RNA (2002), *8*:647–658. Cambridge University Press. Printed in the USA.

DOI: 10.1017.S1355838202029011

In vivo selection of better self-splicing introns in *Escherichia coli*: The role of the P1 extension helix of the *Tetrahymena* intron

FENG GUO and THOMAS R. CECH

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA

ABSTRACT

In vivo selection was used to improve the activity of the *Tetrahymena* pre-rRNA self-splicing intron in the context of heterologous exons. The intron was engineered into a kanamycin nucleotidyltransferase gene, with the pairing between intron bases and the 5′ and 3′ splice sites maintained. The initial construct failed to confer kanamycin resistance on *Escherichia coli*, although the pre-mRNA was active in splicing in vitro. Random mutation libraries were constructed to identify active intron variants in *E. coli*. All the active mutants sequenced contained mutations disrupting a base-paired region above the paired region P1 (referred to as the P1 extension region or P1ex) that involves the very 5′ end of the intron. Subsequent site-directed mutagenesis confirmed that these P1ex mutations are responsible and sufficient to activate the intron splicing in *E. coli*. Thus, it appears that too strong of a secondary structure in the P1ex element can be inhibitory to splicing in vivo. In vitro splicing assays demonstrated that two P1ex mutant constructs splice six to eight times faster than the designed construct at 40 μM GTP concentration. The relative reaction rates of the mutant constructs compared to the original design are further increased at a lower GTP concentration. Possible mechanisms by which the disrupted P1ex structure could influence splicing rates are discussed. This study emphasizes the value of using libraries of random mutations to improve the activity of ribozymes in heterologous contexts in vivo.

Keywords: catalytic RNA; group I intron; kanamycin nucleotidyltransferase; ribozyme; splicing rate enhancement

INTRODUCTION

The group I introns, the family to which the *Tetrahymena* intron belongs, catalyze their own splicing in two sequential transesterification reactions to produce accurately ligated exons and a spliced intron with an exogenous guanosine (G) linked at its 5′ end (Cech, 1990). In the past two decades, extensive genetic and biochemical studies have elucidated the essential elements of this group of self-splicing RNAs or ribozymes. A 5′ region of the intron referred to as the internal guide sequence (IGS), which base pairs with the 5′ exon to form the paired region 1 (P1), is responsible for the recognition of the 5′ exon sequence (Been & Cech, 1986; Waring et al., 1986). A GU wobble base pair at the 5′ splice site is required for the cleavage reaction (Strobel & Cech, 1995). In the first step of the splicing reaction, G or one of its 5′ phosphorylated derivatives binds in a specific site on the ribozyme (Bass & Cech, 1984; Michel et al., 1989) and cleaves at the 5′ splice site to become covalently attached at the 5′ end of the intron (Cech et al., 1981). In the second step, the 3′ hydroxyl of the 5′ exon, which remains bound to the intron, attacks the phosphodiester linkage at the 3′ splice site to form the products.

In recent years, a number of studies have addressed the question of how to make such ribozymes work better in different contexts. For example, in vitro selection was used to identify mutations that activate the *Tetrahymena* ribozyme when Ca^{2+} is substituted for Mg^{2+} in the solution (Lehman & Joyce, 1993) or that increase the folding rate of its P3–P7 domain (Treiber et al., 1998). Small insertions and deletions in the nonconserved joint between paired regions P1 and P2 (re-

Reprint requests to: Thomas R. Cech, Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA; e-mail: Thomas.Cech@colorado.edu.

ferred to as J1/2) were found to enhance the turnover number and sequence specificity of ribozyme-catalyzed cleavage by decreasing the binding affinity of the RNA substrate (Young et al., 1991). A systematic approach based on biochemical knowledge was successfully used to optimize the substrate specificity in a *trans*-splicing reaction of the *Tetrahymena* ribozyme (Zarrinkar & Sullenger, 1999).

The IGS is an essential element for exon recognition (Davies et al., 1982; Cech, 1990), but a complete understanding of how to engineer a group I intron into an arbitrary exon context to activate the splicing reaction is lacking. For example, the work of Woodson has illustrated how different exon structures can affect splicing and reverse splicing in vitro and in vivo in *Escherichia coli* (Woodson & Cech, 1991; Woodson & Emerick, 1993; Roman & Woodson, 1995, 1998). Furthermore, the *Tetrahymena* intron inserted in a luciferase gene gives a 10-fold range of splicing activity in mammalian cells depending on the location of the insertion (Hagen & Cech, 1999). In the present work, we show how a designed insertion of the *Tetrahymena* intron in a kanamycin resistance gene that is initially inactive in *E. coli* can be rescued by in vivo selection for active mutants. In the process, we reveal the importance of a helical extension of the P1 helix, which is referred to as the P1 extension (or P1ex).

RESULTS

In vivo selection of active intron variants

In an attempt to select for intron variants that splice at high temperature in thermophilic bacteria, the *Tetrahymena* group I intron was inserted into a thermally stable kanamycin nucleotidyltransferase (KNT) gene (Matsumura et al., 1984; Liao et al., 1986) at position 119 downstream of the translational start site to produce Tet-119. In the design of Tet-119, bases 15–20 within the IGS (bases 15–27) were modified to base pair with the 3′ exon sequence (the resulting helix is referred to as paired region 10, or P10), as the wild-type intron does with the large subunit rRNA 3′ exon (Fig. 1A,B). At the same time, nt 3 and 4 near the 5′ end of the intron were altered to maintain pairing in the extension of the P1 helix (P1ex). This pre-mRNA construct was tested and found to be competent for RNA splicing in vitro (data not shown). Nevertheless, Tet-119 conferred kanamycin resistance (Kan^R) neither on the thermophile *Thermus thermophilus* nor on *E. coli*.

A kinetic RT-PCR experiment (measurement of amount of the PCR product as a function of PCR cycle number; Wiesner, 1992) was performed to check for any effect of the intron insertion on steady-state mRNA level. The results showed that *E. coli* cells transformed

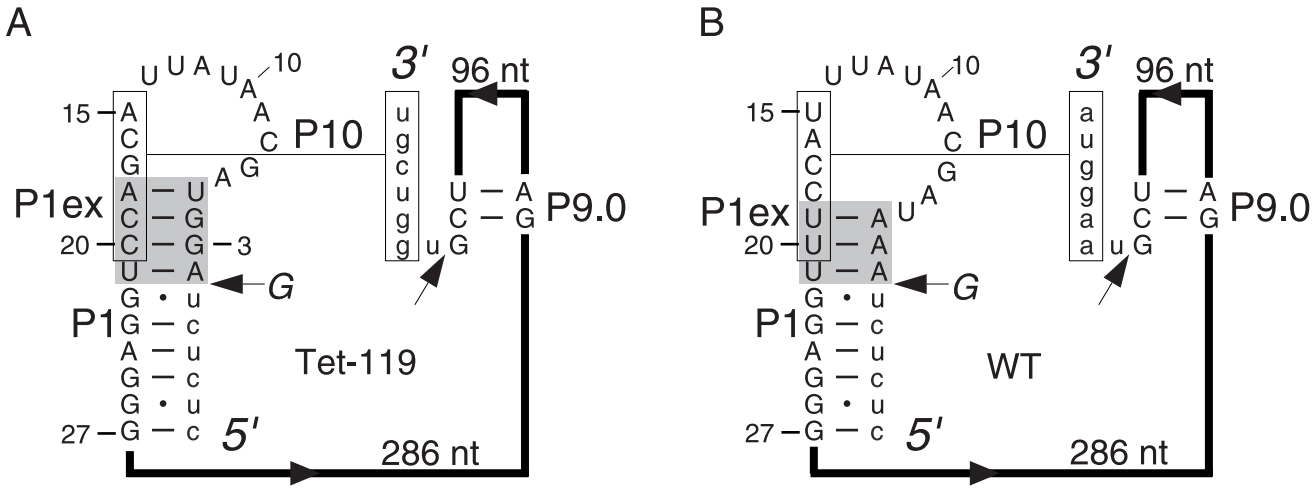

FIGURE 1. Design of the *Tetrahymena* intron insertion Tet-119 into the KNT gene (**A**) and comparison to the natural intron sequence (**B**). The exon sequences are in lower case. One arrow shows the site of attack by guanosine (*G*), and the other arrow indicates the 3′ splice site. The intron sequences of the P1 stem-loop, the P9.0, P10 helices and the 3′ terminal G residue are drawn in capital letters and the interconnecting nucleotides are represented by thick lines. Boxed nucleotides form the P10 helix for the second step of splicing. The shaded area represents the P1 extension (P1ex). **C:** The P1ex mutations identified in each active Tet-119 variant in the in vivo selection.

with Tet-119 contained similar amount of KNT mRNA as did those transformed with the plasmid with no intron insert (data not shown). Thus, the discrepancy between the in vivo and in vitro observations does not appear to be due to differential mRNA stability, but might instead be explained by a relatively slow splicing reaction of the intron. The majority of the pre-mRNA molecules could have been degraded before they underwent proper splicing, resulting in little intact KNT mRNA for subsequent protein synthesis and thus the lack of kanamycin resistance.

To test this explanation and possibly circumvent this problem, we chose to employ in vivo selection. Random mutations were introduced into the intron region using mutagenic PCR (Fromant et al., 1995) at an average level of two mutations per gene. The mutation pool (referred to as Tet-119mut) that contained the unmodified exon sequences was generated by a subsequent high-fidelity PCR reaction using two PCR-amplified exon pieces and the mutated intron as the template. The products of the second PCR were restriction digested and ligated into a pUC19-based vector to create the Tet-119mut library. This library was transformed into *E. coli* and plated on LB agar dishes containing ampicillin or various concentrations of kanamycin (Table 1A). The total number of transformants, which is an approximate estimate for complexity of the pool, was about 2×10^6.

The Tet-119mut library did give colonies on the kanamycin plates, but in much fewer numbers than an equivalent plating on an ampicillin plate (Table 1A). This result suggested that only a small fraction of the pool was active in splicing in vivo. As expected for such a selection experiment, the number of colonies decreased with increasing concentration of kanamycin. In contrast, a control mutation pool based on a splicing-inactive version of the *Tetrahymena* intron (Sullenger & Cech, 1994), Tet-119Δ237-330mut, gave very few transformants on the kanamycin plates (Table 1A). Restriction enzyme cleavage analysis revealed that the colonies on these plates represented a trace amount of contamination from the KNT gene with no intron insert, generated during construction of the mutation pool.

TABLE 1A. Plating results—number of colonies.[a]

	Kanamycin (μg/mL)					
Plasmid transformed	10	20	30	40	50	Ampicillin 100 μg/mL
Kan (no intron)	10^4	10^4	10^4	10^4	10^4	10^4
Tet-119	0	0	0	0	0	10^4
Tet-119mut	200	128	140	80	60	2×10^5
Tet-119Δ237-330mut	1	1	1	2	3	2×10^5
No DNA	0	0	0	0	0	0

[a]The Tet-119mut library, as well as several controls, were transformed into the *E. coli* strain XL-1Blue MRF′. One-tenth of the transformed cells were spread on each plate with the indicated concentrations of kanamycin or ampicillin.

Mutations in the P1 helix extension promote activity

To elucidate the type of mutations rendering the Tet-119 construct active, 10 colonies were cultured and their plasmids were purified and sequenced. The sequencing results showed that each plasmid contained two to four mutations in the intron region (Table 1B). Interestingly, one of the mutations in each plasmid was located in the helical region called P1ex (Table 1B; Fig. 1C). The rest of the mutations were spread throughout the intron. On the basis of previous work, it was not clear how these mutations could affect the splicing properties of the intron.

Previous studies have shown that the base pair interactions in P1 plus tertiary interactions involving its ribose–phosphate backbone are responsible for binding of the 5′ exon to the ribozyme core, setting up the first step of splicing (Been & Cech, 1986; Waring et al., 1986; Pyle et al., 1990; Szewczak et al., 1998). The last residue of the 5′ exon, a uracil, forms an essential G·U wobble pair with IGS, making a tertiary interaction with the ribozyme active site through the exocyclic amine of the G (Strobel & Cech, 1995). In contrast to the well-established role of P1, there has been little report on the P1 extension in the *Tetrahymena* intron. Phylogenetic analysis of group I introns revealed that this region is not conserved (Michel & Westhof, 1990). In fact, many group I introns do not have the P1 extension. Overall, our in vivo selection results suggested that too strong of a P1ex helix may inhibit splicing. This hypothesis is consistent with the fact that the P1ex element of Tet-119 contains 4 bp, with 2 of them GC pairs, whereas the wild-type *Tetrahymena* intron P1ex has only three AU pairs (Fig. 1). The additional interactions in Tet-119 might somehow render the splicing

TABLE 1B. Sequencing results of 10 colonies on the Tet-119mut/kanamycin (50 μg/mL) plate.[a]

M1	**G4A**, U221A, A263U, U412C
M2	**C20A**, C170U
M3	**U21A**, G368U
M4	**U21A**, U130C
M5	**G4A**, U221A, A263U, U412C
M6	**G3U**, C127U
M7	**U21A**, U348A, G388A
M8	**U21A**, U131A
M9	**G3A**, A269U, U291G, A339Δ
M10	**G3A**, C41U, U409C

[a]Plasmids from 10 colonies on the plate containing 50 μg/mL kanamycin were sequenced and the mutations in their intron region are listed. Each plasmid contained a mutation in the P1ex region (highlighted in bold).

of the designed construct slower than that of the wild-type rRNA intron.

To test this hypothesis, the five P1ex mutations observed (Fig. 1C) were individually introduced into the Tet-119 plasmid and tested for their ability to confer resistance to 50 μg/mL kanamycin on *E. coli*. All of these mutants yielded similar numbers of colonies on kanamycin and ampicillin plates, indicating that these mutations are sufficient to confer Kan^R. In addition, because random mutations were introduced throughout the intron during construction of the initial pool, the fact that P1ex is the only mutated region common to all the active molecules sequenced argues that it is the major modification that improves the splicing activity in vivo. Indeed, when the P1ex mutation in clone M1 containing four point mutations was reversed to the original base (A $\rightarrow$ G at position 4), the construct became inactive in *E. coli*. Therefore, these observations suggest that the P1ex mutations are both responsible and sufficient for in vivo splicing activity.

P1ex mutants splice faster in vitro

Although the P1ex mutations rendered the intron-containing KNT gene active in vivo, it was not clear whether this was achieved by directly accelerating intron splicing. One alternative explanation was that these mutations somehow increased the stability of the pre-mRNA from nucleolytic degradation. This possibility was ruled out by a kinetic RT-PCR experiment (Wiesner, 1992) showing similar mRNA levels in *E. coli* transformed with Tet-119 or either of two arbitrarily picked P1ex mutant constructs, G3A and C20A. Another alternative was that the mutations disrupted a binding site for an inhibitory factor in *E. coli*. We investigated this question by comparing these two P1ex mutants with Tet-119 in splicing assays in a protein-free solution.

In a pilot splicing experiment at 37 °C, the G3A mutant was shown to splice much faster than Tet-119 pre-mRNA in the context of a full-length 5′ exon and a truncated (205 bases) 3′ exon (data not shown). However, the splicing of Tet-119 was so slow that it was hard to obtain kinetic parameters. Therefore, the splicing temperature was raised to 50 °C to facilitate the comparison of the mutants with Tet-119. In addition, pre-mRNA with full-length KNT exons was used (Fig. 2A) to better mimic the splicing in vivo.

In the next set of splicing experiments, the pre-mRNA was uniformly labeled with ^{32}P during transcription and purified by denaturing gel electrophoresis. Splicing reactions were incubated at 50 °C and initiated

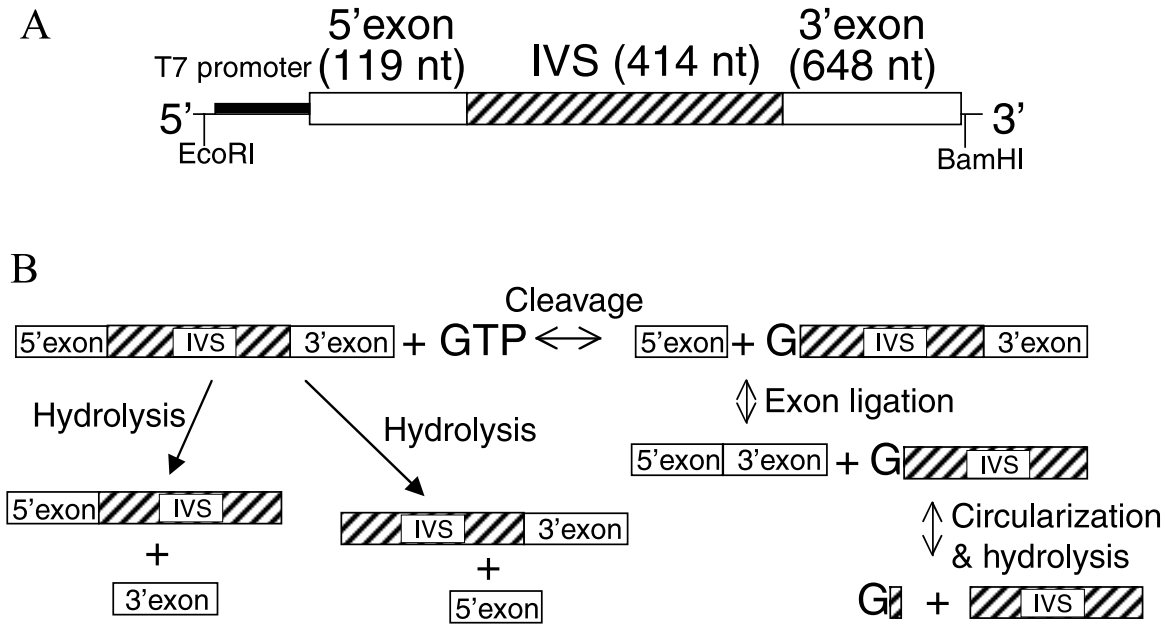

FIGURE 2. A: Constructs of Tet-119 and the P1ex mutants used in the T7 transcription of pre-mRNA. The number of nucleotides in each region of the transcribed pre-mRNA is indicated. **B:** The major reactions that normally occur in *Tetrahymena* intron self-splicing. In the productive pathway, GTP binds to the precursor RNA in a specific site of the intron and cleaves the 5′ exon-intron junction with its 3′ hydroxyl. After this cleavage step, the GTP becomes covalently attached to the 5′ end of the intron. In the subsequent exon ligation reaction, the 3′ hydroxyl of the 5′ exon attacks the intron–3′ exon linkage to form a ligated exon and an intron with the exogenous G attached at its 5′ end. The spliced intron can also undergo a circularization reaction in which its 3′-terminal G residue cleaves at an internal site near the 5′ end to produce a small stretch of 5′ RNA and a circular intron molecule. The newly formed phosphodiester bond in the circular intron can undergo hydrolysis to give a linear intron molecule of a slightly smaller size than that of the full-length intron. The two intron–exon junctions are also subject to a G-independent hydrolysis reaction to give several nonproductive species.

by adding unlabeled guanosine triphosphate (GTP). The pre-mRNAs disappeared over time, and the intensity of the bands corresponding to the ligated exons and the spliced intron increased for the mutants as well as the wild-type RNA. A number of other reaction intermediates and products also formed in the splicing reaction (Fig. 2B), some of which were identified and are indicated on the gel (Fig. 3A).

Quantification of the pre-mRNA as well as the ligated exons and the spliced intron bands clearly revealed that the two mutant constructs had increased splicing rates (Fig. 3B). Fitting of the amount of pre-mRNA remaining as a function of time with an exponential decay curve provided a k_{obs} of 0.06 min^{-1} for Tet-119 compared to 0.31 and 0.42 min^{-1} for G3A and C20A, respectively. The fitting of the ligated exon and spliced intron data also supported the conclusion that the reaction rates of the G3A and C20A mutants are approximately six and eight times, respectively, that of the original construct at a GTP concentration of 40 μM (Table 2A).

The rate differences are more significant at low GTP concentration

We subsequently examined how the GTP concentration affected the splicing reaction for Tet-119 and the P1ex mutants. The Michaelis constant for GTP (K_M^{GTP}) for the wild-type *Tetrahymena* rRNA intron was previously determined to be around 32 μM (Bass & Cech, 1986). The original Tet-119 pre-mRNA and the G3A and C20A mutants were therefore compared in

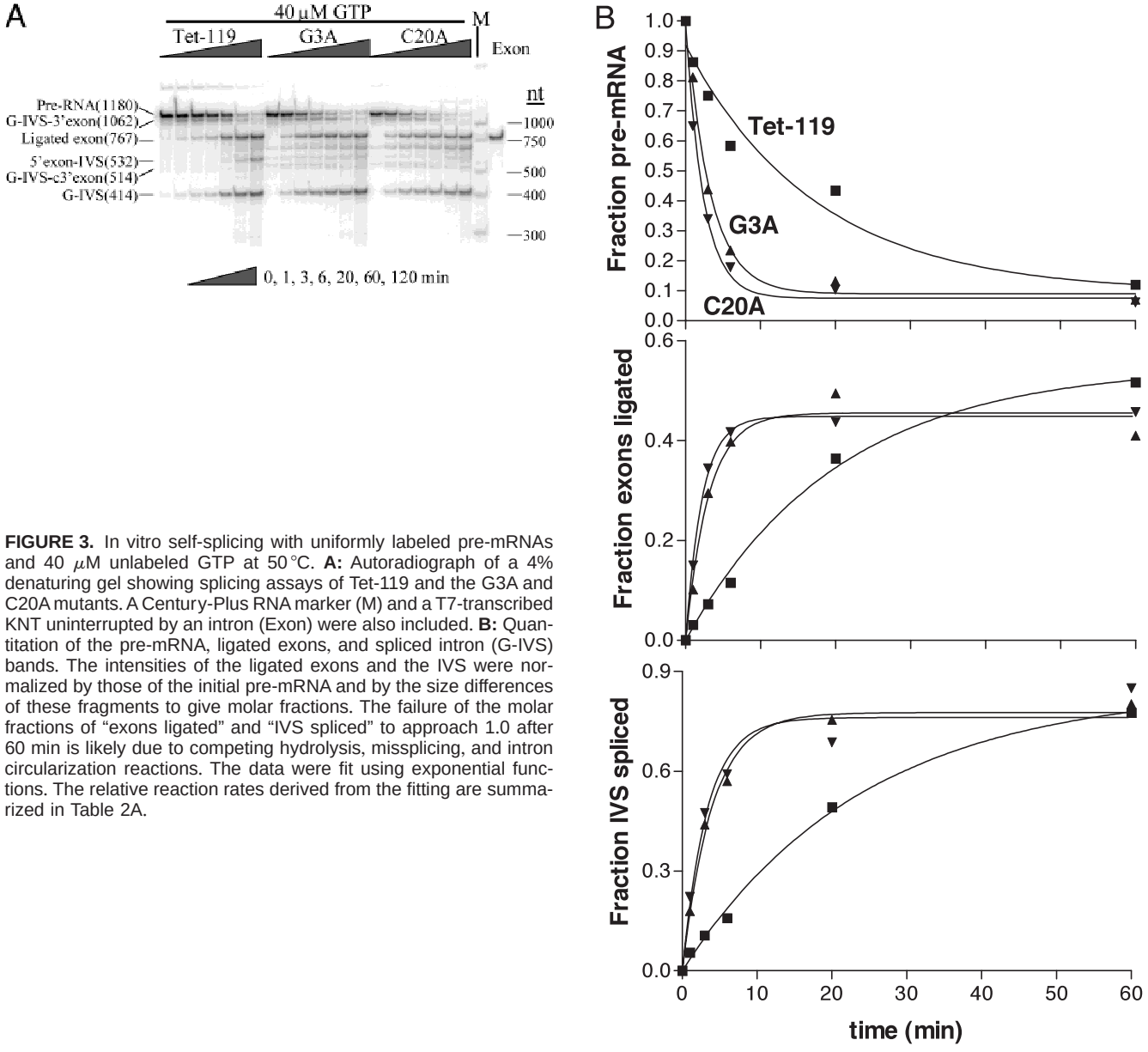

FIGURE 3. In vitro self-splicing with uniformly labeled pre-mRNAs and 40 μM unlabeled GTP at 50 °C. **A:** Autoradiograph of a 4% denaturing gel showing splicing assays of Tet-119 and the G3A and C20A mutants. A Century-Plus RNA marker (M) and a T7-transcribed KNT uninterrupted by an intron (Exon) were also included. **B:** Quantitation of the pre-mRNA, ligated exons, and spliced intron (G-IVS) bands. The intensities of the ligated exons and the IVS were normalized by those of the initial pre-mRNA and by the size differences of these fragments to give molar fractions. The failure of the molar fractions of "exons ligated" and "IVS spliced" to approach 1.0 after 60 min is likely due to competing hydrolysis, missplicing, and intron circularization reactions. The data were fit using exponential functions. The relative reaction rates derived from the fitting are summarized in Table 2A.

TABLE 2A. Relative reaction rates (k/k(Tet-119)) in splicing assays in vitro with uniformly labeled pre-mRNA and 40 μM GTP.

	Tet-119	G3A	C20A
Pre-mRNA	1.0[a]	5.4	7.3
Ligated exon	1.0	6.5	9.0
G-IVS	1.0	6.0	7.2

[a] k_{obs}(Tet-119, pre-mRNA) = 0.06 min^{-1}.

splicing assays at GTP concentrations of 250 nM, which is subsaturating, and 200 μM, which should be close to saturation.

Unlabeled pre-mRNA and [α-^{32}P]GTP were used in these assays, so only those species that became radiolabeled with the exogenous G during splicing were visible on the gel. The G-IVS-3′exon splicing intermediate and G-IVS product were the major bands for Tet-119, whereas an additional product at ~514 nt appeared for the two mutants (Fig. 4A,B). At 250 nM GTP, the rate differences between the mutants and the original design were very significant (Fig. 4C, Table 2B). The ratio of initial velocity values (v) was 1:28:67 (Tet-119:G3A:C20A) for G-IVS RNA accumulation. Splicing at 200 μM GTP shifted this ratio of v values to 1:9:18 (Table 2B). Therefore, the relative rate of the designed construct to the mutants was increased about threefold at high GTP concentration. On the other hand, the reaction rate of Tet-119 was still significantly slower than the P1ex mutants even at the high concentration of GTP, suggesting that the P1ex mutations increase k_{cat} while decreasing K_M^{GTP}.

As mentioned above, the two mutants produced an RNA species estimated to be about 514 bases long (Fig. 4A,B). Because this species was 5′-G labeled during splicing, we identified its 5′ end using enzymatic RNA sequencing (data not shown). Ribonucleases T1 and U2, as well as partial alkaline hydrolysis, were used to distinguish G and A residues (Donis-Keller et al.,

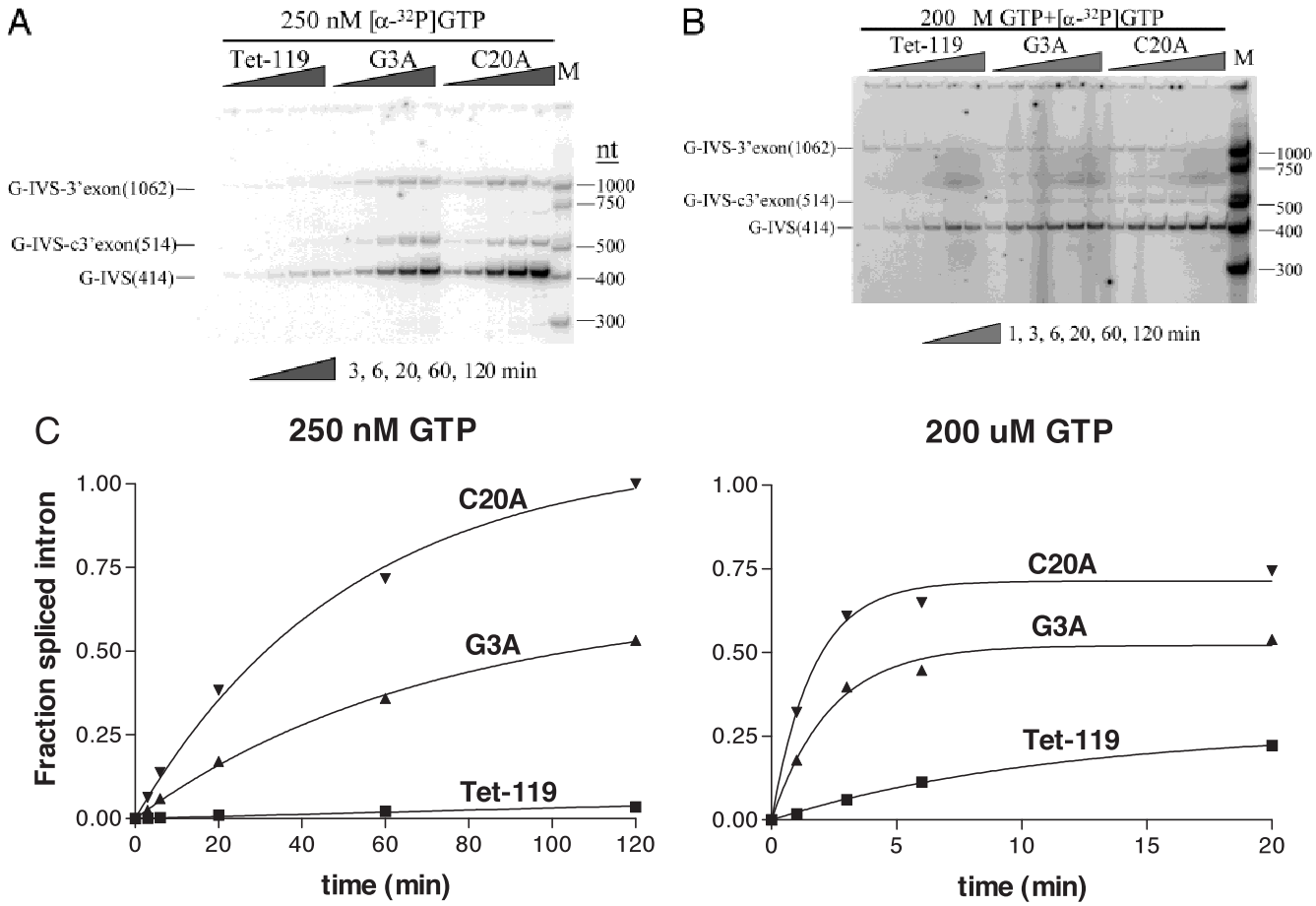

FIGURE 4. In vitro splicing assays with unlabeled pre-mRNAs and [α-^{32}P]GTP at 50 °C. **A:** 8 nM pre-mRNA and 250 nM [α-^{32}P]GTP were included in each reaction. **B:** Same as **A**, except that 200 μM unlabeled GTP were added in addition to the [α-^{32}P]GTP. Additional intense bands of radioactivity appeared at <100 nt in **B** and have not been identified. **C:** Accumulation of the spliced intron in the splicing assays at 250 nM and 200 μM GTP. The data were normalized to the highest amount observed for the C20A RNA in each assay condition and fit using exponential curves. The relative initial velocities are summarized in Table 2B.

TABLE 2B. Ratios of initial velocities (v/v(Tet-119)) of accumulation of the G-IVS in splicing assays in vitro with unlabeled pre-mRNA and [α-^{32}P]GTP.[a]

[GTP]	Tet-119	G3A	C20A
250 nM	1	28	67
200 μM	1	9	18

[a]Because of the absence of pre-mRNA bands in splicing assays with unlabeled pre-mRNA, the absolute catalytic rate constant k_{obs} could not be measured. However, the ratio of initial reaction velocities v of the mutants to Tet-119 could be used to examine the mutational effects on catalysis.

1977). This RNA had exactly the same sequence at its 5′ end as the G-IVS. Therefore, the extra 100 bases must be present at the 3′ end of RNA. Further sequence examination revealed that a stretch of 11 bases, which is located 101–110 bases downstream from the normal 3′ splice site, could potentially serve as a cryptic 3′ splice site. Thus, this G-IVS-c3′exon species appears to arise by a normal first step of splicing followed by attack of the 5′ exon at a cryptic 3′ splice site downstream from the normal 3′ splice site.

The 514-base 3′ cryptic splicing product was also observed in the splicing assays with ^{32}P-labeled pre-mRNA and unlabeled GTP at 40 μM concentration. Because only the two mutants gave rise to significant amounts of the misspliced product, we conclude that the P1ex mutations decrease the splicing specificity of the KNT pre-mRNA.

DISCUSSION

The problem we encountered in designing a self-splicing group I intron insertion into a heterologous gene such that it would be active in vivo is far from unique. Hagen and Cech (1999) observed large differences in splicing efficiency in mammalian cells depending on the intron insertion site used. In parallel work, the percentage of transcripts that underwent self-splicing, again in mammalian cells, was found to range from 0 to 50% (Long & Sullenger, 1999). In addition to low catalytic efficiency, low substrate specificity was shown to be a major limiting factor in the application of the group I ribozymes in therapeutic RNA repair via *trans*-splicing in vivo (Jones et al., 1996). Rational design based on the biochemical knowledge was used to optimize the ability of the *Tetrahymena* ribozyme to discriminate against nonspecific substrates in vitro (Zarrinkar & Sullenger, 1999). However, because of the great complexity of the system, especially that of the intracellular environment, this type of approach has quite variable success. In contrast, in vivo selection allows "irrational" screening of a large number of random mutations in parallel, treating the splicing problem in the cells as a "black box." We demonstrated, in the current study, that in vivo selection resulted in active intron variants that shared mutations in an unexpected region, the P1 extension. The direct role of the P1ex mutations in splicing activation was confirmed by subsequent in vivo and in vitro splicing assays.

The in vivo selection and in vitro characterization results clearly show that the paired region P1ex in the *Tetrahymena* intron can be significant for splicing rate determination. Previous work with the bI5 intron from yeast mitochondria demonstrated that a mutation disrupting the P1ex base pairs reduces the rate of cleavage at the 5′ splice site and blocks the exon ligation step of self-splicing in vitro (Ritchings & Lewin, 1992), opposite to the effect we observe. A possible explanation for this difference is that in contrast to the *Tetrahymena* intron, the bI5 intron contains a short P1 helix (4 bp up to the 5′ splice site) and a long stretch of sequence between the two P1 strands (more than 200 nt). A weakened P1ex interaction may prevent the bI5 intron core from finding the right base-pairing partner in the 5′ exon, and thus result in the reduced cleavage rate. In a second study on bI5 splicing, a mutant with the P1ex helix extended from 2 bp to 5 bp was unable to ligate the exons in either autocatalytic or Cbp2 protein-stimulated splicing (Shaw et al., 1996). This latter observation is consistent with our results that problems arise when P1ex is too strong.

Our observations are consistent with previous in vivo engineering results involving the *Tetrahymena* intron. In the earliest works, this intron was engineered into the α-donor fragment of the β-galactosidase gene and experiments showed that the predicted splicing products could be generated in *E. coli* (Price & Cech, 1985; Waring et al., 1985). In these cases, the intron was unmodified and grafted into the β-galactosidase gene together with small stretches of *Tetrahymena* rRNA exons at both ends. The spliced β-gal exons therefore had an unnatural amino acid insertion at a position that is not sensitive to insertions. When the *Tetrahymena* intron was inserted into the analogous position of *E. coli* 23S rRNA, the intron was shown to splice from pre-rRNA rapidly in *E. coli* (Zhang et al., 1995). Although the IGS was modified to base pair with the 5′ exon, the P1ex base pairs were not modified. In a later work, the *Tetrahymena* intron was implanted into a luciferase gene and splicing efficiency in mammalian cell lines was estimated (Hagen & Cech, 1999). In these designs, the spliced exons did not contain extra sequences. Although the IGS was modified to base pair with both the 5′ and 3′ exons, the intron sequence at the 5′ end was not modified, with the result that the P1ex pairing was disrupted.

How might disruption of P1ex enhance self-splicing?

How does the destabilization of the P1ex helical interaction in the *Tetrahymena* intron increase the splicing

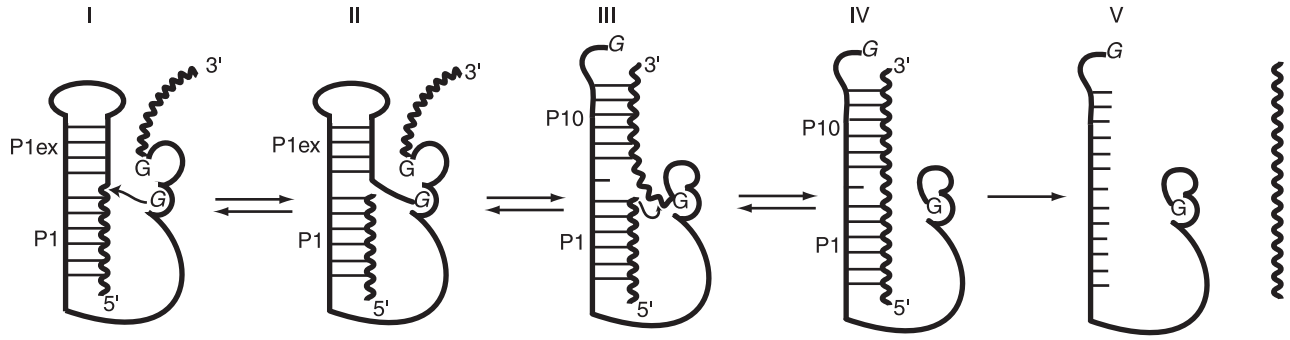

FIGURE 5. Schematics of the *Tetrahymena* intron self-splicing mechanism with the possible roles of the P1ex helix indicated. Wavy lines represent the 5′ and 3′ exons and thick lines stand for the intron. Base pairs are indicated by long thin bars and unpaired bases by short bars. We propose that a strong P1ex interaction could increase the rate for the reverse reaction between states I and II, in addition to decreasing the rate of conformation changes between states II and III. See text for details.

reaction rate? One likely possibility can be explained with reference to Figure 5. In the first step of splicing, an exogenous G (italicized in the figure) binds in a specific pocket of the intron (I) and cleaves the 5′ splice site using its 3′ hydroxyl to become covalently linked to the 5′ end of the intron (II). In order for the reaction to proceed from II to III, a conformational change must occur to the splicing complex. The exogenous G needs to dissociate from the pocket, which requires that any base pairs in P1ex be disrupted. Subsequently, the 3′ exon base pairs with the IGS to form the P10 helix, and the endogenous G residue at the 3′ end of the intron associates with the same binding site in the intron as the exogenous G (III). In the second step of splicing, the 3′ hydroxyl of the 5′ exon cleaves the 3′ splice site to form the ligated exons and the spliced intron (IV). To complete the splicing, the ligated exons need to dissociate from the spliced intron to prevent the reverse reaction of exon ligation (V). If the interaction in the P1ex region is too strong, the free energies of states I and II along the reaction coordinate are lowered and the reaction complex stays longer in state II before proceeding to state III. In state II, the 3′ hydroxyl of the 5′ exon could cleave the G-intron linkage as a reverse reaction of the first cleavage step. In addition, because the formation of the P1ex and P10 helices compete with each other, a strong P1ex could reduce the rate of the conformational change (II $\rightarrow$ III) required to initiate exon ligation.

It has been reported that folding of the *Tetrahymena* pre-rRNA into a catalytically active conformation can be rate determining for in vitro splicing at physiological temperatures (Emerick & Woodson, 1994; Emerick et al., 1996). We have no evidence as to whether the P1ex mutations increase the splicing rate by facilitating the folding of the intron. In addition, Woodson and Cech (1991) have demonstrated that an alternative secondary structure in the 5′ exon [referred to as P(−1)] can compete with the formation of P1 and affect the splicing rate of the *Tetrahymena* intron from the natural pre-rRNA. However, only two out of the five P1ex mutations that were observed to rescue pre-mRNA splicing are involved in a P(−1) interaction in the engineered KNT gene. Therefore, it is unlikely that the P1ex mutations promote the splicing rate primarily by modulating the equilibrium between the P1 and P(−1) helices.

The P1ex mutants tested for splicing in vitro have slightly different reaction rates (Table 2A,B). These could perhaps be due to differential effect on the stability of P1ex, or on differential disruption of P10 helical interactions. G3A is located in the 5′ terminal region of the intron and only destabilizes the P1ex helix. In contrast, C20A is in an overlapping region between P1 and P10, and it shortens P10 from the side close to the 3′ splice site by 1 bp. Yet another possibility is that some base substitutions cause more or less steric clash between the partially disrupted P1ex helix and the intron active site.

The initial velocities of the P1ex mutants relative to the original Tet-119 increase with decreased concentration of GTP, suggesting that their K_M^{GTP} values are decreased. The K_M^{GTP} value for a self-splicing intron typically does not represent the dissociation constant of GTP from the intron (Herschlag & Cech, 1990), although the two can be equal under special circumstances such as low pH or with deoxynucleotides in the 5′ exon. In the present case, the change in K_M^{GTP} value could simply result from an alteration of the rate-limiting step in the splicing pathway. However, it is still possible that the GTP binding affinity of the intron is modified by the P1ex mutations. Although P1ex and the G-binding site in P7 (Michel et al., 1989) are far away from each other in the secondary structure, they must be close together in the folded RNA: P1ex is immediately next to the 5′ splice site, and the 3′ hydroxyl of GTP is the attacking group for the 5′ cleavage reaction. Furthermore, it was shown that the formation of the P1 helix is

thermodynamically coupled with G binding (Bevilacqua et al., 1993; McConnell et al., 1993), so there could be coupling to the nearby P1ex as well.

P1ex mutations compromise accuracy

Although the P1ex mutants splice faster than the original Tet-119 in terms of both disappearance of pre-mRNA and accumulation of ligated exons and excised intron, a small fraction (about 20% and 10% for G3A and C20A, respectively, as estimated from the earliest time points) of the splicing product utilizes a cryptic 3′ splice site 100 bases downstream of the normal site (Fig. 4A,B). So it appears that P1ex mutations increase the reaction rate for splicing at the expense of accuracy. The cryptic 3′ splice site could potentially form P10 interactions of 9 bp with the IGS for Tet-119 and for the G3A mutant, and 7 bp for C20A. It was shown previously that the P10 helix of the *Tetrahymena* intron contributes to 3′ splice site selection (Suh & Waring, 1990), although it is not required for the splicing activity (Been & Cech, 1985). A possible explanation of the increased missplicing at the 3′ exon in the P1ex mutants is that the disruption of P1ex region makes the formation of P10 easier, and in turn increases the chance of a cryptic splice site being utilized. In the wild-type *Tetrahymena* rRNA intron, the P1ex region consists of three AU base pairs (Fig. 1B). We speculate that this is a compromise between the requirements for splicing rate and specificity in the eukaryotic cellular environment.

Conclusions

In vitro selection has been used previously to evolve RNA molecules with improved enzymatic properties (Beaudry & Joyce, 1992; Pan & Uhlenbeck, 1992; Bartel & Szostak, 1993; Illangasekare et al., 1995; Jabri & Cech, 1998). However, the environment in cells is much more complex and very hard to control. Thus, if the goal is to obtain a variant ribozyme with improved activity in cells, in vivo selection is a more direct approach. In the current study, a group I intron that evolved naturally in a eukaryotic organism, *T. thermophila*, was engineered into a kanamycin resistance gene and tested in a prokaryote, *E. coli*. There had been several reports on self-splicing activity of the *Tetrahymena* intron in *E. coli* in the context of different exon sequences (Price & Cech, 1985; Waring et al., 1985; Zhang et al., 1995; Roman & Woodson, 1998). It was therefore somewhat surprising that the designed construct did not work in vivo. Given the complex nature of the system, it was hard to predict the reason. The in vivo selection provided an excellent tool to solve this problem. In addition, the selection results and subsequent in vitro splicing studies strongly suggest a previously unappreciated feature of group I intron splicing, that too strong a helical structure in the P1ex region can inhibit the splicing reaction.

MATERIALS AND METHODS

Molecular cloning and construction of mutation libraries

Molecular cloning was performed according to standard protocols (Sambrook et al., 1989). The *E. coli–T. thermophilus* shuttle vector pUC19EKF-Tsp3 was kindly provided by Dr. Shuang-yong Xu at New England Biolabs (Wayne & Xu, 1997). The plasmid was reengineered so that the original P_{taq} promoter, which directs the transcription of the KNT gene in *T. thermophilus*, was replaced by the slpA promoter that is functional in both *E. coli* and *T. thermophilus* (de Grado et al., 1998). The KNT gene was flanked by unique restriction sites *Nde*I and *Kpn*I. The size of the plasmid was also reduced to 6 kb.

Mutagenic PCR (Fromant et al., 1995) was used to introduce random mutations to the intron region of both Tet-119 and Tet-119Δ237-330 constructs. The PCR mix contained 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 4.76 mM $MgCl_2$, 0.56 mM dATP, 0.90 mM dCTP, 0.20 mM dGTP, 1.40 mM dTTP, 0.5 mM $MnCl_2$, 0.5 μM each primer, 3 ng/μL Tet-119 plasmid template, 0.08 U/μL Taq DNA polymerase (Roche). The thermal cycles included 1× (94 °C, 5 min), 10× (94 °C, 1 min; 52 °C, 1 min; 72 °C, 3 min), and 1× (72 °C, 10 min). The 5′ and 3′ exon DNAs were amplified by high-fidelity PCR using Pfu Turbo DNA polymerase (Stratagene), the mix of which contained 1× cloned Pfu buffer (Stratagene), 0.2 mM each of dNTP, 0.4 μM primers, and 0.2 ng/μL template. The PCR cycles included 1× (94 °C, 1 min), 30× (94 °C, 0.5 min; 55 °C, 0.5 min; 72 °C, 1 min), 1× (72 °C, 10 min). The exons and the mutated intron PCR products were purified on a 1% agarose gel. The mutation libraries were constructed by a second Pfu Turbo PCR reaction, at the same condition as described, except that 0.4 ng/μL each of the 5′ and 3′ exons, and 15 ng/μL mutated intron were used as the template. The elongation time in the PCR cycles was increased to 2 min. The second PCR products were purified on a 1% agarose gel and digested with *Nde*I and *Kpn*I. The mutation libraries were ligated into the *Nde*I/*Kpn*I-digested modified pUC19EKF-Tsp3 vector. The libraries were then transformed into 50 μL of the XL-1 Blue MRF′ electrocompetent *E. coli* (Stratagene). Immediately after the electroporation, 930 μL Super Broth (37 °C) was added and the cells were shaken at 37 °C for 1 h. Portions (100 μL) of the cells were spread on LB plates containing 10, 20, 30, 40, and 50 μg/mL kanamycin. The total number of transformants was estimated by spreading serial dilutions of the cells on LB/ampicillin plates, which contained 100 μg/mL carbnicillin (a modified form of ampicillin). All the plates were incubated at 37 °C overnight.

For in vitro transcription purposes, Tet-119 and the P1ex mutants were subcloned into pUC19 by high-fidelity PCR with a forward primer containing an *Eco*RI site and a T7 promoter, and a reverse primer containing a *Bam*HI site (Fig. 2A). The entire KNT and intron sequences were then verified by DNA sequencing.

Transcription of precursor RNAs

The precursor RNAs were transcribed in vitro by T7 RNA polymerase. By using an decreased concentration of $MgCl_2$ and increased concentrations of ribonucleotide triphosphate (NTP), the NTPs chelate most of the Mg^{2+} ion so that splicing of the precursor during transcription is almost completely inhibited (Michel et al., 1992). Transcription reaction mixtures for preparing uniformly labeled RNA contained 40 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 4 mM each of NTP, 1 μg of template plasmid cleaved with the restriction enzyme *Bam*HI (New England Biolabs), 40 μCi of [α-^{32}P]CTP (New England Nuclear), and 1 μL of T7 RNA polymerase (Davanloo et al., 1984). The mixture was incubated at 37 °C for 2 h and subsequently cleaned using P-30 spin columns (Bio-Rad). The RNA transcripts were purified by electrophoresis on 4% polyacrylamide [29:1 acrylamide:bis (acrylamide)]/8 M urea gels. The gels were then covered with Saran Wrap and exposed to Kodak X-Omat films. The pre-mRNA bands were excised according to the films, crushed and soaked in a TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA) overnight. The RNA extractions were then filtered through a 0.22 μm membrane to remove the polyacrylamide, and ethanol precipitated using standard protocol. Purified RNA was resuspended in 50 μL sterile water, a 2 μL fraction of which was mixed with 4 mL ScintiSafe Econo1 scintillation fluid (Fisher) and counted using a Beckman LS-3801 scintillation counter.

The unlabeled pre-mRNAs were transcribed in 2-mL reactions in a similar condition to the one described for uniformly labeled pre-mRNAs, except that no ^{32}P-labeled nucleotide was added. The transcription reactions were incubated at 37 °C for 17 h. The RNA samples were purified on 3-mm thick denaturing gels, which were then visualized using ultraviolet (UV) shadowing. The gel-extracted RNAs were concentrated using UltraFree-15 concentrator (Millipore) with a molecular weight cutoff of 30 kDa to a volume of around 200 μL. Finally the concentrations of the pre-mRNAs were determined by UV absorbance.

To make sure that the comparison of Tet-119 and the P1ex mutants was fair in the splicing assays, the concentration and length of the purified unlabeled pre-mRNAs were checked on an analytical 4% polyacrylamide/8 M urea gel. The gel was soaked in a staining solution containing 10,000:1 dilution of SYBR Green II (Molecular Probes) in 1× TBE buffer for 10 min after the electrophoresis and then was scanned using a Typhoon 8600 imager (Molecular Dynamics) with an excitation wavelength of 532 nm and a short-pass emission filter 526SP. The Century-Plus RNA markers (Ambion) on this gel indicated that the pre-mRNAs were at the expected size. Quantitation of the pre-mRNA bands confirmed that their relative concentrations are correct within an error range of 10%.

In vitro splicing assays

Uniformly labeled pre-mRNAs were preincubated in 50 μL of 30 mM Tris-HCl, pH 7.5, 100 mM $(NH_4)_2SO_4$, and 10 mM $MgCl_2$ at 50 °C for 3 min. The splicing was initiated by adding GTP to the concentration specified and the incubation was continued at 50 °C. Fractions of 5 μL were removed from the reaction at indicated time points and were mixed immediately with 5 μL of 2× loading buffer, which contains 10 M urea, 30 mM EDTA, and 0.01% bromophenol blue, 0.025% xylene cyanol and 0.1× TBE electrophoresis buffer (1× TBE is 0.1 M Tris base, 0.083 M boric acid, and 1 mM EDTA), to stop the splicing. Samples were run on 4% polyacrylamide/8 M urea gels, which were then attached to filter papers, dried in vacuum and exposed to image plates. The images were scanned by a Typhoon 8600 imager (Molecular Dynamics) at 100 μm resolution. The splicing assays with unlabeled pre-mRNA and [α-^{32}P]GTP were performed similarly. The 50-μL reactions with 250 nM GTP were started by adding 10 μCi of [α-^{32}P]GTP, whereas those with 200 μM GTP were initiated by adding 50 μCi of [α-^{32}P]GTP and 5 μL 2 mM unlabeled GTP.

The RNA bands on the gel images were quantified using ImageQuant (version 5.2, Molecular Dynamics). The background was corrected by counting a box of the same size as the one around the RNA band located in the same lane. The kinetic data obtained were fit and plotted using the program PRISM (version 3.0, GraphPad). The data on disappearance of the pre-mRNA were fit to a one-phase exponential decay curve

$$Y = Y_0 * \exp(-k * t) + Plateau$$

where Y is the fraction remaining, t is time, k is the observed reaction rate k_{obs}, *Plateau* represents the unreacted fraction of pre-mRNA, and Y_0 is the amount of the active starting pre-mRNA. The data on accumulation of the ligated exons and G-IVS were fit to a one-phase exponential association curve

$$Y = Y_{max} * (1 - \exp(-k * t))$$

where Y is the fraction accumulated, t is time, Y_{max} is the maximum amount that could be generated, and k is the rate constant. The initial velocity v was calculated as $Y_{max} * k$ as the data were fit to the exponential curves.

Kinetic RT-PCR

Total RNAs were isolated from 2 mL *E. coli* cultures containing 100 μg/mL ampicillin using the RNeasy Mini kit (QIAGEN), and were digested by RQ1 RNase-free DNase I (Promega). The RNA concentrations were determined by UV absorbance. Equal amounts of RNAs were used in RT reactions containing SuperScript II reverse transcriptase (Invitrogen), which were followed by PCR reactions using Taq DNA polymerase (Roche). The primer used in the RT reactions and also as the reverse PCR primer has the sequence GAAG ATCTGATTGCTTAACTGCTTCAGTT, and the sequence of the forward PCR primer is GGTGGAATCAGATTGGCCGCTT. The PCR products contain 330 bp, which span positions 274 to 603 downstream of the translational start site of the KNT mRNA. Samples were taken from the PCR reactions at the end of cycles 12, 14, 16, 18, 20, 22, 24, and 26, and were analyzed by a 1.4% agarose gel. After electrophoresis, the gel was stained in 10,000:1 diluted Vistra Green (Amersham Pharmacia) and was scanned using a Typhoon 8600 imager with an excitation filter at 532 nm and a short-path emission filter 526SP. The bands were quantified using ImageQuant and the range of exponential growth was identified. The data

were fit to the one-phase exponential association curve using the program PRISM (version 3.0, GraphPad).

ACKNOWLEDGMENTS

We are grateful to Dr. Shuang-yong Xu, who generously provided the pUC19EKF-Tsp3 plasmid and *T. thermophilus* cells. We thank Prof. Norm Pace for advice on transformation of *T. thermophilus*. We thank Anne Gooding for providing T7 RNA polymerase and Elaine Podell and Karen Goodrich for oligonucleotide synthesis. We also thank Art Zaug and Peter Baumann for insightful discussions and Andy Berglund for critical reading of the manuscript.

Received January 17, 2002; returned for revision February 13, 2002; revised manuscript received March 4, 2002

REFERENCES

Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science 261*:1411–1418.

Bass BL, Cech TR. 1984. Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: Implications for biological catalysis by RNA. *Nature 308*:820–826.

Bass BL, Cech TR. 1986. Ribozyme inhibitors: Deoxyguanosine and dideoxyguanosine are competitive inhibitors of self-splicing of the *Tetrahymena* ribosomal ribonucleic acid precursor. *Biochemistry 25*:4473–4477.

Beaudry AA, Joyce GF. 1992. Directed evolution of an RNA enzyme. *Science 257*:635–641.

Been MD, Cech TR. 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.

Been MD, Cech TR. 1986. One binding site determines sequence specificity of *Tetrahymena* pre-rRNA self-splicing, *trans*-splicing, and RNA enzyme activity. *Cell 47*:207–216.

Bevilacqua PC, Johnson KA, Turner DH. 1993. Cooperative and anticooperative binding to a ribozyme. *Proc Natl Acad Sci USA 90*:8357–8361.

Cech TR. 1990. Self-splicing of group I introns. *Annu Rev Biochem 59*:543–568.

Cech TR, Zaug AJ, Grabowski PJ. 1981. In vitro splicing of the ribosomal RNA precursor of *Tetrahymena*: Involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell 27*:487–496.

Davanloo P, Rosenberg AH, Dunn JJ, Studier FW. 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci USA 81*:2035–2039.

Davies RW, Waring RB, Ray JA, Brown TA, Scazzocchio C. 1982. Making ends meet: A model for RNA splicing in fungal mitochondria. *Nature 300*:719–724.

de Grado M, Lasa I, Berenguer J. 1998. Characterization of a plasmid replicative origin from an extreme thermophile. *FEMS Microbiol Lett 165*:51–57.

Donis-Keller H, Maxam AM, Gilbert W. 1977. Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res 4*: 2527–2538.

Emerick VL, Pan J, Woodson SA. 1996. Analysis of rate-determining conformational changes during self-splicing of the *Tetrahymena* intron. *Biochemistry 35*:13469–13477.

Emerick VL, Woodson SA. 1994. Fingerprinting the folding of a group I precursor RNA. *Proc Natl Acad Sci USA 91*:9675–9679.

Fromant M, Blanquet S, Plateau P. 1995. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem 224*:347–353.

Hagen M, Cech TR. 1999. Self-splicing of the *Tetrahymena* intron from mRNA in mammalian cells. *EMBO J 18*:6491–6500.

Herschlag D, Cech TR. 1990. Catalysis of RNA cleavage by the *Tetrahymena* thermophila ribozyme. 2. Kinetic description of the reaction of an RNA substrate that forms a mismatch at the active site. *Biochemistry 29*:10172–10180.

Illangasekare M, Sanchez G, Nickles T, Yarus M. 1995. Aminoacyl-RNA synthesis catalyzed by an RNA. *Science 267*:643–647.

Jabri E, Cech TR. 1998. In vitro selection of the Naegleria GIR1 ribozyme identifies three base changes that dramatically improve activity. *RNA 4*:1481–1492.

Jones JT, Lee SW, Sullenger BA. 1996. Tagging ribozyme reaction sites to follow *trans*-splicing in mammalian cells. *Nat Med 2*: 643–648.

Lehman N, Joyce GF. 1993. Evolution in vitro of an RNA enzyme with altered metal dependence. *Nature 361*:182–185.

Liao H, McKenzie T, Hageman R. 1986. Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. *Proc Natl Acad Sci USA 83*:576–580.

Long MB, Sullenger BA. 1999. Evaluating group I intron catalytic efficiency in mammalian cells. *Mol Cell Biol 19*:6479–6487.

Matsumura M, Katakura Y, Imanaka T, Aiba S. 1984. Enzymatic and nucleotide sequence studies of a kanamycin-inactivating enzyme encoded by a plasmid from thermophilic bacilli in comparison with that encoded by plasmid pUB110. *J Bacteriol 160*:413–420.

McConnell TS, Cech TR, Herschlag D. 1993. Guanosine binding to the *Tetrahymena* ribozyme: Thermodynamic coupling with oligonucleotide binding. *Proc Natl Acad Sci USA 90*:8362–8366.

Michel F, Hanna M, Green R, Bartel DP, Szostak JW. 1989. The guanosine binding site of the *Tetrahymena* ribozyme. *Nature 342*:391–395.

Michel F, Jaeger L, Westhof E, Kuras R, Tihy F, Xu MQ, Shub DA. 1992. Activation of the catalytic core of a group I intron by a remote 3′ splice junction. *Genes & Dev 6*:1373–1385.

Michel F, Westhof E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol 216*:585–610.

Pan T, Uhlenbeck OC. 1992. In vitro selection of RNAs that undergo autolytic cleavage with Pb^{2+}. *Biochemistry 31*:3887–3895.

Price JV, Cech TR. 1985. Coupling of *Tetrahymena* ribosomal RNA splicing to beta-galactosidase expression in *Escherichia coli*. *Science 228*:719–722.

Pyle AM, McSwiggen JA, Cech TR. 1990. Direct measurement of oligonucleotide substrate binding to wild-type and mutant ribozymes from *Tetrahymena*. *Proc Natl Acad Sci USA 87*:8187–8191.

Ritchings BW, Lewin AS. 1992. Mutational evidence for competition between the P1 and the P10 helices of a mitochondrial group I intron. *Nucleic Acids Res 20*:2349–2353.

Roman J, Woodson SA. 1995. Reverse splicing of the *Tetrahymena* IVS: Evidence for multiple reaction sites in the 23S rRNA. *RNA 1*:478–490.

Roman J, Woodson SA. 1998. Integration of the Tetrahymena group I intron into bacterial rRNA by reverse splicing in vivo. *Proc Natl Acad Sci USA 95*:2134–2139.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Shaw LC, Thomas J Jr, Lewin AS. 1996. The Cbp2 protein suppresses splice site mutations in a group I intron. *Nucleic Acids Res 24*:3415–3423.

Strobel SA, Cech TR. 1995. Minor groove recognition of the conserved G.U pair at the *Tetrahymena* ribozyme reaction site. *Science 267*:675–679.

Suh ER, Waring RB. 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.

Sullenger BA, Cech TR. 1994. Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing. *Nature 371*:619–622.

Szewczak AA, Ortoleva-Donnelly L, Ryder SP, Moncoeur E, Strobel SA. 1998. A minor groove RNA triple helix within the catalytic core of a group I intron. *Nat Struct Biol 5*:1037–1042.

Treiber DK, Rook MS, Zarrinkar PP, Williamson JR. 1998. Kinetic intermediates trapped by native interactions in RNA folding. *Science 279*:1943–1946.

Waring RB, Ray JA, Edwards SW, Scazzocchio C, Davies RW. 1985. The *Tetrahymena* rRNA intron self-splices in *E. coli*: In vivo evidence for the importance of key base-paired regions of RNA for RNA enzyme function. *Cell 40*:371–380.

Waring RB, Towner P, Minter SJ, Davies RW. 1986. Splice-site se-

lection by a self-splicing RNA of *Tetrahymena*. *Nature 321*: 133–139.
Wayne J, Xu SY. 1997. Identification of a thermophilic plasmid origin and its cloning within a new *Thermus–E. coli* shuttle vector. *Gene 195*:321–328.
Wiesner RJ. 1992. Direct quantification of picomolar concentrations of mRNAs by mathematical analysis of a reverse transcription/exponential polymerase chain reaction assay. *Nucleic Acids Res 20*:5863–5864.
Woodson SA, Cech TR. 1991. Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry 30*: 2042–2050.
Woodson SA, Emerick VL. 1993. An alternative helix in the 26S rRNA promotes excision and integration of the *Tetrahymena* intervening sequence. *Mol Cell Biol 13*:1137–1145.
Young B, Herschlag D, Cech TR. 1991. Mutations in a nonconserved sequence of the *Tetrahymena* ribozyme increase activity and specificity. *Cell 67*:1007–1019.
Zarrinkar PP, Sullenger BA. 1999. Optimizing the substrate specificity of a group I intron ribozyme. *Biochemistry 38*:3426–3432.
Zhang F, Ramsay ES, Woodson SA. 1995. In vivo facilitation of *Tetrahymena* group I intron splicing in *Escherichia coli* pre-ribosomal RNA. *RNA 1*:284–292.