures. The results (Table 1) show that, while binding is too weak to measure with ribozymes containing short RE2 lengths, a progressive increase in binding affinity is observed as the length of RE2 increases over six. Therefore, elongating RE2 enhances the ability of the ribozyme to bind the bridging region of the 36-mer RNA substrate. Note that, while this assay allows assessment of the P9.0 interaction independent of the P10 and P1 interactions, ωG at the 3′ splice junction is contained within the bridging region and could potentially bind the guanosine binding site of the ribozyme. Thus, binding values represent P9.0 and ωG interactions. Each of these interactions could conceivably be enhanced with RE2-elongated ribozymes.

Because the RE2-elongated ribozymes are likely to have properties distinct from the unmodified ribozyme, we further characterized the rP-8/4x-10A ribozyme (Figure 6). The reaction is complete in 30 min and requires approximately 40 nM ribozyme to attain maximal activity. The k_{obs} for the first step is 0.16 min^{-1} and that for the first and second step together is 0.16 min^{-1}. Thus, the rate of the second step for rP-8/4x-10A is limited by the rate of the first step. Compared to the unmodified ribozyme (*1*), the rate of the first step decreases 10-fold, while the rate of the second step increases at least 3-fold. This increase in the rate of the second step will likely have the desired effect of decreasing the time that the bridge-3′ exon intermediate can dissociate between the two reaction steps.

It is thought that one purpose of RE3 (via P10 formation) is to help bind the bridge-3′ exon intermediate after the first reaction step (*4*, *7*, *8*, *16*, *17*). If this is true, then it might be

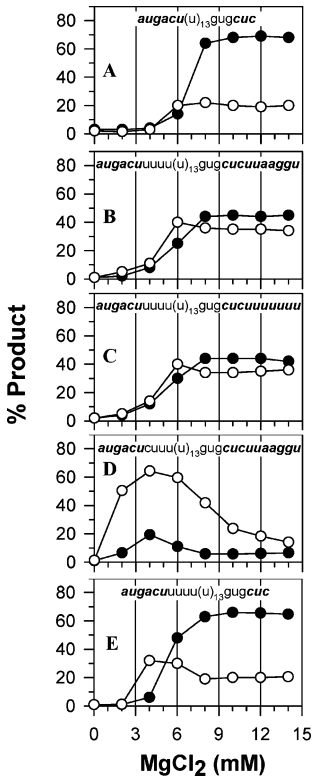

FIGURE 7: Magnesium dependence of TES reactions using rP-8/4x with various substrates. Reactions were conducted at 44 °C for 1 h using 166 nM ribozyme and 1.3 nM 5′-end-radiolabeled substrate at the indicated magnesium concentration. The open circles represent the 5′ exon intermediate, and the closed circles represent the TES product. The substrates are shown in the graphs and in the corresponding ribozyme diagrams. The data points are the average of two independent assays.

possible to compensate for the loss of P10 by using ribozymes that can form elongated P9.0 helices. To test this, we synthesized the ribozyme rP-8/4x-noP10-10A, which lacks RE3 but contains an extended RE2 of 10 adenosines. Figure 3 shows that rP-8/4x-noP10-10A can generate the complete TES product, but to a smaller extent (15%) than rP-8/4x (30%). Apparently, elongating P9.0 does permit the second step of the TES reaction to occur in the absence of P10, but elongating P9.0 is not as effective as even a relatively short P10 helix.

The Sequence of the Bridging Region and the Presence of Substrate Nucleosides Immediately Downstream from P10 Can Inhibit the Second Step of the TES Reaction. To determine why the 36-mer substrate, in particular, results in relatively low second-step TES yields, a series of magnesium-dependent TES reactions were conducted, each of which utilize both the unmodified rP-8/4x ribozyme and a substrate that is a sequence variant of the 36-mer. When both the r(CACC) and r(*UUAAGGU*) regions of the 36-mer substrate are omitted, the TES reaction improves substantially (from 30% in Figure 2A to 70% in Figure 7A), even at high magnesium concentrations, and particularly in terms of the second reaction step. This result shows that there is nothing fundamentally inhibitory regarding the native P1, P9.0, and P10 helix sequences.

When the r(CACC) region of the excised fragment is replaced with r(UUUU), the amount of substrate that undergoes the first step (defined as the amount of 5′ cleavage product plus the amount of TES product) does not change; however, the yield of the second step increases (from 30% in Figure 2A to 45% in Figure 7B). In addition, product formation is no longer inhibited at high magnesium concentrations. TES reactions using a substrate where r(CUUU) [Figure 7D] replaces r(CACC) [Figure 2A] are very similar for the second reaction step, and both are very different from TES reactions utilizing the r(UUUU)-containing substrate [Figure 7B]. When taken together, these studies isolate second-step inhibition in the presence of the 5′ C in the substrate region to be excised. This 5′ C could form a Watson−Crick base pair with the 3′ G of RE3, which could result in the formation of an extended P1 helix (called P1ex). Such P1ex helices have been shown to be inhibitory in other contexts with group I intron-based catalytic RNAs (*11*−*14*).

Using substrates that have r(UUUU)-containing excised regions, we isolated the inhibition potential of the 3′ exon region on the TES reaction. When the seven 3′ nucleosides of the 36-mer substrate (Figure 7B) are replaced with seven uridines (Figure 7C) or adenosines (data not shown), the magnesium dependence of all the TES reactions shows maximum second step yields no greater than 45%. Evidently, the product yield is largely independent of the sequence of this downstream region. To test whether the mere presence of the downstream region is inhibitory, we analyzed a substrate where just the 3′ exon region was deleted. Note that this substrate also does not contain r(CACC). The product yield went from 45% in the presence of a 3′ exon region (parts B and C of Figure 7) to 70% in its absence (Figure 7E). Apparently, the mere presence of a 3′ exon region inhibits the second step of the TES reaction.

Note that the extent of second-step reactivity for the ribozymes with long P9.0 regions is the same as that seen when using the unmodified ribozyme and the 25-mer substrate (in which case both the r(CACC) segment and sequences downstream from P10 are omitted). Apparently, lengthening P9.0 can simultaneously overcome the inhibition because of both of these components.

The Bridging Region of the Substrate Can Inhibit the TES Reaction by Forming a P1ex Helix. In accordance with the presupposition that the 5′ end of the excised region of TES substrates can base pair with RE3 and compete with the 3′ exon for binding RE3 (*11*−*14*), we analyzed a series of modified ribozymes that are expected to form stability-enhanced P1ex or P10 helices *when paired with the 36-mer substrate*. By modulating the strength of these base-pairing elements, we can deduce whether TES inhibition likely results from P1ex formation interfering with P10 formation.

We first analyzed whether increasing the strength of P1ex decreases second-step reactivity. The ribozyme rP-8/4x-P1ex, which contains an extended IGS region that is completely complementary to the r(CACC) segment of the substrate, can potentially form a four base pair P1ex helix with the 36-mer substrate instead of the one or two base pair P1ex helix that forms with the unmodified ribozyme. The results (Figure 2D) show that rP-8/4x-P1ex is prohibited from catalyzing the second step of the TES reaction (the first step is reduced approximately 2-fold; compare with Figure 2A). Therefore, enhancing the stability of P1ex can completely inhibit the second reaction step of the TES reaction.

Conversely, we synthesized ribozymes that are expected to thermodynamically favor P10 over P1ex formation. These ribozymes are expected to restore TES activity by increasing

the ability of P10 to form and reducing the ability of P1ex to form (part B and C of Figure 2). With each of the modified ribozymes, approximately 50% TES product is formed, compared with a maximum of 30% for the native ribozyme sequence. Therefore, enhancing the stability of P10 can overcome the inhibitory aspects of P1ex formation. In addition, the extent of the reaction using elongated P10 helices is similar to that seen when r(CACC) is replaced with r(UUUU) (Figure 7B), which can also be expected to alter P1ex formation.

DISCUSSION

In the present work, we identify components of a substrate that can lead to the inhibition of the TES reaction. Specifically, these are the bridging region of the substrate, which can form a P1ex helix with the IGS of the ribozyme, and the mere presence of nucleobases immediately downstream from P10. We rationally engineered ribozymes that effectively overcome these inhibitory interactions. In the process, we have gained insights into the mechanism by which the TES reaction (and self-splicing) occurs for the *P. carinii* intron.

P1ex. It has previously been reported for self-splicing group I introns (and their derived ribozymes) that the RE3 region of the IGS can base pair with the 5′ end of the intron, as well as 3′ exons (*11*−*14*). In this way, P10-helix formation is disrupted because the IGS can also form an extended P1 helix. The results of our study confirm that such competition can occur in the *P. carinii* ribozyme and that it can occur in the context of the TES reaction.

The inhibitory effects of P1ex formation can be quite acute; a P1ex of only one Watson−Crick base pair can inhibit the reaction. When P1ex is elongated to four base pairs, the extent of the first reaction step is cut in half, while the extent of the second reaction step is completely eliminated. The reduction of the first reaction step necessarily leads to a reduction of the second step. Independent of this, however, if P1ex is too strong, P1ex would not dissociate after the first step, preventing P10 binding and interfering with or eliminating second-step reactivity.

The fact that a relatively long P1ex helix inhibits both the first and second reaction steps, which are analogous to the steps of the self-splicing reaction, indicates that P1ex helices are usually short in group I introns to circumvent this inhibition. Nevertheless, P1ex helices are common so they likely perform a beneficial biological role, perhaps by reducing premature P10-helix formation or by preventing the premature structural activation of the 5′ splice junction (for hydrolysis, for example).

P10. Previous reports have shown that, while P10 formation is not required for self-splicing reactions (*19*−*22*), its presence is beneficial for the second reaction step, both to enhance binding of the 3′ exon (*4*, *7*, *8*, *16*, *17*) and to reduce cryptic 3′ splice sites (*23*). The results of our study show that the presence of P10 is very beneficial for the effective execution of the second reaction step in the TES reaction. Note that we do not observe an increase in cryptic 3′ splice sites using any of the ribozymes outlined in this paper.

The amount of TES substrate that undergoes the first reaction step is independent of whether P10 is 0, 3, 5, or 10 base pairs in length. This suggests that either P10 formation occurs after the first reaction step or that P10 formation occurs before the first step, yet it does not affect the first reaction step enough (for example, in terms of rate changes) to be detected in this assay (i.e., using a reaction endpoint of 1 h). The latter seems unlikely, because the P10 helix would be directly adjacent to the P1-helix 5′ cleavage site and so its relative strength would be expected to alter this reaction step appreciably.

The amount of substrate that undergoes the second reaction step is dependent on the presence of a P10 helix. No TES product forms when using a ribozyme with a two base pair P9.0 and no RE3, showing how critical P10 formation is in the reaction. Similar results have been shown for the self-splicing reaction using a *Saccharomyces cerevisiae* intron (*17*), which is in contrast to results reported from a *Tetrahymena* intron (*19*). When P10 is present, its length influences the yield of the second step of the TES reaction (Figure 2). Interestingly, these magnesium-dependent product curves are basically superimposable, except for the high-magnesium inhibition that is occurring in the cases with shorter P10 helices. Apparently, the second reaction step is inhibited at high magnesium by something for which the stepwise elongation of P10 can progressively overcome. On the basis of our results and previous reports (*11*−*14*), it is likely that this inhibition is due to high magnesium P1ex formation interfering with P10 formation at shorter P10 lengths.

The fact that elongating P10 from 5 to 10 base pairs has such little affect on the reaction suggests that either the ribozyme cannot form long P10 helices or that the stability gained from the inclusion of these extra base pairs (above 5) does not have a significant effect on the reaction. For example, P10 formation could be helping to stabilize 3′ exon binding to the ribozyme but might not be helping to productively dock ωG into the G-binding site. The fact that relatively long P10 helices do not significantly enhance this reaction correlates with the fact that most native group I intron P10 helices are relatively short in length (less than 4 base pairs). With regard to the self-splicing reaction, a strong P10 would be undesirable because it would likely lead to premature docking of ωG before the first reaction step, as well as poor substrate dissociation.

Evidently, similar to that seen in the self-splicing reaction (*11*−*14*), P1ex and P10 are dynamic and compete with one another during each step of the TES reaction. Note, however, that in our studies lacking RE3, in which case both P10 and P1ex are not permitted, the second step of the TES reaction (only) goes poorly. Apparently, P10 does not exist exclusively to outcompete P1ex; P10 formation is beneficial for the TES reaction even when P1ex formation is not possible.

P9.0. Although P9.0 in most group I introns is quite short, commonly being only two base pairs, its presence acts to enhance the stability and specificity of binding the bridge-3′ exon to the ribozyme in the self-splicing reaction, thereby playing an important role in defining the 3′ splice site (*3*−*6*, *8*, *15*, *24*−*27*). Note that, in the self-splicing reaction, P9.0 is intramolecular with regard to the ribozyme. For the TES reaction, P9.0 represents the ribozyme binding the excised region of the target. Our results show that P9.0 is a beneficial component of the TES reaction in instances where relatively lengthy substrate bridge regions are excised. Moreover, we show that P9.0 can be elongated appreciably for the purpose of enhancing the second step of the TES reaction.

The amount of substrate that undergoes the first reaction step of TES is independent of the length of P9.0. The rate of the first step, however, is slower for systems with longer P9.0 lengths. The origin of this effect could be that the actual rate of chemistry decreases or that the folding rate of the ribozyme−substrate complex into its initial "active" conformation decreases.

The amount of substrate that maximally undergoes the second reaction step is dependent on the length of P9.0. The fact that second-step reactivity increases with increasing P9.0 length and that P9.0 acts in part to stabilize the bridge-3′ exon region of the TES substrate suggests that unproductive dissociation of the bridge-3′ exon region can be overcome by strengthening this molecular interaction. Note, however, that a threshold for product formation is reached using a total P9.0 length of around eight base pairs. This could be due to any number of factors, including the inability of more than 10 bases of RE2 to pair with the bridge of this particular substrate (perhaps because of structural constraints). Elongating P9.0 also enhances the rate of the second step of the reaction. A likely explanation is that the extra stability in P9.0 is acting to preorganize the bridge-3′ exon in the catalytic site for second-step reactivity, perhaps by driving ωG to bind the catalytic core of the ribozyme (*8*).

The stability increase observed upon binding the 3′ reaction intermediate to P9.0-elongated ribozymes is less than predicted (*18*). This is also true when adding an additional base pair to P9.0 in a *Tetrahymena* ribozyme (*15*), in which case the stability gain might be attributed primarily to an increase in base stacking (*6*). In our cases that could be measured, the stability due to P9.0 increases more than linearly with the stepwise addition of each nearest neighbor pair. For example, lengthening P9.0 from 8 to 10 base pairs enhances the stability by 0.6 kcal/mol; lengthening from 10 to 12 base pairs enhances the stability by 1 kcal/mol; and lengthening from 12 to 14 base pairs enhances the stability by 1.5 kcal/mol. This is compared with a predicted increase for each doublet addition of approximately 1.2 kcal/mol at 22 °C (*18*). There are many possible explanations for this result; however, the outcome is that P9.0 becomes progressively more stable and thus less transient, as P9.0 gets longer, thus allowing less bridge-3′ exon intermediate dissociation and producing more TES product.

Our results suggest that, with regard to group I introns and their derived ribozymes, (1) P9.0 is not optimized for thermodynamic stability, (2) P9.0 is not just two base pairs because of an overwhelming steric constraint, and (3) the distance between P9 and ωG is not an important determinant of the 3′ splice site position. The fact that long P9.0 helices are not prevalent in nature suggests that the benefit of longer P9.0 helices is outweighed by a detrimental factor. Stable P9.0 helices might slow the rate of the first reaction step (as outlined above), or the intron might bind the bridge-3′ exon region (and ωG) before the first reaction step (*6*), thus decreasing first-step reactivity.

General Considerations. Our results also suggest that P9.0 and P10 play somewhat different roles in stabilizing and orienting the bridge-3′ exon for subsequent catalysis. The primary role of P9.0 appears to be stabilization of the bridge-3′ exon, as ribozymes with more stable P9.0 helices catalyze the TES reaction to greater extents. Strengthening P9.0 considerably, however, cannot completely overcome the lack of P10 in the second step of the TES reaction, which is relatively weak to begin with. Therefore, it appears that the role of P10 is not simply to stabilize the binding of the bridge-3′ exon intermediate. It could exist primarily to spatially orient the exon for catalysis. Nevertheless, elongating P9.0 does eliminate P1ex formation at high magnesium, probably by acting synergistically with P10 to support the formation of P10. Therefore, P9.0 and P10 do have overlapping roles.

We have previously shown (*1*) that excising single nucleotides is a relatively efficient process (70% yield). A possible explanation is that P1 and P10 (without P9.0) are situated favorably to act in concert in binding ωG into the G-binding site. In other words, ωG is at least partially preorganized for the second reaction step in terms of it binding the G-binding site. In systems where longer RNA regions are excised, attaining the same level of preorganization can be compounded by the fact that additional inhibitory factors come into play. For example, P1ex can occur, the number of possible bound substrate conformations increases, and the catalytic core of the ribozyme has to internalize a longer excised region. The result is that increasing the stability of binding the bridge-3′ exon intermediate to the ribozyme might regularly be required to overcome the inhibitory factors associated with longer excised regions. We have shown that elongating P9.0 can provide this extra stabilization.

Note that we have previously shown that TES reactions in a myotonic dystrophy model system, in which 28 nucleotides are excised, produce up to 70% of the TES product yield (*1*). Such relatively high product yield could be connected to the fact that the myotonic dystrophy substrate is not expected to form a P1ex helix and only contains a one base 3′ exon (as opposed to a seven base pair 3′ exon in this report). Nevertheless, attempts were made to enhance this TES reaction by separately elongating P9.0 and P10 in the myotonic dystrophy system. The results indicate, however, that elongating P9.0 and P10 does not increase TES product yields in this system, nor does it reduce the magnesium concentration requirement for the reaction (data not shown). Apparently, elongating P9.0 and P10 does not enhance an already effective reaction (i.e., one in which the inhibitory substrate components are absent). Thus, the usefulness of elongating P9.0 and P10, as a reactivity enhancement strategy, needs to be assessed on a case-by-case basis.

Design Principles for TES Reactions. For effective TES reactions to occur, minimizing P1ex formation while maintaining P10 is important. Because P1ex inhibits the reaction and a strong P1ex helix cannot be effectively overcome by elongating P10, the target sequence immediately downstream from the 5′ splice site (in the region to be excised) should be chosen such that it does not base pair with the 5′ end of RE3 in the IGS. Although P10 elongation has been useful for in vivo TES reactions using *Tetrahymena* ribozymes (*28*, *29*), P10 lengths greater than five in our model system were not particularly beneficial for reactivity. Whether elongating P10 enhances the specificity of TES reactions remains to be determined. We found that elongating P9.0 can be very beneficial for the TES reaction. It slows down the rate of the first step, which is expected to enhance specificity (*30*, *31*), and speeds up the second step, thus permitting less time

for bridge-3′ exon intermediate dissociation. Reduced intermediate dissociation improves the reaction substantially. This is important because both steps of the reaction have to occur for TES to be successful. Note, however, that elongating P9.0 is only a useable design principle when excising RNA regions of appreciable length. Nevertheless, in cases where elongating P9.0 is possible, it can help the reaction by reducing the negative effects associated with both P1ex formation and the presence of nucleosides immediately downstream from P10, which will always be present in transcript targets. Thus, in exploiting the adaptability of P9.0 and P10, the potential for success in TES reactions can be substantially increased.

ACKNOWLEDGMENT

The authors thank Dana Baum for critical reading of the manuscript and all members of the Testa lab for helpful discussions and technical assistance.

SUPPORTING INFORMATION AVAILABLE

Time-dependent graphs of representative TES reactions (from reactions shown in Figures 4C and 7D). This material is available free of charge via the Internet at http://pubs.acs.org.

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. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) A *Pneumocystis carinii* group I intron that does not require 2′ OH groups on its 5′ exon mimic for binding to the catalytic core, *Biochemistry 36*, 15303−15314.
3. Burke, J. M. (1989) Selection of the 3′-splice site in group I introns, *FEBS Lett. 250*, 129−133.
4. Michel, F., Hanna, M., Green, R., Bartel, D., and Szostak, J. W. (1989) The guanosine binding site of the Tetrahymena ribozyme, *Nature 342*, 391−395.
5. Burke, J. M., Esherick, J. S., Burfeind, W. R., and King, J. L. (1990) A 3′ splice site-binding sequence in the catalytic core of a group I intron, *Nature 344*, 80−82.
6. Russell, R., and Herschlag, D. (1999) Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme, *RNA 5*, 158−166.
7. Karbstein, K., Caroll, K. S., and Herschlag, D. (2002) Probing the *Tetrahymena* Group I Ribozyme Reaction in Both Directions, *Biochemistry 41*, 11171−11183.
8. Karbstein, K., and Herschlag, D. (2003) Extraordinarily slow binding of guanosine to the *Tetrahymena* group I intron ribozyme: Implications for RNA preorganization and function, *Proc. Natl. Acad. Sci. U.S.A. 100*, 2300−2305.
9. Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990) Direct measurement of oligonucleotide substrate binding to wild-type and mutant ribozymes from Tetrahymena, *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187−8191.
10. Johnson, A. K., Baum, D. A., Tye, J., Bell, M. A., and Testa, S. M. (2003) Molecular recognition properties of IGS-mediated reactions catalyzed by a *Pneumocystis carinii* group I intron, *Nucleic Acids Res. 31*, 1921−1934.
11. Fernandez, A. (1992) On how hydrolysis at the 3′ end is prevented in the splicing of a sequentially folded group I intron, *FEBS Lett. 297*, 201−204.
12. Ritchings, B. W., and Lewin, A. S. (1992) Mutational evidence for competition between the P1 and the P10 helices of a mitochondrial group I intron, *Nucleic Acids Res. 20*, 2349−2353.
13. Shaw, L. C., Thomas, J., and Lewin, A. S. (1996) The Cbp2 protein suppresses splice site mutations in a group I intron, *Nucleic Acids Res. 24*, 3415−3423.
14. Guo, F., and Cech, T. R. (2002) In vivo selection of better self-splicing introns in *Escherichia coli*: The role of the P1 extension helix of the *Tetrahymena* intron, *RNA 8*, 647−658.
15. Moran, S., Kierzek, R., and Turner, D. H. (1993) Binding of Guanosine and 3′ Splice Site Analogues to a Group I Ribozyme: Interactions with Functional Groups of Guanosine and with Additional Nucleotides, *Biochemistry 32*, 5247−5256.
16. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982) Making ends meet: A model for RNA splicing in fungal mitochondria, *Nature 300*, 719−724.
17. Partono, S., and Lewin, A. S. (1990) Splicing of COB intron 5 requires pairing between the internal guide sequence and both flanking exons, *Proc. Natl. Acad. Sci. U.S.A. 87*, 8192−8196.
18. Xia, T., Santa Lucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs, *Biochemistry 37*, 14719−14735.
19. Been, M. D., and Cech, T. R. (1985) Sites of circularization of the Tetrahymena rRNA IVS are determined by sequence and influenced by position and secondary structure, *Nucleic Acids Res. 13*, 8389−8408.
20. Price, J. V., and Cech, T. R. (1988) Determinants of the 3′ splice site for self-splicing of the Tetrahymena pre-rRNA, *Genes Dev. 2*, 1439−1447.
21. Cech, T. R. (1990) Self-splicing of group I introns, *Annu. Rev. Biochem. 59*, 543−568.
22. Davies, R. W., Waring, R. B., and Towner, P. (1987) Internal guide sequence and reaction specificity of group I self-splicing introns, *Cold Spring Harbor Symp. Quant. Biol. 52*, 165−171.
23. Suh, E. R., and Waring, R. B. (1990) Base Pairing between the 3′ Exon and an Internal Guide Sequence Increases 3′ Splice Site Specificity in the *Tetrahymena* Self-Splicing rRNA Intron, *Mol. Cell. Biol. 10*, 2960−2965.
24. Tanner, N. K., and Cech, T. R. (1987) Guanosine binding required for cyclization of the self-splicing intervening sequence ribonucleic acid from *Tetrahymena thermophila*, *Biochemistry 26*, 3330−3340.
25. Michel, F., Netter, P., Xu, M.-Q., and Shub, D. A. (1990) Mechanism of 3′ splice site selection by the catalytic core of the *sun*Y intron of bacteriophage T4: the role of a novel base-pairing interaction in group I introns, *Genes Dev. 4*, 777−788.
26. Berzal-Herranz, A., Chowrira, B. M., Polsenberg, J. F., and Burke, J. M. (1993) 2′Hydroxyl Groups Important for Exon Polymerization and Reverse Exon Ligation Reactions Catalyzed by a Group I Ribozyme, *Biochemistry 32*, 8981−8986.
27. Bevilacqua, P. C., Sugimoto, N., and Turner, D. H. (1996) A Mechanistic Framework for the Second Step of Splicing Catalyzed by the *Tetrahymena* Ribozyme, *Biochemistry 35*, 648−658.
28. Kohler, U., Ayre, B. G., Goodman, H. W., and Haseloff, J. (1999) *Trans*-splicing Ribozymes for Targeted Gene Delivery, *J. Mol. Biol. 285*, 1935−1950.
29. 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.
30. Herschlag, D. (1991) Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: more isn't always better, *Proc. Natl. Acad. Sci. U.S.A. 88*, 6921−6925.
31. Zarrinkar, P. P., and Sullenger, B. A. (1999) Optimizing the Substrate Specificity of a Group I Intron Ribozyme, *Biochemistry 38*, 3426−3432.

BI035874N