4602–4611 *Nucleic Acids Research, 2005, Vol. 33, No. 14*
doi:10.1093/nar/gki770

A peripheral element assembles the compact core structure essential for group I intron self-splicing

Mu Xiao, Tingting Li, Xiaoyan Yuan, Yuan Shang, Fu Wang, Shoudeng Chen and Yi Zhang*

Department of Biotechnology, State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China

Received May 17, 2005; Revised July 20, 2005; Accepted July 31, 2005

ABSTRACT

The presence of non-conserved peripheral elements in all naturally occurring group I introns underline their importance in ensuring the natural intron function. Recently, we reported that some peripheral elements are conserved in group I introns of IE subgroup. Using self-splicing activity as a readout, our initial screening revealed that one such conserved peripheral elements, P2.1, is mainly required to fold the catalytically active structure of the *Candida* ribozyme, an IE intron. Unexpectedly, the essential function of P2.1 resides in a sequence-conserved short stem of P2.1 but not in a long-range interaction associated with the loop of P2.1 that stabilizes the ribozyme structure. The P2.1 stem is indispensable in folding the compact ribozyme core, most probably by forming a triple helical interaction with two core helices, P3 and P6. Surprisingly, although the ribozyme lacking the P2.1 stem renders a loosely folded core and the loss of self-splicing activity requires two consecutive transesterifications, the mutant ribozyme efficiently catalyzes the first transesterification reaction. These results suggest that the intron self-splicing demands much more ordered structure than does one independent transesterification, highlighting that the universally present peripheral elements achieve their functional importance by enabling the highly ordered structure through diverse tertiary interactions.

INTRODUCTION

The fundamental roles of ribozymes in processing various RNA molecules have been appreciated in the past two decades (1–4). It has been realized only recently that RNA molecules are active players in synthesizing polypeptides and probably pre-mRNA splicing in all the living organisms and the eukaryotes, respectively (5,6). RNA catalyzing self-cleavage also controls gene expression by responding to natural metabolites in microorganisms (7) and by promoting RNA polymerase II termination in mammalian cells (8). Therefore, understanding how functional RNAs achieve the active structure becomes increasingly important. As the first discovered class of catalytic RNA, group I introns have fundamentally contributed to the current understanding of RNA structure, folding and function (9–13).

Most naturally occurring ribozymes catalyze a self-cleavage reaction through hydrolysis, while group I and group II introns catalyze the self-splicing reactions through two consecutive transesterifications (1–4). Consistent with that mechanism, the self-splicing reaction requires a more complicated ribozyme structure and the self-splicing ribozymes are much larger than the self-cleaving ribozymes (2,11). Although the crystal structures of a number of self-cleaving ribozymes and group I introns have been resolved (14–20), the fundamental structure required for catalysis of the two consecutive transesterification reactions has not yet been clearly identified. In addition to the self-splicing reaction that group I introns naturally catalyze, various other reactions have been developed to reflect the activity of group I introns, where most of them undergo only a one-step transesterification or hydrolysis (21–24). Nevertheless, it is unclear yet if the group I ribozyme folds to the same structure to catalyze those different reactions.

Group I introns are characterized by a conserved core structure that is formed by the assembly of the P4–P6 and the P3–P9 domains (16,25). Appendant to the core structure, all group I introns contain at least one variable paired region peripheral to the conserved core, termed the peripheral element that significantly enlarges the size of group I introns (25). The universal presence of peripheral elements underscores their importance in the natural function of group I introns. Interestingly, deletion of any of the peripheral elements of the *Tetrahymena* intron generally reduces but does not eliminate self-splicing *in vitro* (22,26). Our study of the *Candida* ribozyme, a group I intron belonging to IE subgroup (27), revealed a very different

*To whom correspondence should be addressed. Tel: +86 27 68756207; Fax: +86 27 68754945; Email: yizhang@whu.edu.cn

folding scenario compared with that of the *Tetrahymena* intron leading us to wonder how the unique set of peripheral elements of each intron modulates intron folding and activity *per se* (28,29). Our recent identification of the phylogenetically conserved peripheral elements in group I introns belonging to IE subgroup suggests the necessity of these structural elements in enabling the intron structure required for the natural splicing activity of these introns (30).

After the crystal structures of a number of group I ribozymes have been resolved and our understanding of the ribozyme folding and catalysis has been greatly enriched, exploring the function of the highly variable and abundant peripheral elements should be valuable to fully appreciate the structure responsible for the biological activity of those self-splicing introns.

Comparative sequence analysis of a large number of group I introns and the experimental study of the *Tetrahymena* intron reveal that many peripheral elements are capable of stabilizing the intron tertiary structure via long-range loop–loop or loop–receptor interactions, enabling the maximal ribozyme activity in the presence of low magnesium concentrations (11,22,26,31,32). Recent studies have showed that the interaction between the peripheral loops I and II of the hammerhead ribozyme is essential for the self-cleavage activity under low magnesium concentrations, as well as for the intracellular activity (33,34), suggesting that the peripheral interactions could be essential in forming the catalytic core in living cells.

The importance of the peripheral elements conserved among IE introns, including P2, P2.1, P9.1 and P9.2, in the folding of the native structure of these group I ribozymes was assessed in this study by using the *Candida* group I ribozyme as a model. Co-transcriptional self-splicing activity was used to reflect the importance of these peripheral elements in supporting the native ribozyme structure required for the intron self-splicing, a natural activity of a group I intron. It is not surprising that deletion of any of these conserved elements resulted in the reduced self-splicing activity, while the reduced extent greatly varied among these elements. However, it came as a big surprise that the deletion of P2.1 completely eliminated the intron self-splicing activity that could not be compensated even by high concentration of magnesium. Structural and functional analysis showed that a sequence-conserved short stem of P2.1 plays an unexpectedly important role in organizing the compact tertiary structure of the *Candida* ribozyme through helical interactions with both P3 and P6. The mutant ribozyme lacking the short P2.1 stem assumes the loosely folded structure(s) that is (are) only capable of catalyzing a one-step transesterification reaction. These findings highlight that a highly ordered ribozyme structure is required to catalyze intron self-splicing, and the peripheral elements could acquire an essential function through organizing such a compact structure.

MATERIALS AND METHODS

RNA preparation

Precursor RNA and *trans*-acting ribozyme for the wild-type intron were transcribed from the genomic DNA of *Candida albicans* (28) and those for the mutant introns were from the corresponding plasmid constructs. All the *trans*-acting ribozymes lack the first 11 nt at the 5′ end of the intron and cleave the substrate RNA Ca/sub-11 supplied *in trans* (28). PCR-directed mutagenesis was used to introduce deletion or point mutations to the introns (35). Flanking exons of 129 and 100 nt upstream and downstream of the intron, respectively, were included in each construct. DNA sequencing analysis was performed to confirm each mutation. RNA concentrations were determined by measuring the ultraviolet absorbance at 260 nm.

Precursor RNA for self-splicing was randomly labeled by the incorporation of [α-^{32}P]GTP (3000 Ci/mmol; PerkinElmer, Boston, MA) during *in vitro* transcription (36). Ribozyme RNA for the footprinting experiments was labeled at the 5′ end by the incorporation of [γ-^{32}P]ATP (3000 Ci/mmol; PerkinElmer) (28). RNA oligonucleotide Ca/sub-11 and the P2.1 RNA were synthesized by Dharmacon Research, Inc. (Lafayette) and labeled at the 5′ end as described previously (28).

Analysis of the activity of pure ribozymes

Self-splicing of the labeled precursor RNA was performed at 37°C for 20 min in 10 µl reaction under standard conditions, except as otherwise indicated (29). The substrate cleavage activity of the *trans*-acting ribozymes was measured as described previously (28).

Co-transcriptional splicing of the precursor RNA

PCR-amplified DNA containing the wild-type or mutant introns and adjacent exons was transcribed in the standard transcription reaction (36) except that 0.02 µCi of [α-^{32}P]-GTP (3000 Ci/mmol; PerkinElmer) was included in each reaction. Co-transcriptional splicing was run for 20 min and fractionated on a 5% PAGE–8 M urea gel and was then exposed against X-ray films.

Footprinting analysis of the ribozyme structure

The 5′ end labeled *trans*-acting ribozymes were cleaved by ribonuclease T1 or hydroxyl radical to probe the ribozyme folding. T1 footprinting was performed as described previously (28). Hydroxyl radical footprinting experiment was performed under similar conditions to T1 footprinting as reported previously (37). The 5′ end labeled RNA (~20 ng) in 4 µl of 10 mM sodium cacodylate (pH 7) and 0.1 mM EDTA was denatured at 95°C for 1 min, annealed at 37°C for 1 min and then chilled on ice. Each sample was then allowed to fold at 37°C for 5 min after the addition of 1 µl of 5× folding buffer containing 85 mM sodium cacodylate and the desired concentrations of $MgCl_2$. Cleavage was initiated by adding 6 µl of hydroxyl radical solution using a 12-channel pipette. The hydroxyl radical solution contains equal volumes of 15 mM $(NH_4)Fe(SO_4)_2$, 25 mM EDTA, 45 mM sodium ascorbate and 0.3% H_2O_2 that were mixed 30–45 s before the cleavage. The OH-dependent RNA cleavage was allowed to proceed for 40 s and then quenched by adding 10 µl of 40 mM thiourea. RNA samples were then precipitated with 0.1 vol of 3 M sodium acetate (pH 5.4), 10 µg of tRNA and 2.5 vol of pure ethanol. The RNA pellets were dissolved in RNA sequence loading buffer for further analysis.

Gel-shift analysis

Varying concentrations of the purified ribozymes were denatured at 95°C for 1 min in 10 mM Tris–HCl (pH 7.5) and were then incubated with 2 nM of the radiolabeled P2.1 RNA at 37°C for 30 min in the presence of 50 mM Tris–HCl (pH 7.5) and 10 mM $MgCl_2$. All samples were then chilled on ice and an equal volume of the loading dye containing 30% glycerol was added, which was followed by electrophoresis on 5% native polyacrylamide gels in 1× running buffer (50 mM Tris–HCl, 30 mM borate, 0.1 mM EDTA, pH 8.0 and 10 mM $MgCl_2$) at ~10 V/cm for ~1 h at 4°C.

In this study, gels were exposed to phosphor screens that were scanned and analyzed using Typhoon 9200 variable scanner (Amersham Pharmacia Biotech) to obtain the quantitative data. All data were plotted using the GraphPad Prism 4.0 program (www.graphpad.com).

RESULTS

Deletion of the peripheral element P2.1 eliminates the intron self-splicing activity that is not compensated by high concentrations of Mg^{2+}

We attempted to study how each of the four peripheral elements conserved in IE introns influences folding of the *Candida* group I intron using co-transcriptional intron self-splicing as a readout. Mutant introns with each of these elements being deleted were constructed. Two additional deletion mutants lacking a non-conserved peripheral element P5abc or a peripheral element P9 conserved among all group I introns were included in the assay as controls. Splicing activity of mutant introns during transcription of the corresponding precursors was monitored by the levels of spliced intron and ligated exon (Figure 1A and data not shown). The splicing capability significantly differs among the mutant introns, ranking as WT > ΔP5abc > ΔP9 ≈ ΔP9.2 > ΔP2 > ΔP9.1 > ΔP2.1 (see Figure 2A for a summary of the nomenclature of the structural elements of the intron and the sequences of the mutants). Consistent with that the conserved peripheral elements play more important roles than the non-conserved ones in supporting the native fold required for the self-splicing activity of an IE intron, deletion of each of the conserved peripheral elements, including P2, P2.1, P9.1 and P9.2, rendered the *Candida* ribozyme a significantly lower splicing activity than did the non-conserved P5abc. It is noteworthy that the intron self-splicing activity of the mutant intron (ΔP9) lacking the P9 element conserved among all group I intron only resulted in a moderate impaired splicing activity, which could be explained as the importance of this highly conserved element is not reflected at our assay condition (Figure 1A).

We were particularly interested in studying the peripheral elements critical for the folding of the *Candida* ribozyme to the active structure required for intron self-splicing. P2.1 and P9.1 were the best candidates because their deletion fully or near fully compromised the splicing activity in the co-transcriptional assay. Peripheral elements are known to facilitate ribozyme activity through establishing tertiary interactions that stabilize the ribozyme structure; such a function is usually replaceable by high concentrations of magnesium (26,31–34). We studied if magnesium compensates for the loss of P2.1 and P9.1 structures. The purified RNA precursors containing the wild-type intron (Pre-WT) or the mutant introns (Pre-Δ2.1 and Pre-Δ9.1) were self-spliced at the same conditions. Self-splicing of the wild-type intron reached the half-maximal splicing at 1 mM Mg^{2+}, and the optimal splicing was usually observed at 1.5–2.5 mM Mg^{2+} (Figure 1B and data not shown). Although Pre-Δ9.1 did not self-splice at low Mg^{2+} concentrations, it spliced well at Mg^{2+} concentrations ⩾8 mM, indicating that P9.1 primarily functions in stabilizing the ribozyme structure (Figure 1B). However, increase in Mg^{2+} up to 50 mM neither produced spliced intron nor ligated exons

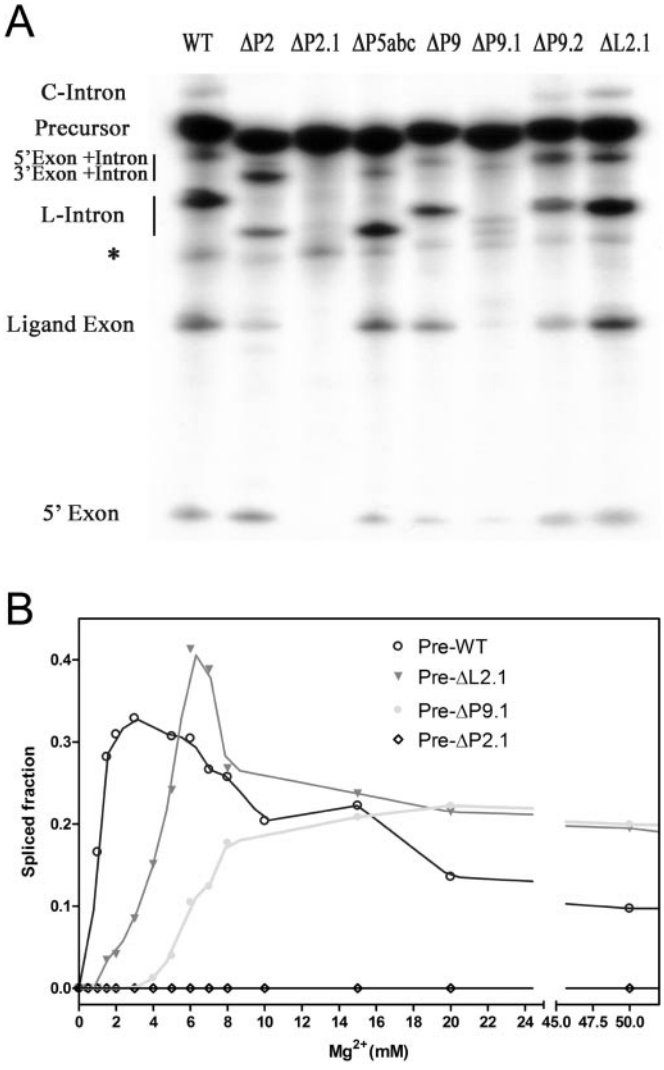

Figure 1. Self-splicing of the precursor RNA. (**A**) Co-transcriptional self-splicing of the wild-type and mutant introns lacking P2 (ΔP2, deletion of A21-U71), P2.1 (ΔP2.1, deletion of G73-C110), P5abc (ΔP5abc, A139-A188 replaced by GAAA), P9 (ΔP9, deletion of C299-G317), P9.1 (ΔP9.1, deletion of G319-C344), P9.2 (ΔP9.2, deletion of A346-U369) and L2.1 (ΔL2.1, G86-G99 replaced by GAAA). (**B**) Self-splicing of the gel-purified precursor RNAs. Pre-WT, Pre-ΔP2.1, Pre-ΔP9.1 and Pre-ΔL2.1 indicate precursors containing wild-type intron or corresponding mutant introns. These precursor RNAs were transcribed and gel-purified *in vitro*, and then subjected to a self-splicing reaction in the presence of the indicated concentrations of magnesium. The reaction was quenched by the EDTA-containing loading buffer and was analyzed on 5% PAGE-8M Urea gels. The band intensity on the PAGE gels was analyzed and plotted as described previously (29).

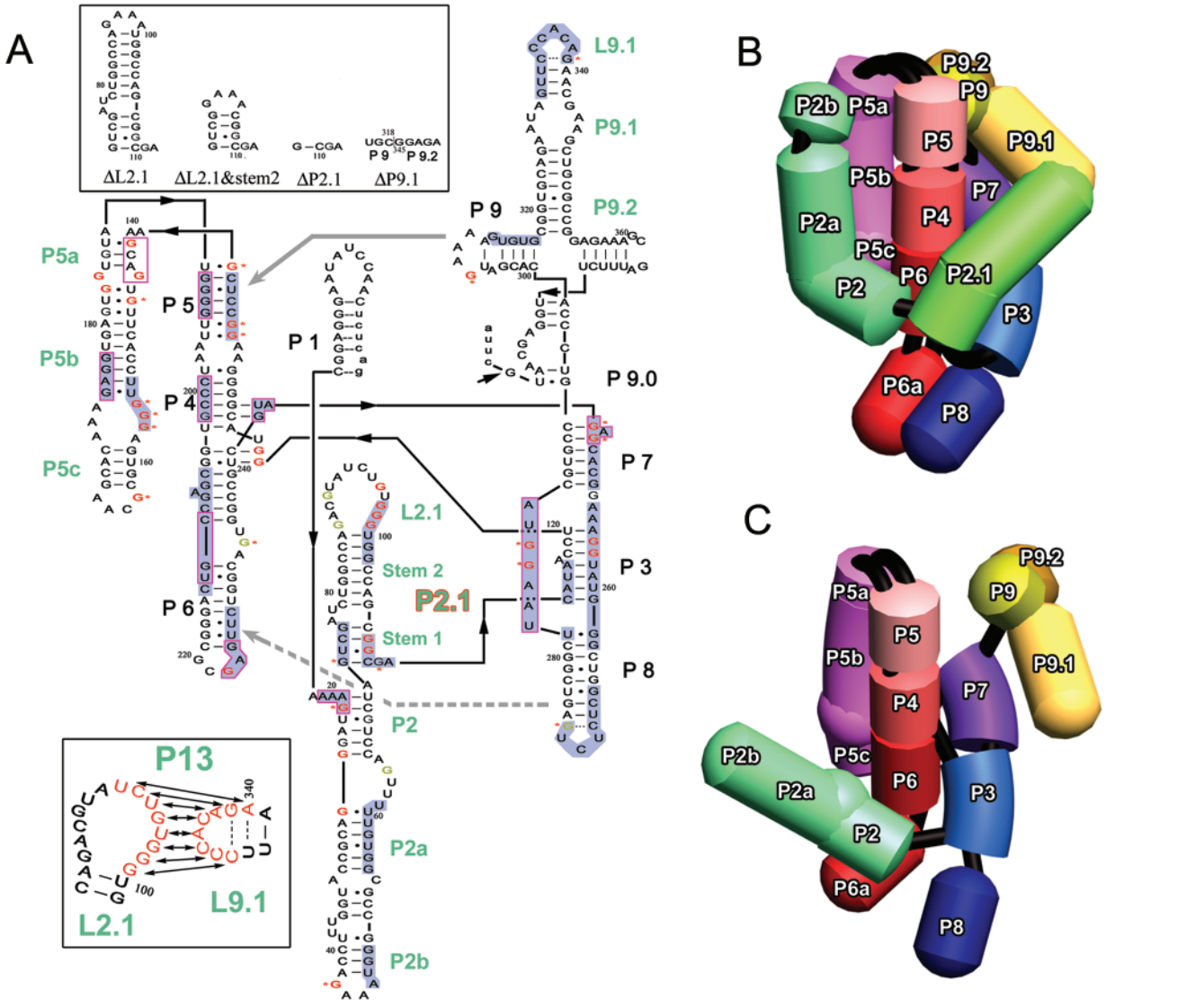

Figure 2. Structures of the *Candida* ribozyme. (**A**) The secondary structure with the footprinting data being labeled. Thirteen paired regions were denoted P1 to P9.0. The P13 stem formed by long-range pairing of L2.1 and L9.1 is shown in the lower left inset. The conserved elements are numbered in black and the peripheral elements are numbered in green. The local structures of the major deletion mutants in P2.1 and P9.1 are inset in the upper left. In the folded structure of the wild-type ribozyme, the backbone region protected from Fe(II)–EDTA hydroxyl radical cleavage are shaded in blue; guanosine residues protected from RNase T1 cleavage are in red font. In the presence of 50–100 mM magnesium, the backbone region of the mutant ribozyme $E^{\Delta P2.1}$ demonstrating protection from hydroxyl radical cleavage footprinting is boxed by purple lines, and nucleotides that become more accessible to T1 than the wild-type ribozyme are labeled by red asterisks. Representative gels of the footprinting results are shown in Supplementary Figures 2 and 3, which are published as supporting information available at *Nucleic Acids Research* Online. The schematic tertiary structures of the wild-type ribozyme (**B**) and mutant ribozyme $E^{\Delta P2.1}$ (**C**) are shown. P1 is not included in these 3D structures.

when P2.1 is deleted from the intron (Figure 1B). Apparently, P2.1 is absolutely required for self-splicing of the *Candida* intron through a function not replaceable by magnesium. These results could readily be interpreted as that, even in the presence of high concentration of magnesium, the P2.1 deletion leads to RNAs without native fold required for intron self-splicing.

High concentrations of Mg^{2+} compensate for the P2.1 function in enabling a one-step transesterification reaction

Self-splicing of group I introns consists of two consecutive transesterification reactions. We determined whether P2.1 is essential for a *trans*-cleavage reaction requiring only one transesterification. Figure 3 shows that no *trans*-cleavage activity is observed for the ribozyme lacking P2.1 ($E^{\Delta P2.1}$) at physiological concentrations of Mg^{2+}. However, unlike the self-splicing reaction, high concentrations of Mg^{2+} efficiently compensated for the loss of P2.1 in the *trans*-cleavage reaction. $E^{\Delta P2.1}$ requires >16-fold more Mg^{2+} than that of E^{WT} to reach the half-maximal cleavage. Gel retardation experiments showed that the apparent K_d for the substrate binding by $E^{\Delta P2.1}$ and E^{WT} were 5.495 and 3.465 nM, respectively (data not shown), excluding the possibility that the recovered *trans*-acting activity of $E^{\Delta P2.1}$ is due to an increased affinity of substrate binding. Therefore, it is more probable that, at high concentrations of magnesium, $E^{\Delta P2.1}$ folds to an active

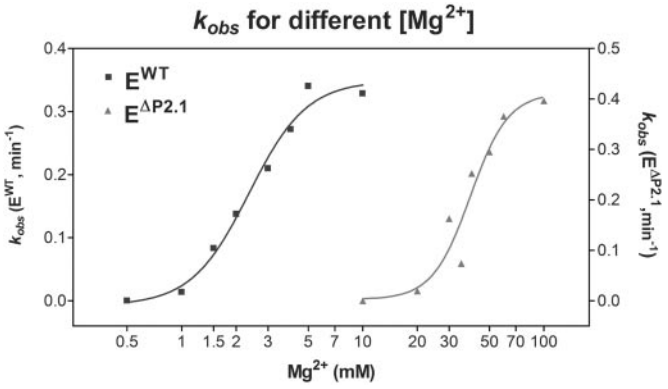

Figure 3. Magnesium-dependent *trans*-cleavage activity of the *Candida* ribozymes. *Trans*-cleavage activity of the wild-type (E^{WT}) and mutant ribozyme lacking P2.1 ($E^{\Delta P2.1}$) was studied as described previously (28). The fraction of the cleaved products (f_p) versus time was fitted to $f_p = f_{max}(1-e^{k_{obs}\cdot t})$ to obtain k_{obs} for each magnesium concentration. Then k_{obs} versus $[Mg^{2+}]$ was fitted to the equation $k_{obs} = k_{max}\times[Mg^{2+}]^n/(Mg^n_{1/2}[Mg^{2+}]^n)$ to represent the magnesium-dependent activity of the ribozymes. Calculations yield $n \approx 2.73$, $Mg_{1/2} \approx 2.36$ mM for E^{WT}, and $n \approx 4.16$, $Mg_{1/2} \approx 39.28$ mM for $E^{\Delta P2.1}$.

structure that supports the one-step but not two consecutive transesterification reactions.

The peripheral elements P2–2.1 and P5abc of the *Tetrahymena* ribozyme could act as *trans*-activators in restoring the ribozyme activity of mutants lacking these elements at low magnesium concentration (38,39). However, the presence of a large excess of P2.1 RNA (15 μM) *in trans* only slightly reduced the Mg^{2+} requirement for the *trans*-acting activity of $E^{\Delta P2.1}$ (75 nM) but did not restore the self-splicing activity of the *Candida* ribozyme even in the presence of up to 100 mM Mg^{2+} (data not shown). These results suggest that the P2.1 element of the *Candida* ribozyme primarily acts *in cis* to promote the native ribozyme fold required for the intron self-splicing, which significantly differs from that of the *Tetrahymena* ribozyme.

The long-range base paired P13 stabilizes the ribozyme structure and enables the ribozyme activity under physiological Mg^{2+} concentrations

We, then, started to dissect the roles of each structural motif of P2.1 in the folding of the catalytically active ribozyme. The terminal loop of P2.1 was first examined because both sequence analysis and T1 digestion predict the P13 helix formed by base pairing of this loop and the loop of P9.1 (Figure 2A). This long-range interaction was first identified in some group I introns belonging to IC1 subgroup (31) and has been demonstrated that it is highly conserved in IE introns (30). Gel retardation experiments were carried out to examine if the P2.1 RNA specifically binds to the P9.1 structure of the intron. The mutant ribozymes, $E^{\Delta P2.1}$ and $E^{\Delta P2.1\Delta P9.1}$, were incubated with the labeled P2.1 RNA for 20 min, allowing the binding to reach equilibrium. The P2.1 RNA bound to the mutant ribozyme $E^{\Delta P2.1}$ efficiently in a single-site binding mode with a K_d of 374.3 nM, consistent with a model that the binding results from stable base pairing between the free P2.1 RNA and the ribozyme P9.1 stem. The mutant lacking P9.1 ($E^{\Delta P2.1\Delta P9.1}$) lost the ability to bind the P2.1 RNA (Supplementary Figure 1, available at *Nucleic Acids Research* Online).

To study the function of the P13 structure, a mutant ribozyme $E^{\Delta L2.1}$ in which the P13 base pairing is disrupted by replacing the L2.1 sequence by a GAAA tetraloop was constructed (Figure 2A). In the presence of increasing concentrations of magnesium, T1 ribonuclease and Fe(II)–EDTA hydroxyl radical were used to monitor the tertiary interactions including long-range base pairing and helical–helical contacts (28,29,37). As shown in Figure 4, Supplementary Figure 2 and Supplementary Table 1, available at *Nucleic Acids Research* Online, the guanosine residues involved in the formation of the tertiary structure of the wild-type ribozyme were protected from T1 cleavage. For example, significant protections were observed at G95, 97–99 in L2.1 forming the P13 helix, at G121–122 in J3/4 forming a base triple and at G256–257 in P3′ forming the P3 helix. When the P13 base pairing was abolished, higher concentrations of Mg^{2+} were required to reach the compact ribozyme structure that confers resistance to T1 cleavage, suggesting that the P13 structure stabilizes the ribozyme tertiary structure.

Hydroxyl radical (OH) attacks the C4′ position of the sugar resulting in ribose decomposition and phosphodiester cleavage; thus it is a useful probe to study RNA tertiary structure, especially the helical–helical contacts in a folded RNA (37). We showed that Mg^{2+} cooperatively protected a number of helical regions of the *Candida* ribozyme from hydroxyl radical cleavage, reflecting global tertiary folding of the ribozyme (Figure 4, Supplementary Figure 3 and Supplementary Table 2, available at *Nucleic Acids Research* Online). The tertiary base pairing detected by T1 cleavage and the tertiary structure detected by hydroxyl radical cleavage were consistent with each other (Figure 2), validating the reliability of both methods in studying the ribozyme tertiary structure. Consistent with a model that the P13 base pairing stabilizes the ribozyme tertiary structure, the mutant ribozyme $E^{\Delta L2.1}$ requires much higher concentrations of magnesium to protect the helices than does the wild-type ribozyme (Figure 4, Supplementary Figure 3 and Supplementary Table 2, available at *Nucleic Acids Research* Online).

We then addressed how P13 affects the ribozyme activity. The mutant ribozyme $E^{\Delta L2.1}$ catalyzing the *trans*-cleavage reaction and its precursor Pre-ΔL2.1 catalyzing the self-splicing reaction were not active at low magnesium concentrations (Figure 4). Similar higher Mg^{2+} concentrations were required to recover both the activity of the mutant ribozyme lacking a P13 structure as that required for stabilizing its tertiary structure. On the other hand, the mutant ribozyme activity reached the levels of the wild-type at higher Mg^{2+} concentrations. We, therefore, conclude that the P13 tertiary base pairing enables the ribozyme activity under the physiological magnesium concentrations mainly through the stabilization of the ribozyme tertiary structure.

The essential function of P2.1 is localized to the sequence-conserved short stem

Owing to the essential role that P2.1 plays in the folding of the *Candida* intron into the structure required for intron self-splicing is not associated with the terminal loop that forms the P13 tertiary base pairing, we predicted that the P2.1 stem

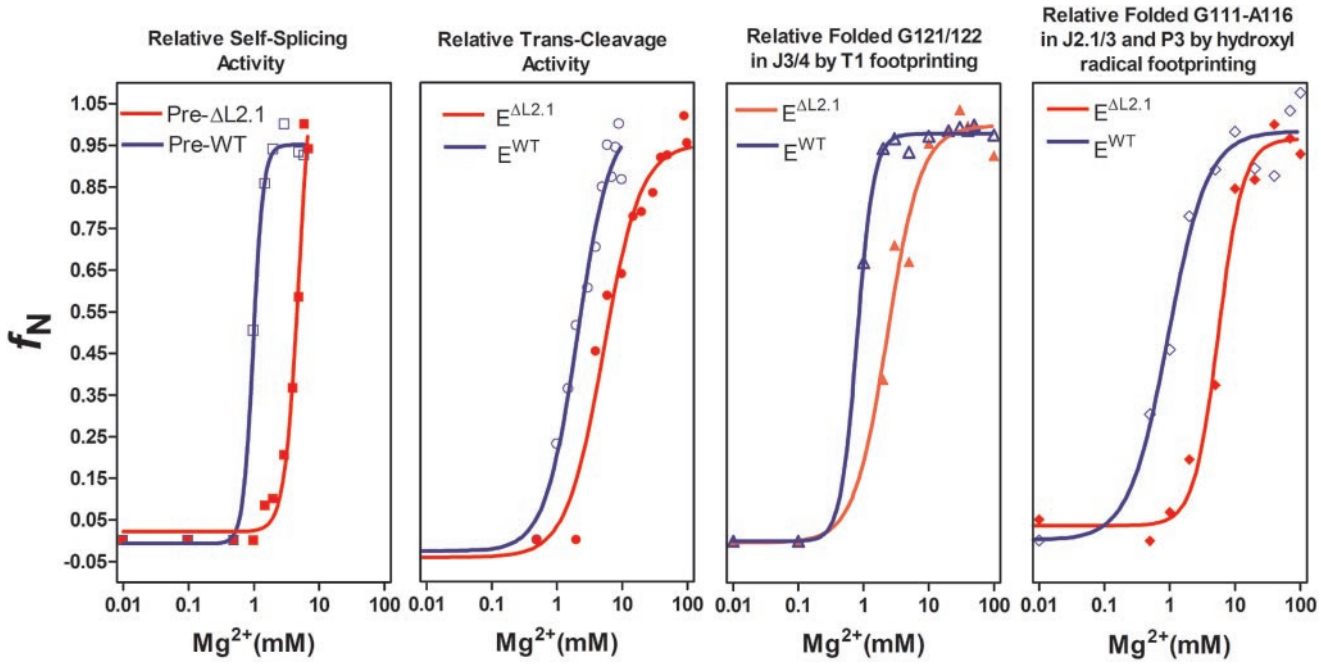

Figure 4. Mg^{2+}-dependent folding and catalytic activity of the wild-type and mutant ribozyme lacking L2.1 ($E^{\Delta L2.1}$). Self-splicing data was from Figure 1B with only the data at $Mg^{2+} \leqslant 7$ and 6 mM for the wild-type and mutant ribozyme $E^{\Delta L2.1}$, respectively, being plotted. Higher Mg^{2+}_2 concentrations inhibited self-splicing activity and were not plotted. Cleavage of Ca/sub-11 (50 nM) by the *trans*-acting ribozymes (10 nM) was run under the standard condition for 3 min. For activity assays, the fraction of native RNA (f_N) was determined as the fraction of the products of self-splicing or *trans*-cleavage reactions; the ribozyme structure under the Mg^{2+} concentration yielding maximal activity was assumed to be completely folded. Mg^{2+}-dependent folding of ribozymes was assayed by both RNase T1 footprinting and hydroxyl radical footprinting that were processed under the similar conditions. Representative gels are shown in Supplementary Figures 2 and 3, which are published as supporting information available at *Nucleic Acids Research* Online. The fraction folded (f_N) was normalized by taking the observed intensities at 0 mM $MgCl_2$ in $E^{\Delta P2.1}$ as unfolded and at 50 mM $MgCl_2$ in E^{WT} as completely folded. The band intensity change at each readable region/site was quantified and analyzed; the results are shown in Supplementary Tables 1 and 2, which are published as supporting information available at *Nucleic Acids Research* Online. T1 footprinting result at G121-122 of J3/4 and hydroxyl radical footprinting result at G111-A116 were plotted and shown here. The fraction of native RNA or folded (f_N) versus $[Mg^{2+}]$ was fitted to the Hill equation $f_N = f_{max}[Mg^{2+}]^n/(Mg^n_{1/2} + [Mg^{2+}]^n)$. Different symbols are used to indicate self-splicing activity (square), *trans*-cleavage activity (circle), T1 cleavage (triangle) and hydroxyl radical cleavage (diamond). The Hill coefficient (n) and half-maximal Mg^{2+} concentration ($Mg^{2+}_{1/2}$) for self-splicing activity were $n \approx 5.562$ (E^{WT}) and 3.851 ($E^{\Delta L2.1}$), $Mg^{2+}_{1/2} \approx 0.9802(E^{WT})$ and 4.987 ($E^{\Delta L2.1}$), respectively. For *trans*-cleavage activity, $n \approx 1.723$ (E^{WT}) and 1.528 ($E^{\Delta L2.1}$), $Mg^{2+}_{1/2} \approx 2.069(E^{WT})$ and 5.171 ($E^{\Delta L2.1}$).

Table 1. Co-transcriptional self-splicing of different intron mutants

	Self-splicing
P2.1 mutants (based on ΔL2.1&stem2)	
ΔL2.1&stem2	Yes
ΔG76-C107	Yes
ΔC75-G108	Yes
ΔU74-G109	No
J2.1/3 mutant	
G111→U	Yes
A112→U	Yes
ΔG111	Yes
ΔA112	Yes
G111→U, A112→U	Yes
ΔG111 & A112	No

exerts a critical function. We deleted the second stem of P2.1 and constructed a mutant ribozyme $E^{\Delta L2.1\&stem2}$ that contains the 4 bp stem1 and a GAAA tetraloop (Figure 2A). $E^{\Delta L2.1\&stem2}$ demonstrated a normal self-splicing activity during transcription (Table 1), suggesting that this 4 bp stem is sufficient to promote the native intron fold essential for the self-splicing activity. The second to the fourth pairs of this 4 bp stem were individually deleted, and the resulted ribozymes were subjected to the co-transcriptional self-splicing analysis. Interestingly, deletion of the second U•G pair abolished the ribozyme self-splicing activity during transcription, while deletion of neither of the other two individual deletions significantly impaired the ribozyme self-splicing activity (Table 1), reducing the essential P2.1 stem promoting the ribozyme native fold to 3 bp. The second U•G pair might exert the critical role through tertiary contacts with some structures of the ribozyme. The importance of the J2.1/3 in the folding of the *Candida* ribozyme to the native structure required for intron self-splicing was also studied, demonstrating that the length of the junction but not the base identity is essential for the native ribozyme fold. This was evident from the persistence of the self-splicing activity when J2.1/3 residues G111 and A112 were singly deleted or singly or both mutated, but loss of activity when both were deleted (Table 1).

The *Candida* group I intron belongs to the IE subgroup (27), which constitutes nearly 20% of the group I introns annotated to a known subgroup (www.rna.icmb.utexas.edu). Our comparative analysis of 211 group I introns in IE subgroup produced high-quality alignments (30). Analysis of the core sequence variation divides IE introns into three distinct minor subgroups, IE1 through IE3. The alignments showed that all the IE1 and IE2 introns display essentially the same

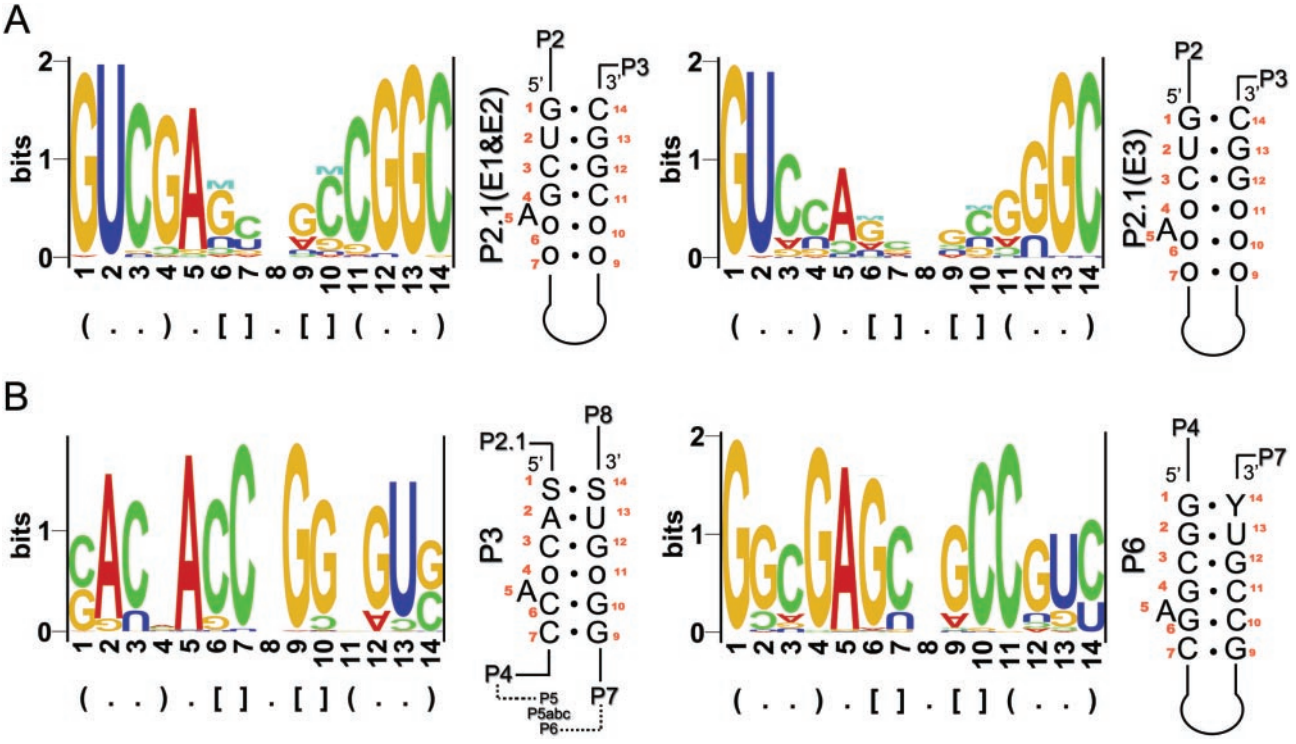

Figure 5. The conserved sequences and structure motif in P3, P6 and P2.1 paired regions. The base identity in the each position of the '4 bp stem-single A bulge-stem' was calculated and plotted by the program RNA Structure Logo (www.cbs.dtu.dk/~gorodkin/appl/slogo.html). Please note that the stem after the single A bulge was usually 4–6 bp, and only 2 bp were shown for simplicity.

base identity in the first 4 bp (stem1) of P2.1, while only the first 2 bp of P2.1 in IE3 introns are highly conserved (Figure 5A). The sequence conservation in the short P2.1 stem, especially the first 2 bp, suggests its critical role in the folding of the active structure of IE introns and is consistent with our experimental results of the study of the *Candida* intron.

The short P2.1 stem is essential in assembling the compact ribozyme core

T1 ribonuclease and hydroxyl radical footprinting assays were then carried out to elucidate the structural function of the P2.1 stem1. Formation of the compact core of a group I intron requires the assembly of the two independently folded domains P4–P6 and P3–P9 through extensive tertiary interactions. The known interactions include base triples J3/4:P6 and J6/7:P4, as well as the L9/P5 tetraloop–receptor interaction (25,31). Consistent with the model that the P4–P6 and P3–P9 domains are assembled into a compact core in the wild-type ribozyme, T1 and hydroxyl radical results showed that all those motifs involved in the core assembly were strongly protected in the folded structure at the physiological concentrations of magnesium (Figure 2).

Although deletion of L2.1 or both L2.1 and the stem2 of P2.1 produced the requirement for higher concentrations of magnesium to correctly fold the ribozyme, the final folded ribozyme structure is similar to the wild-type ribozyme (Figure 4, Supplementary Figures 2 and 3 and Supplementary Tables 1 and 2, available at *Nucleic Acids Research* Online). Nonetheless, the J2.1-P3 region of the mutant ribozymes is much more protected from the hydroxyl radical cleavage than is the corresponding region of the wild-type ribozyme, suggesting that the folded structures of the mutant ribozymes are not identical to the wild-type. Moreover, as reflected by the slope of the sigmoid curve drawn by the Hill equation, the folding cooperativity (n) of the wild-type and mutant ribozymes in response to magnesium concentrations showed obvious difference in a number of regions (Supplementary Table 2, available at *Nucleic Acids Research* Online). For example, folding of J2.1/3-P3 (G111-A116) and a part of J1/2 and P2 (A19–G22) were much more cooperative in $E^{\Delta L2.1}$ than in the wild-type ribozyme as indicated by hydroxyl radical cleavage protection, but folding of most other regions became less cooperative, suggesting that the folding pathways of these mutant ribozymes might be altered as well.

Surprisingly, the deletion of P2.1 resulted in a very loose structure even in the presence of up to 100 mM Mg^{2+} (Figure 2, Supplementary Figures 2 and 3, available at *Nucleic Acids Research* Online). In this loosely packed structure, the L9/P5 tertiary interaction is disrupted as demonstrated by the loss of the T1 protection in G308 of L9 and that of the hydroxyl radical protection at the P5 receptor region. Another striking feature of the loose $E^{\Delta P2.1}$ structure is that the P3–P7 pseudoknot lost almost all tertiary interactions with surrounding helices, as indicated by the lack of hydroxyl radical protection at this region. In addition, most hydroxyl radical protected regions in P8 stem, P2 and P6 disappeared. Interestingly, one of the P5a strand became more protected from hydroxyl radical attack when P2.1 was not present, suggesting that the

folding of this region in the mutant ribozyme was altered. Strikingly, such a loosely folded core is sufficient for catalyzing the *trans*-cleavage reaction consisting a one-step transesterification reaction (Figure 3).

Combining these observations with the catalysis data (Table 1) and the unpublished data provided by M. Xiao and Y. Zhang, obtained by footprinting analysis of the structures of other deletion mutants, we were able to outline the compact tertiary fold of the wild-type ribozyme (Figure 2B) and the loosely folded mutant ribozyme $E^{\Delta P2.1}$ (Figure 2C). In this model, two sides of the P2.1 stem1 helix makes intimate contacts with one side of P3 stem and P6 stem, respectively, tethering the P4–P6 domain with the P3–P9 domain. This assembly allows P9, P2 and P8 to make intimate contacts with P5, P5abc and P6, respectively. When the P2.1 stem1 is deleted, the P4–P6 and the P3–P9 domains cannot assemble, resulting in the disappearance of the tertiary interactions between P2 and P5abc, P8 and P6, and as well as P5 and L9. The loss of P2–P5abc and P9–P5 interactions allows the formation of a new local tertiary interaction between P5a and P5. As the three-partner interaction of P3–P2.1 stem1–P6 is helical–helical interaction, we named such an interaction as triple helical interaction. Apparently, this triple helical interaction bridged by P2.1 is essential in assembling the compact core structure of the *Candida* ribozyme. We interpret elimination of intron self-splicing by a 2-base deletion in J2.1/3 (Table 1), as that the P2.1 is unable to correctly position the two independent domains with a shorter J2.1/3.

Interestingly, analysis of the alignments and structures of all 211 IE introns revealed the presence of a unique structure motif '4 bp stem-single A bulge-stem' (30). Strikingly, this structure motif is highly and exclusively conserved in P2.1, P3 and P6 (Figure 5). As the bulged A is capable of establishing the A-minor interaction to stabilize the RNA tertiary structure, the universal and exclusive presence of the '4 bp stem-single A bulge-stem' motif in P3, P6 and P2.1 may suggest a network of A-minor interactions among these helices. This observation supports our finding that the P3–P2.1–P6 triple helical interaction may be established to assemble the compact core of the *Candida* ribozyme. However, the presence of the highly conserved '4 bp stem-single A bulge-stem' in all three helices involving in the P3–P2.1–P6 triple helical interaction suggests not only the important role of the A-minor interaction in establishing the triple helical interaction but also the general presence of this novel tertiary interaction among IE introns.

DISCUSSION

Novel structural role of the peripheral elements: lessons from a peripheral element essential for the natural ribozyme activity

In contrast to the conserved elements, the peripheral elements of group I introns vary considerably in size, structure and location (25). However, the universal presence of at least one peripheral element in all studied group I introns underscores the importance of this non-conserved structure (11).

The well-known function of peripheral elements is to stabilize the tertiary structure of group I ribozyme through peripheral interactions (31,32,40,41). The P2 stem functions as a spacer in efficient docking of P1 containing the 5′ splice site to the catalytic active site (42). Nevertheless, no peripheral elements have been reported so far to be absolutely required to produce the native fold of a group I ribozyme. Here, we reported that the peripheral element P2.1 is essential in folding of the compact tertiary structure of the *Candida* ribozyme that is required for intron self-splicing. This essential function, however, is not performed by the terminal loop of P2.1 that participates in the formation of the P13 tertiary interaction. Unexpectedly, the very short P2.1 stem linking P2 and P3 performs the essential role, through fostering the triple helical interaction with both P3 and P6. Although the detailed structure is not yet known, such an interaction is absolutely required for a compact assembly of the P4–P6 and P3–P9 domains. High concentrations of magnesium cannot compensate for the function of the short P2.1 stem1 in assembling the ribozyme core. Compared with the L9/P5 tetraloop–receptor interaction linking P4–P6 and P3–P9 domains, the P3–P2.1–P6 triple helical interaction is obviously more important in that it is essential for promoting the native ribozyme structure required for the natural intron activity.

Although some naturally occurring group I introns lack the P2.1 element, this peripheral element is present in many subgroups of group I introns (25,31) and essentially in all group I introns belonging to the IE subgroups (30). Furthermore, the sequence of the short P2.1 stem is highly conserved in IE subgroup, supporting the possible universal role of the triple helical interaction in folding of IE introns.

A picture has emerged from the present study and many previous studies to depict the function of peripheral elements in assisting the ribozyme folding, in which the diverse structural motifs of peripheral elements encode a large variety of tertiary interactions unique to different group I introns (18–20,30,31,43,44). Difference in peripheral interactions also characterizes different folding pathways (28,29,45,46).

Group I self-splicing requires a compact folded core: interpretation of the universal presence of the peripheral elements

It has been known for decades that group I introns catalyze two consecutive transesterification reactions to accomplish intron excision and exon ligation (21). Although the catalytic mechanism is somewhat different, the splicing of group II introns and nuclear introns also occurs via two sequential transesterification reactions (4). However, the specific structural components enabling the precise two-step transesterifications are not yet clear. *Trans*-acting ribozymes mimicking a one-step transesterification are widely used to understand the structure and function of group I ribozymes (21), but it has not yet been possible to dissect the ribozyme structure required to catalyze the one-step reaction and the self-splicing reaction. In this study, we demonstrated that the loose tertiary structure lacking a well-assembled core is capable of efficiently catalyzing the one-step transesterification reaction but is not able to catalyze the self-splicing reaction. These results reveal, for the first time, that the compactly assembled core structure of group I intron is not required for a one-step transesterification. Consistent with this finding, a *td* intron-derived ribozyme lacking the P4–P6 domain was reported to catalyze two independent transesterification reactions at both 5′ and 3′ splice sites, but is not capable of self-splicing (47). Therefore, we propose that

group I introns acquire the compact core structure to meet the challenge of the precise and efficient self-splicing requiring two consecutive and intimately linked transesterifications.

As discussed earlier, most of the studied peripheral elements display a function in stabilizing or organizing the compact tertiary structure of group I introns. Although this function in stabilizing the ribozyme core is replaceable by magnesium in many cases, such a function might be essential for the ribozyme activity in living cells where the magnesium concentrations are usually lower than the levels needed to replace the requirement for these peripheral elements (33). The finding that the compact folded structure is required for intron self-splicing offers an attractive explanation of the universal presence of peripheral elements, i.e. the peripheral elements establishing various tertiary interactions to stabilize and/or assemble the ordered ribozyme core to enable the intron splicing in the cellular environment.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank M. J. Leibowitz (UMDNJ) for careful reading of the manuscript and E. Westhof (IBMC-CNRS, France) for critical comments on this work. This work is supported by the NSFC funds (30170213 and 30330170) and the Wuhan University Fund (0000028) granted to Y.Z. The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Symons,R.H. (1997) Plant pathogenic RNAs and RNA catalysis. *Nucleic Acids Res.*, **25**, 2683–2689.
2. Frank,D.N. and Pace,N.R. (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.*, **67**, 153–180.
3. Doudna,J.A. and Cech,T.R. (2002) The chemical repertoire of natural ribozymes. *Nature*, **418**, 222–228.
4. Lambowitz,A.M. and Zimmerly,S. (2004) Mobile group II introns. *Annu. Rev. Genet.*, **38**, 1–35.
5. Valadkhan,S. and Manley,J.L. (2001) Splicing-related catalysis by protein-free snRNAs. *Nature*, **413**, 701–707.
6. Moore,P.B. and Steitz,T.A. (2003) The structural basis of large ribosomal subunit function. *Annu. Rev. Biochem.*, **72**, 813–850.
7. Winkler,W.C., Nahvi,A., Roth,A., Collins,J.A. and Breaker,R.R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, **428**, 281–286.
8. West,S., Gromak,N. and Proudfoot,N.J. (2004) Human 5′→3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, **432**, 522–525.
9. Brion,P. and Westhof,E. (1997) Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.*, **26**, 113–137.
10. Tinoco,I.,Jr and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol.*, **293**, 271–281.
11. Doherty,E.A. and Doudna,J.A. (2001) Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 457–475.
12. Thirumalai,D., Lee,N., Woodson,S.A. and Klimov,D. (2001) Early events in RNA folding. *Annu. Rev. Phys. Chem.*, **52**, 751–762.
13. Schroeder,R., Barta,A. and Semrad,K. (2004) Strategies for RNA folding and assembly. *Nature Rev. Mol. Cell Biol.*, **5**, 908–919.
14. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell*, **81**, 991–1002.
15. Ferre-D'Amar,A.R., Zhou,K. and Doudna,J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature*, **395**, 567–574.
16. Golden,B.L., Gooding,A.R., Podell,E.R. and Cech,T.R. (1998) A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science*, **282**, 259–264.
17. Rupert,P.B. and Ferre-D'Amare,A.R. (2001) Crystal structure of a hairpin ribozyme–inhibitor complex with implications for catalysis. *Nature*, **410**, 780–786.
18. Adams,P.L., Stahley,M.R., Kosek,A.B., Wang,J. and Strobel,S.A. (2004) Crystal structure of a self-splicing group I intron with both exons. *Nature*, **430**, 45–50.
19. Golden,B.L., Kim,H. and Chase,E. (2005) Crystal structure of a phage Twort group I ribozyme–product complex. *Nature Struct. Mol Biol.*, **12**, 82–89.
20. Guo,F., Gooding,A.R. and Cech,T.R. (2004) Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol. Cell*, **16**, 351–362.
21. Cech,T.R. (1990) Self-splicing of group I introns. *Annu. Rev. Biochem.*, **59**, 543–568.
22. Beaudry,A.A. and Joyce,G.F. (1990) Minimum secondary structure requirements for catalytic activity of a self-splicing group I intron. *Biochemistry*, **29**, 6534–6539.
23. van der Horst,G. and Inoue,T. (1993) Requirements of a group I intron for reactions at the 3′ splice site. *J. Mol. Biol.*, **229**, 685–694.
24. Inoue,T., Sullivan,F.X. and Cech,T.R. (1986) New reactions of the ribosomal RNA precursor of *Tetrahymena* and the mechanism of self-splicing. *J. Mol. Biol.*, **189**, 143–165.
25. Michel,F. and Westhof,F. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.*, **216**, 585–610.
26. Joyce,G.F., van der Horst,G. and Inoue,T. (1989) Catalytic activity is retained in the *Tetrahymena* group I intron despite removal of the large extension of element P5. *Nucleic Acids Res.*, **17**, 7879–7889.
27. Suh,S.O., Jones,K.G. and Blackwell,M. (1999) A Group I intron in the nuclear small subunit rRNA gene of *Cryptendoxyla hypophloia*, an ascomycetous fungus: evidence for a new major class of Group I introns. *J. Mol. Evol.*, **48**, 493–500.
28. Xiao,M., Leibowitz,M.J. and Zhang,Y. (2003) Concerted folding of a Candida ribozyme into the catalytically active structure posterior to a rapid RNA compaction. *Nucleic Acids Res.*, **31**, 3901–3908.
29. Zhang,L., Xiao,M., Lu,C. and Zhang,Y. (2005) Fast formation of the P3–P7 pseudoknot: a strategy for efficient folding of the catalytically active ribozyme. *RNA*, **11**, 59–69.
30. Li,Z. and Zhang,Y. (2005) Predicting the secondary structures and tertiary interactions of 211 group I introns in IE subgroup. *Nucleic Acids Res.*, **33**, 2118–2128.
31. Lehnert,V., Jaeger,L., Michel,F. and Westhof,E. (1996) New loop–loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.*, **3**, 993–1009.
32. Doherty,E.A., Herschlag,D. and Doudna,J.A. (1999) Assembly of an exceptionally stable RNA tertiary interface in a group I ribozyme. *Biochemistry*, **38**, 2982–2990.
33. Khvorova,A., Lescoute,A., Westhof,E. and Jayasena,S.D. (2003) Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature Struct. Biol.*, **10**, 708–712.
34. De la Pena,M., Gago,S. and Flores,R. (2003) Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity. *EMBO J.*, **22**, 5561–5570.
35. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, NY.
36. Zhang,Y. and Leibowitz,M.J. (2001) Folding of the group I intron ribozyme from the 26S rRNA gene of *Candida albicans*. *Nucleic Acids Res.*, **29**, 2644–2653.
37. Hampel,K.J. and Burke,J.M. (2001) Time-resolved hydroxyl-radical footprinting of RNA using Fe(II)-EDTA. *Methods*, **23**, 233–239.

38. van der Horst,G., Christian,A. and Inoue,T. (1991) Reconstitution of a group I intron self-splicing reaction with an activator RNA. *Proc. Natl Acad. Sci. USA*, **88**, 184–188.
39. Ikawa,Y., Shiraishi,H. and Inoue,T. (1998) *Trans*-activation of the *Tetrahymena* ribozyme by its P2-2.1 domains. *J Biochem. (Tokyo)*, **123**, 528–533.
40. Engelhardt,M.A., Doherty,E.A., Knitt,D.S., Doudna,J.A. and Herschlag,D. (2000) The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry*, **39**, 2639–2651.
41. Zarrinkar,P.P. and Williamson,J.R. (1996) The P9.1–P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme. *Nucleic Acids Res.*, **24**, 854–858.
42. Peyman,A. (1994) P2 fuctions as a spacer in the *Tetrahymena* ribozyme. *Nucleic Acids Res.*, **22**, 1383–1388.
43. Pan,J. and Woodson,S.A. (1999) The effect of long-range loop–loop interactions on folding of the *Tetrahymena* self-splicing RNA. *J. Mol. Biol.*, **294**, 955–965.
44. Waldsich,C., Masquida,B., Westhof,E. and Schroeder,R. (2002) Monitoring intermediate folding states of the *td* group I intron *in vivo*. *EMBO J.*, **21**, 5281–5291.
45. Zhuang,X., Bartley,L.E., Babcock,H.P., Russell,R., Ha,T., Herschlag,D. and Chu,S. (2000) A single-molecule study of RNA catalysis and folding. *Science*, **288**, 2048–2051.
46. Rangan,P., Masquida,B., Westhof,E. and Woodson,S.A. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA*, **100**, 1574–1579.
47. Ikawa,Y., Shiraishi,H. and Inoue,T. (2000) Minimal catalytic domain of a group I self-splicing intron RNA. *Nature Struct. Biol.*, **7**, 1032–1035.