
In vitro characterization of the splicing efficiency and fidelity of the Rmlnt1 group II intron as a means of controlling the dispersion of its host mobile element

ISABEL CHILLÓN,^{1,2,4,5} MARÍA DOLORES MOLINA-SÁNCHEZ,^{1,4} OLGA FEDOROVA,^{2,3} FERNANDO MANUEL GARCÍA-RODRÍGUEZ,¹ FRANCISCO MARTÍNEZ-ABARCA,¹ and NICOLÁS TORO¹

¹Grupo de Ecología Genética, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, 18008 Granada, Spain

²Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA

³Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA

ABSTRACT

Group II introns are catalytic RNAs that are excised from their precursors in a protein-dependent manner in vivo. Certain group II introns can also react in a protein-independent manner under nonphysiological conditions in vitro. The efficiency and fidelity of the splicing reaction is crucial, to guarantee the correct formation and expression of the protein-coding mRNA. Rmlnt1 is an efficient mobile intron found within the *ISRM2011-2* insertion sequence in the symbiotic bacterium *Sinorhizobium meliloti*. The Rmlnt1 intron self-splices in vitro, but this reaction generates side products due to a predicted cryptic IBS1* sequence within the 3' exon. We engineered an Rmlnt1 intron lacking the cryptic IBS1* sequence, which improved the fidelity of the splicing reaction. However, atypical circular forms of similar electrophoretic mobility to the lariat intron were nevertheless observed. We analyzed a run of four cytidine residues at the 3' splice site potentially responsible for a lack of fidelity at this site leading to the formation of circular intron forms. We showed that mutations of residues base-pairing in the tertiary EBS3–IBS3 interaction increased the efficiency and fidelity of the splicing reaction. Our results indicate that Rmlnt1 has developed strategies for decreasing its splicing efficiency and fidelity. Rmlnt1 makes use of unproductive splicing reactions to limit the transposition of the insertion sequence into which it inserts itself in its natural context, thereby preventing potentially harmful dispersion of *ISRM2011-2* throughout the genome of its host.

Keywords: RNA circles; Rmlnt1; group II intron; ribozyme; splicing

INTRODUCTION

Group II introns are catalytic RNAs that are excised from their precursor RNA in a protein-dependent manner in vivo (Pyle and Lambowitz 2006; Lambowitz and Zimmerly 2010). Self-splicing can occur in the absence of proteins in vitro, if these introns are incubated under nonphysiological conditions (Pyle 2010). As in spliceosomal introns, the ribozyme of group II introns excises the intron as a branched, lariat structure, through two sequential transesterification reactions (Fig. 1A). In the first step, the 2' OH group of a bulged adenosine in domain 6 acts as the nucleophile for an attack on the 5' splice site (5' SS), producing an intron lariat–3' exon intermediate. In the second step, the 3' OH of the cleaved 5' exon, which remains tightly bound to the intron via EBS1–IBS1 and

EBS2–IBS2 (exon and intron binding sites 1 and 2, respectively) base-pairing interactions, acts as the nucleophile for an attack on the 3' splice site (3' SS), resulting in exon ligation and the excision of an intron lariat RNA (van der Veen et al. 1986; Pyle and Lambowitz 2006; Toro et al. 2007). A hydrolytic pathway, in which a water molecule acts as the nucleophile during the first step of splicing, can also occur, producing ligated exons and a linear intron (Fig. 1B; Jarrell et al. 1988; Daniels et al. 1996; Fedorova and Zingler 2007). A third pathway involves the formation of intron circles, in which the 5' and 3' extremities of the intron are covalently linked by a putative 5'–2' bond (Fig. 1C; Murray et al. 2001). This alternative splicing pathway has been described both in vitro (Murray et al. 2001; Nagy et al. 2013) and in vivo (Molina-Sanchez et al. 2006, 2011). It requires the presence of a 5' exon–intron intermediate formed by *trans*-splicing, the prior release of the 3' exon and the successive attack of the 2' OH group of the 3' end of the intron

⁴These authors contributed equally to this work.

⁵Present address: Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA

Corresponding author: nicolas.toro@eez.csic.es

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.047407.114>. Freely available online through the RNA Open Access option.

© 2014 Chillón et al. This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

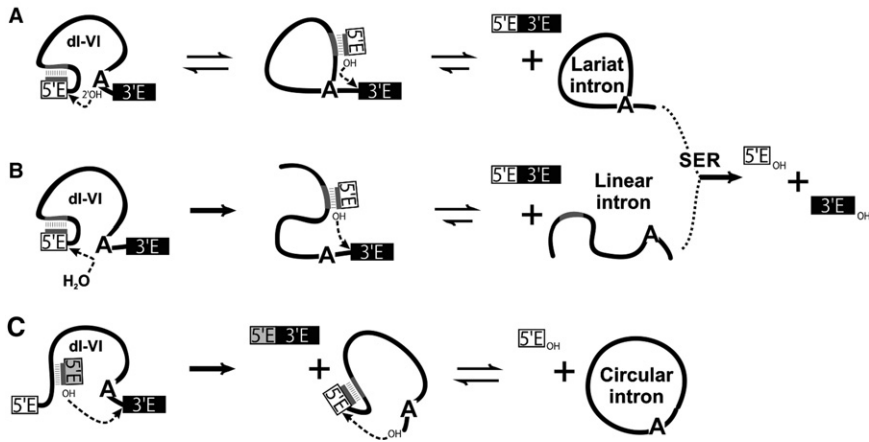


FIGURE 1. Mechanisms of group II intron excision. Schematic representations of the branching (A), hydrolytic (B), and circle formation (C) pathways are shown yielding lariat, linear, and circular intron, respectively. The different mechanisms are described in the text. The intron is indicated by a black line, while the exon sequences are represented as white/gray (5' exon) or black (3' exon) boxes. The bulged adenosine in the intron domain VI is represented as A. Dotted lines indicate the nucleophilic attack during each step of reaction.

to the 5' SS. In vitro, group II introns can catalyze spliced-exon reopening (SER), a reaction involving the hydrolysis of ligated exons to generate free 5' and 3' exons (Jarrell et al. 1988). Regardless of the splicing pathway, the free spliced intron can reinsert itself in vivo into target loci within the genome of the host organisms and can be reverse-transcribed by the intron-encoded reverse transcriptase during retrohoming (Cousineau et al. 2000; Muñoz-Adelantado et al. 2003; Zhong and Lambowitz 2003; Martínez-Abarca et al. 2004).

RmInt1 is a mobile bacterial group II intron found within the *ISRm2011-2* insertion sequence in *Sinorhizobium meliloti*, the nitrogen-fixing symbiont of alfalfa (*Medicago sativa*) (Martínez-Abarca et al. 1998). It is a bacterial IIB3/class D intron and has a class IIB-like RNA structure, with some IIA features (Fig. 2; Zimmerly et al. 2001; Toro et al. 2003). RmInt1 self-splices in vitro through the branching and hydrolytic pathways (Costa et al. 2006a). However, the second step in splicing is particularly inefficient, resulting in the accumulation of a lariat/3' exon intermediate and poor exon ligation. RmInt1 splicing also generates aberrant products, demonstrating a poor splicing fidelity for this intron. For instance, an unexpected truncated lariat/3' exon intermediate, with only 10 or 11 nt in its 3' exon, has been detected (Costa et al. 2006a). It has been suggested that this aberrant product is generated by ribozyme-promoted cleavage 3' to the GACGAA or GACGAAG sequence (hereafter designated IBS1*) in the 3' exon, which better matches the EBS1 (UUUCGUC) sequence than the authentic IBS1 (GAUGAGA) sequence in the 5' exon (Costa et al. 2006b). We found that in vivo RmInt1 splicing led to the accumulation of lariat forms but that this process was inefficient, shown by a low detection of ligated exons (Chillón et al. 2011). Thus, RmInt1 functions more like a retroelement than spliceosomal intron preventing the spread of *ISRm2011-2* in the *S. meliloti* genome.

In group II introns, certain nucleotides are conserved at intron boundaries. The consensus sequence at the 5' end of the intron is GUGYG, whereas the consensus sequence at the 3' end is AY (where Y is C or U). Some variations have been reported, associated with the various specific RNA subgroups (Dai et al. 2008; Candales et al. 2012). Remarkably, the RmInt1 intron shows a deviation from the consensus sequence at its 3' end (CC) (Fig. 2). The nucleotide at the 3' end is engaged in a Watson–Crick interaction with a purine located in the joining segment separating domain 2 from domain 3 (J2/3) (γ – γ' interaction) (Jacquier and Michel 1990). The penultimate nucleotide is predicted to be involved in a non-Watson–Crick interaction with the first nucleotide of the intron, occurring after lariat formation

(Chanfreau and Jacquier 1993). The presence of a pyrimidine at the penultimate position of RmInt1 differentiates this intron from all other group II introns in terms of consensus sequences at the 3' end (AY), but also from eukaryotic pre-mRNA introns (consensus sequence at the 3' end: AG), which are thought to be the phylogenetic descendants of group II introns (Sharp 1991; Fica et al. 2013). In addition, the CC dinucleotide at the 3' end of RmInt1 is followed by another pair of cytidine residues at the 5' end of the 3' exon. The first C nucleotide in the 3' exon is involved in a Watson–Crick base-pairing, as part of the EBS3–IBS3 interaction (exon and intron binding sites 3, respectively) with a guanosine residue located in the ID⁽ⁱⁱⁱ⁾–ID^(iv) internal loop. Thus, the 3' SS of RmInt1 forms a characteristic run of four cytidine residues (CCCC), potentially leading to the formation of an ambiguous 3' SS.

We investigated the reasons for the low efficiency and fidelity of RmInt1 in vitro. We found that deletion of the predicted IBS1* resulted in a significant decrease of side products. We also analyzed the role of the run of four cytidine residues at the 3' SS and revealed that this sequence was responsible for the formation of aberrant products (truncated circles or circles carrying extra residues) migrating with the lariat intron. These observations provide insight into the biological role of RmInt1, which requires impairment of its enzyme to limit the potentially harmful dispersion of other mobile genetic elements in the host genome.

RESULTS

The splicing fidelity of the RmInt1 intron is affected by an alternative EBS1–IBS1 pairing in vitro

It has been suggested that a putative alternative IBS1 sequence (IBS1*) located in the 3' exon, may compete with

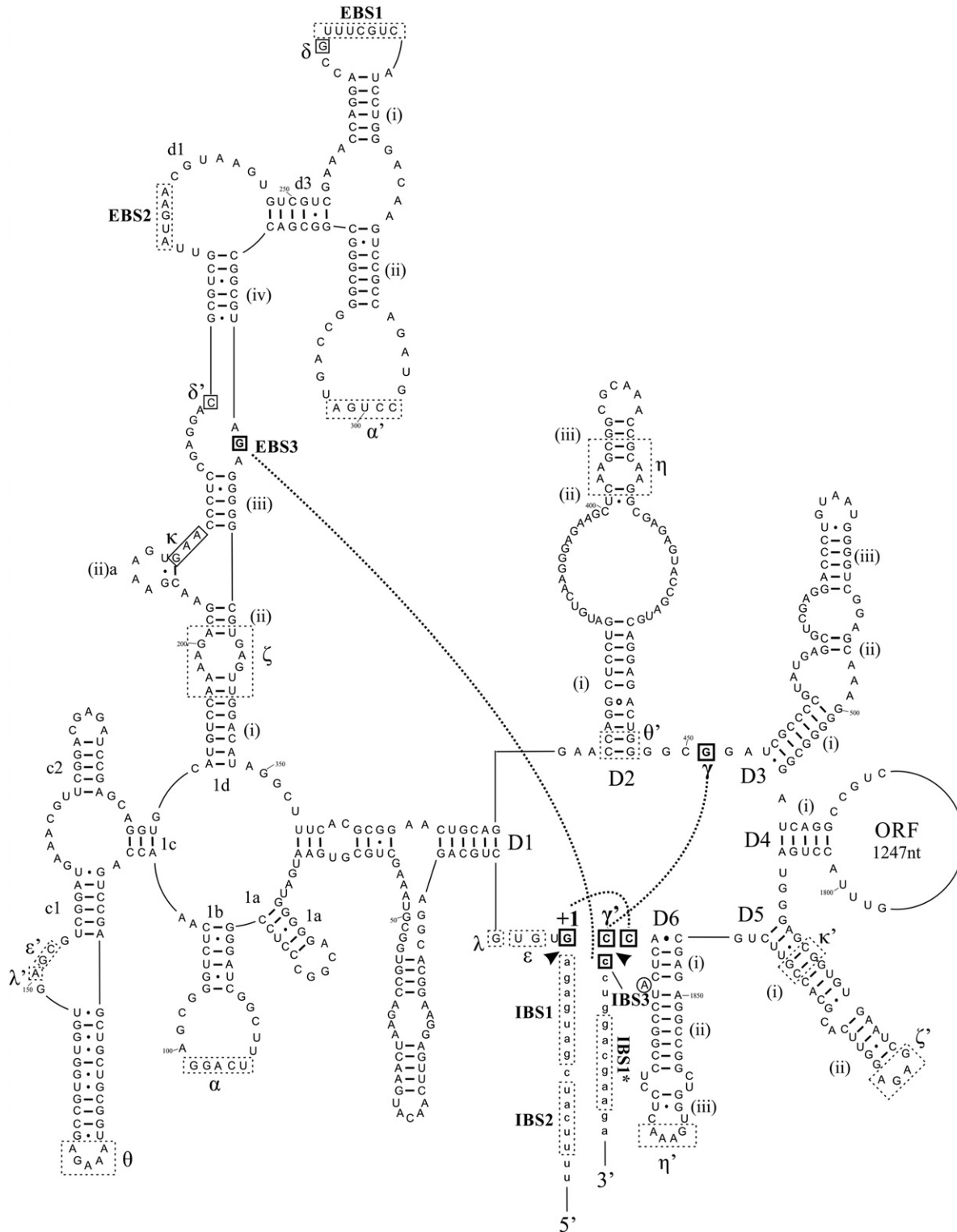


FIGURE 2. Diagram of the secondary structure of the RmInt1 group II intron. The nucleotides at the 3' splice site (3' SS) and the nucleotides with which they have been reported to interact are highlighted in bold letters and enclosed in boxes with solid lines. Long-range interactions between these nucleotides are denoted by dashed bold lines. Sequence elements involved in other long-range interactions are indicated by dashed boxes. Arrowheads indicate the 5' and 3' SS, respectively. Uppercase letters correspond to intron sequences, whereas lowercase characters correspond to exon sequences.

the authentic sequence (located in D1) for the RmInt1 intron during splicing in vitro, leading to the formation of an unconventional product (Fig. 3B, right panel; Costa et al. 2006a). We investigated the effect of the IBS1* sequence on RmInt1 splicing in vitro, by deleting this sequence from the 3' exon and analyzing the splicing rate and products (Fig. 3B).

We trimmed the 3' exon of the Δ ORF-WT RmInt1 intron construct (Costa et al. 2006a) to its first 5 nt, retaining the IBS3 sequence but not the IBS1* sequence. We then added a region corresponding to the *Escherichia coli lacZ* gene, to enlarge the trimmed 3' exon, such that it was of similar length to Δ ORF-WT (Fig. 3A). The Δ ORF-WT construct and the newly engineered Δ ORF-WT2 construct were then used in a self-splicing assay in vitro, with analyses carried out at various time points during the reaction (Fig. 3B). As expected, Δ ORF-WT gave rise to multiple side products, in both the branching and hydrolytic pathways, and these products had previously been shown to contain an extra 10 or 11 nt from the 3' SS (Costa et al. 2006a). In contrast, Δ ORF-WT2 yielded products corresponding in size to the lariat intron or circle, lariat intron–3' exon intermediate, and linear intron. The linear intron–3' intermediate could not be distinguished from the intron precursor due to the short extension of the 5' exon, although this form was taken into account for the calculation of reaction kinetics (see Materials and Methods). Interestingly, kinetic analyses showed similar apparent first-order rate constants for the conversion of precursor RNA into products for both the Δ ORF-WT and Δ ORF-WT2 transcripts (0.0120 ± 0.0008 and $0.0127 \pm 0.0007 \text{ min}^{-1}$, respectively). However, the degree of completion of the reaction differed significantly for the two constructs: 82% of the Δ ORF-WT2 precursor reacted, versus only 65% of the Δ ORF-WT precursor, suggesting a role for IBS1* in affecting the amplitude of the reaction (Fig. 3C). We also ruled out a possible effect of the *lacZ* sequence in the 3' exon, by reacting another transcript trimmed to 5 nt but lacking the *lacZ* sequence (Fig. 3C). The rate and amplitude of the reaction were similar for this short transcript ($0.0119 \pm 0.0006 \text{ min}^{-1}$ and 86%, respectively) and for Δ ORF-WT2.

Thus, the IBS1* sequence located in the 3' exon decreases the fidelity and amplitude of the RmInt1 self-splicing reaction in vitro.

RmInt1 displays a lack of fidelity for circle formation, but that is not translated to ligated exons

We characterized the bands of sizes corresponding to the lariat intron or circle from the Δ ORF-WT2 self-splicing reaction by gel extraction, followed by reverse transcription with the Ect1 primer and further amplification with primers LL and +97, followed by the sequencing of individual clones (Fig. 4A). Together with the expected intron lariat form, we detected circular products: an intron form with a truncated domain 6 linked either to the extremity of the 5' exon or the beginning of the intron, and a circular intron form con-

taining an extra residue (cytidine in position +1 with respect to the 3' SS) (Table 3, below; Fig. 4B). We investigated the formation of correct ligated exons by examining the sequence of the splice junction, which was amplified from an unfractionated in vitro splicing reaction of Δ ORF-WT2 by RT-PCR with primers binding to the 3' and 5' exons (Fig. 4A). All the sequences analyzed (35) corresponded to the expected exon junction.

Our sequencing data show that, along with the lariat intron, atypical circular forms were produced during the splicing reaction of the Δ ORF-WT2 transcript whereas the ligated exons we analyzed were likely formed from typical splicing reactions leading to lariat formation. Nevertheless, we cannot rule out that other low abundant exonic products were formed alongside the atypical circles could be produced, and that our methods were not sensitive enough to detect them.

Analyses of mutations at the 3' splice site reveal an intrinsic defect of the RmInt1 intron affecting splicing fidelity at the 3' splicing junction

We hypothesized that the unusual run of four cytidines at the 3' SS of RmInt1 might be responsible for the lack of fidelity and the formation of atypical circle molecules. We tested this hypothesis by introducing nucleotide substitutions at positions 1883 (penultimate) and 1884 (γ') of the intron and at positions +1 (IBS3) of the 3' exon of Δ ORF-WT2 (we numbered intron positions in accordance with the full-length RmInt1 intron, although this construct lacks the ORF sequence) (Fig. 2). As these nucleotides at the 3' SS are involved in long-range interactions with nucleotides at other positions in the intron sequence (Fig. 2; Jacquier and Michel 1990; Costa et al. 2000), we also introduced the corresponding compensatory substitutions at positions 329 (EBS3) and 452 (γ) of the intron. We analyzed the reaction rates and products (Table 1; Fig. 5), and compared them with the ligated exon rates calculated for each mutant (Table 2). We also cloned and sequenced the splicing products of selected mutants (Table 3).

Substitutions at the penultimate position of the intron (C1883) did not impair the first step of splicing, but all such substitutions resulted in defects in the second splicing reaction, with the 1883G and 1883U mutants the most severely affected (Table 1). As expected, this defect in the second splicing reaction resulted in a decrease in the rate of exon ligation (Table 2). The 1883A mutation is the substitution most frequently observed in natural conditions, potentially accounting for the kinetics of the splicing reaction and the ligated exon rate, which were similar to those for the wild-type sequence. In addition, sequence analysis on the products of the excision reaction for the 1883U and 1883A mutants showed lariat and circular products, including truncated domain 6 linked to 5' exon/intron circles and circular products with an extra cytidine residue (Table 3). Interestingly, true circles were detected for the 1883U mutant.

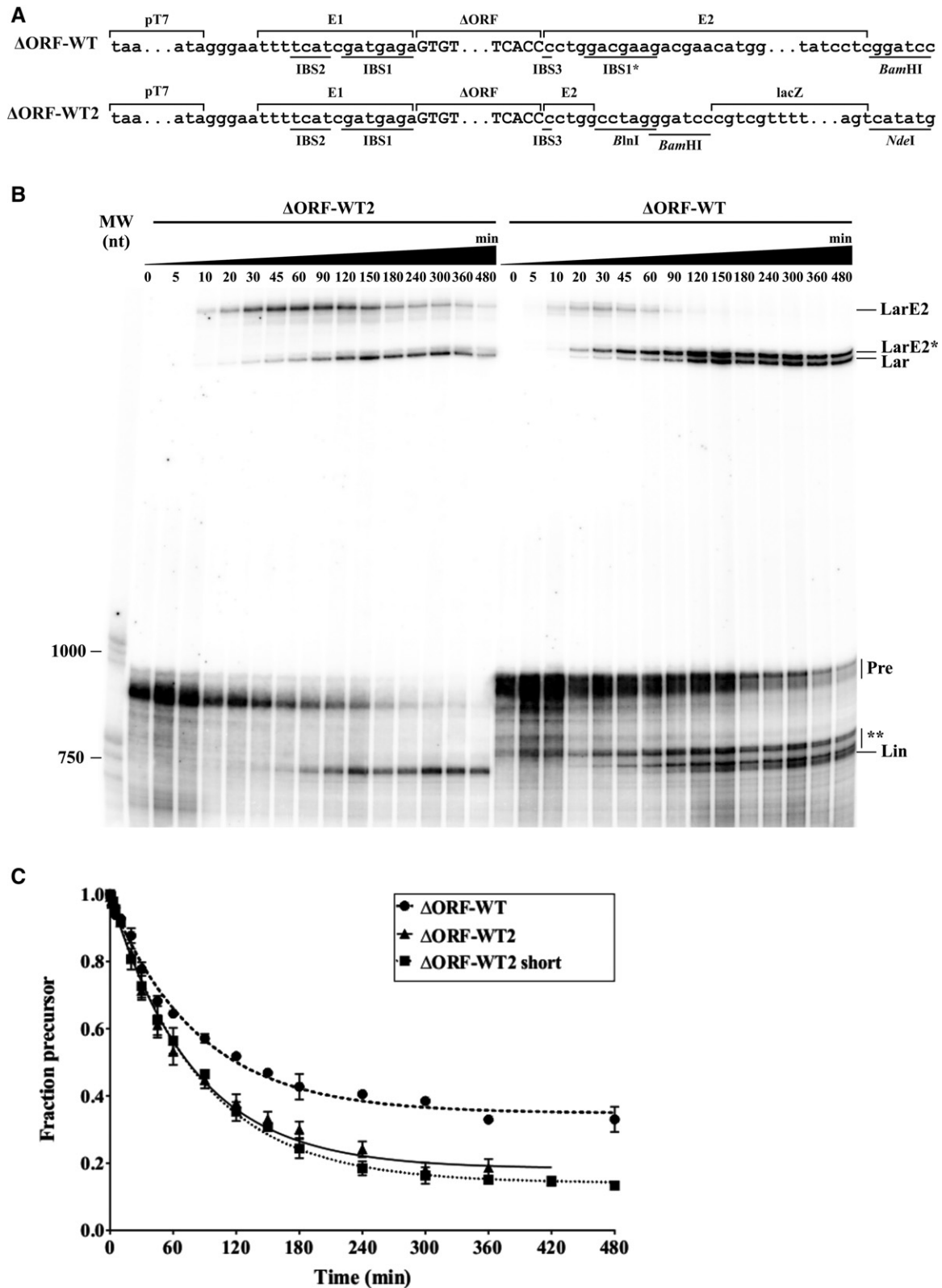


FIGURE 3. Self-splicing assay for RmInt1 Δ ORF-WT and Δ ORF-WT2 transcripts. (A) Schematic diagram of the RmInt1 WT constructs used in this study. Δ ORF-WT was previously described as Δ 15 or pLM1 (Costa et al. 2006a,b). (B) Self-splicing assay for the Δ ORF-WT (right panel) and Δ ORF-WT2 (left panel) transcripts. Reaction time points and nucleotide size markers are indicated at the top and the left of the gel, respectively. Possible interpretations of the various bands observed are provided on the right. (Lar) intron lariat; (LarE2) intron lariat-3' intermediate; (LarE2*) intron lariat-3' intermediate with truncation of the first 10 or 11 nt of the 3' intron; (Pre) precursor; (Lin) intron linear, and (**) the presence of other linear side products of various lengths. (C) The reactivity of three RmInt1 WT transcripts is shown. Experimental data for the disappearance of precursor RNA over time were fitted with a first-order exponential equation with endpoint correction. Error bars indicate the standard error of the mean over two or three trials.

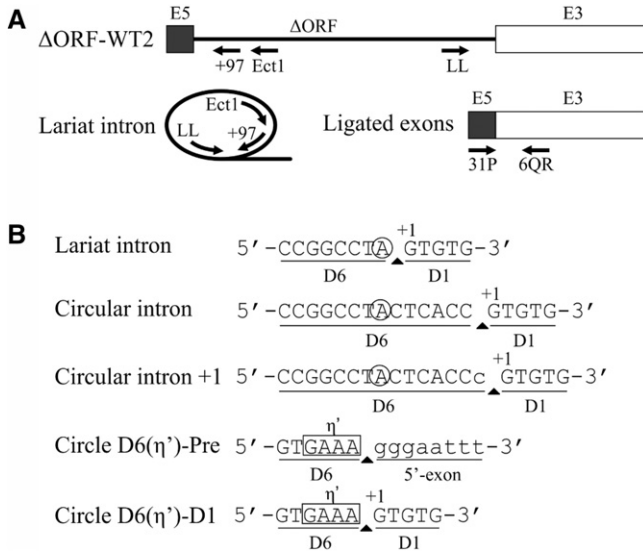


FIGURE 4. Identification of splicing forms. (A) Schematic diagram of the Δ ORF-WT2 precursor and ligated exons with the primers (arrows) used for reverse transcription and PCR. (B) Sequences of the lariat and circular forms produced after splicing of the Δ ORF-WT2 transcript. A circular intron, which was not observed for the wild-type Δ ORF-WT2 transcript, is also shown for clarity. Uppercase letters correspond to intron sequences, whereas lowercase characters correspond to exon sequences. The covalent bond joining the two ends to form a circle is indicated by an arrowhead. The bulging A is surrounded by a circle. The first nucleotide of the intron is indicated by +1. D1 is domain 1 of the intron and D6 is domain 6 of the intron. η' is the tetraloop involved in the tetraloop-receptor η - η' tertiary interaction between intron domains 2 and 6.

Substitutions involving the γ - γ' interaction resulted in a lower efficiency of the first step of splicing, regardless of the maintenance of the Watson-Crick interaction (Table 1). The second step of splicing was also impaired in all the mutants, but mutations resulting in disruption of the Watson-Crick interaction resulted in greater defects, with the production of larger amounts of some low-mobility forms (Fig. 5). Interestingly, sequencing of the rapidly migrating bands after gel extraction for the γ - γ' -CC mutant revealed the presence of true circles, circles carrying an extra C and atypical circles formed by a truncated domain 6 linked to the end of the 5' exon, but no lariat product was detected (Table 3). The γ - γ' mutants maintaining Watson-Crick pairings displayed a more moderate effect on the second splicing step and ligated exon rates, and only the γ - γ' -AU mutant, which is as frequent in natural conditions as the WT G(γ)C(γ') (Dai et al. 2003; Candales et al. 2012) displayed lower levels of accumulation of atypical low-mobility forms (Fig. 5), and an enrichment in the lariat product (Table 3).

Finally, substitutions involving the IBS3-EBS3 interaction did not impair the first step of splicing. Moreover, only the EBS3-IBS3-GG and EBS3-IBS3-CC mutants displayed defects in the second step of splicing, resulting in lower ligated exon rates (Tables 1, 2). The EBS3-IBS3-CG and EBS3-IBS3-UA mutants, for which Watson-Crick base-pairing was maintained, and intriguingly, also the EBS3-IBS3-AA

mutant, for which Watson-Crick interaction was not conserved, displayed no defects in the second splicing step (Tables 1, 2). The EBS3-IBS3-CG and EBS3-IBS3-AA mutants resulted in an enrichment in lariat forms, whereas the EBS3-IBS3-UA mutant yielded a mixture of lariat and circular splicing products.

Our results suggest that the run of four cytidine residues in this intron plays an important role in regulating the fidelity of the RmInt1 splicing reaction and the formation of circular molecules, reflecting the high degree of complexity of these long-range interactions.

DISCUSSION

Group II intron splicing is a complex process involving a large number of tertiary intramolecular interactions, with fine structural rearrangements of the intron occurring between the first and the second splicing step reactions (Marcia and Pyle 2012; Marcia et al. 2013). We showed that the removal of a sequence predicted to interfere with the splicing reaction from the 3' exon led to an increase in splicing product fidelity. However, other atypical circular forms comigrating with the lariat intron were nevertheless observed. We hypothesize that the presence of a run of four cytidine residues in the 3' SS might be responsible for the activation of cryptic sites and the formation of atypical circular forms. We observed that certain mutations affecting the EBS3-IBS3 interaction led to an increase in the rate of lariat formation (fidelity) while maintaining the ligated exon rate (efficiency). Based on these observations, we suggest that RmInt1 has evolved to preserve a splicing reaction with an intrinsic low efficiency and fidelity. We interpret this as an adaptation favoring survival within the bacterial genome in which the RmInt1 lives, thus preventing the harmful insertion of the intron into essential sequences during its dispersion within the host genome.

A predicted alternative IBS1 sequence in the 3' exon affects splicing specificity

The RmInt1 intron has been shown to generate unusual splicing side products in vitro (Costa et al. 2006a). A sequence present within the 3' exon has been predicted to match the EBS1 site better than the authentic IBS1 sequence at the 3' end of the 5' exon, leading to hydrolysis at positions +10 or +11 of the 3' exon (Costa et al. 2006a). We deleted this sequence from the 3' exon and showed that this deletion greatly decreased the number of atypical splicing forms generated (Fig. 3B). Interestingly, the rate of reaction was not affected, as expected given that the competing pathway occurs after the completion of the first splicing step (Costa et al. 2006a). However, the amplitude of the reaction was almost 20% higher in the absence of IBS1* sequence. This suggests that the IBS1* sequence interacting with the EBS1 led to the misfolding of a fraction of the population of precursor

TABLE 1. Absolute and relative efficiencies for the first and second steps of the branching pathway during Rmlnt1 self-splicing in vitro

Construct assayed	First step (min ⁻¹) (%)	Second step (min ⁻¹) (%)
WT	0.01255 ± 0.00173 (100)	0.00944 ± 0.00434 (100)
γ-γ'-CC	0.00713 ± 0.00394 (57)	0.00096 ± 0.00052 (10)
γ-γ'-GG	0.00572 ± 0.00176 (46)	0.00059 ± 0.00035 (6)
γ-γ'-CG	0.00364 ± 0.00136 (29)	0.00179 ± 0.00083 (19)
γ-γ'-AU	0.00400 ± 0.00064 (32)	0.00230 ± 0.00061 (24)
γ-γ'-AA	0.00304 ± 0.00003 (24)	0.00047 ± 0.00011 (5)
EBS3-IBS3-CC	0.01935 ± 0.00672 (154)	0.00379 ± 0.00025 (40)
EBS3-IBS3-GG	0.00990 ± 0.00354 (79)	0.00165 ± 0.00059 (17)
EBS3-IBS3-CG	0.01260 ± 0.00156 (100)	0.01036 ± 0.00303 (110)
EBS3-IBS3-UA	0.01162 ± 0.00393 (93)	0.02115 ± 0.00106 (224)
EBS3-IBS3-AA	0.02685 ± 0.00926 (214)	0.01620 ± 0.00368 (172)
1883A	0.01350 ± 0.00438 (108)	0.00642 ± 0.00065 (68)
1883G	0.02025 ± 0.00544 (161)	0.00211 ± 0.00012 (22)
1883U	0.03075 ± 0.01520 (245)	0.00374 ± 0.00021 (40)

The reported rates were calculated with KinTek Explorer software (KinTek Corporation). Errors correspond to the standard deviation of the mean over two, three, or four independent experiments.

molecules, which precluded them from reacting. Alternative splice site sequences have been reported for several group II introns (Jacquier and Jacquesson-Breuleux 1991; Tourasse et al. 2005; Toor et al. 2006; Nagy et al. 2013). These regions generally include modified IBS1 motifs that translocate the intron interaction, generating alternative excision products, as observed for Rmlnt1 (Costa et al. 2006a). Rmlnt1 splicing in vivo is known to be assisted by the intron-encoded protein (IEP), which may even modulate the ratio of lariat and circular products (Molina-Sánchez et al. 2006). We therefore cannot rule out the possibility that the IEP impedes these alternative base-pairings in vivo or even that exon shuffling may be of biological relevance.

Mutations of the run of cytidines at the 3' splice site favor the production of lariat introns over unproductive circular forms

An analysis of the low-mobility splicing products of the ΔORF-WT2 transcript lacking the IBS1* sequence revealed the presence of unexpected circular forms comigrating with lariat products. These circular products included circular introns containing an extra residue belonging to the first nucleotide of the 3' exon and intron circles consisting of an intron form with a truncated domain 6 linked to the end of the 5' exon or the beginning of the intron. It has been suggested

that circles consisting of an intron with a truncated domain 6 linked to the end of the 5' exon may be generated by transesterification between the 3' OH group of the 3' end of the truncated intron and the precursor triphosphate 5' end (Costa et al. 2006a). Rmlnt1 domain 6 was generally truncated at the level of the η' motif, consisting of a GAAA tetraloop. We hypothesize that the sequence 5'-GGUGAAA-3' located at the distal portion of domain 6 can be recognized by EBS1, which is partially complementary to the above sequence, thus leading to hydrolysis and intron truncation at this position (Fig. 4B). However, the mechanism underlying the formation of intron forms with a truncated domain 6 linked to the 5' end of the intron remains to be elucidated. We observed no true circles for the wild-type intron, although these forms have been characterized for other group II introns in vitro (Murray et al. 2001; Nagy et al. 2013) and for Rmlnt1 intron in vivo (Molina-Sánchez et al. 2006). It has been suggested that circles are generated by two successive transesterification reactions: A free 5' exon attacks the 3' SS, generating a 5' exon-intron intermediate, and the 2' OH of the terminal cytidine then reacts at the 5' splice site (5' SS) (Murray et al. 2001). This *trans*-splicing mechanism produces a circular intron together with ligated exons and a

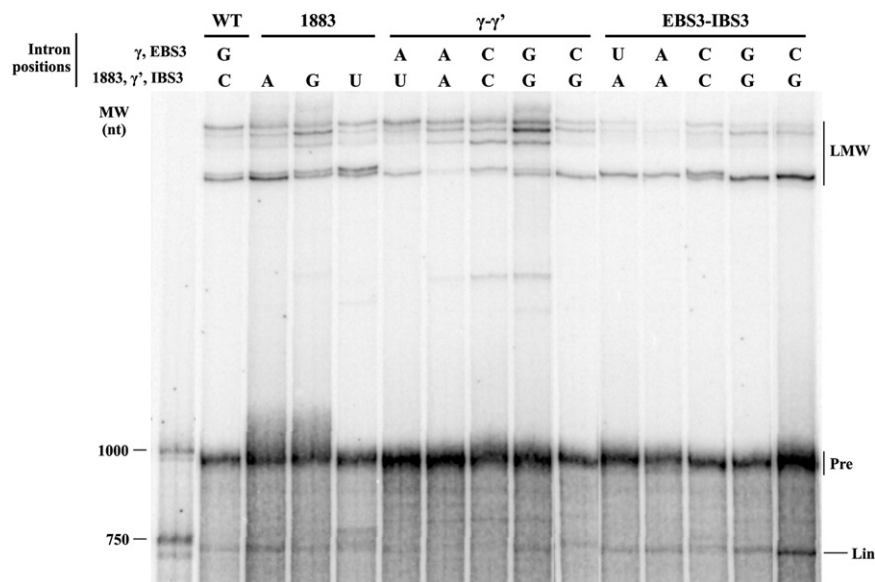


FIGURE 5. Self-splicing activity of the WT and mutant Rmlnt1 introns at positions 329 (EBS3), 452 (γ), and 1883 and 1884 (γ') of the intron and +1 (IBS3) of the 3' exon. Splicing reactions were carried out for 2 h and the products were run in parallel. LMW includes all low-mobility molecules of similar electrophoretic mobility to the lariat-3' intron intermediate and lariat intron; (Pre) the intron precursor; (Lin) the linear intron.

TABLE 2. WT and mutant RmInt1 splicing activity, as measured by the percentage of exons ligated

Construct assayed	LE (%) ^a
WT	100
γ - γ' -CC	0.1 \pm 0.0
γ - γ' -CG	1.4 \pm 0.2
γ - γ' -CG	3.4 \pm 0.9
γ - γ' -AU	8.8 \pm 2.2
γ - γ' -AA	0.1 \pm 0.0
EBS3-IBS3-CC	5.6 \pm 1.0
EBS3-IBS3-GG	5.0 \pm 1.3
EBS3-IBS3-CG	83.6 \pm 45.1
EBS3-IBS3-UA	152.5 \pm 108.4
EBS3-IBS3-AA	55.7 \pm 12.3
1883A	82.8 \pm 4.2
1883G	14.9 \pm 0.8
1883U	36.1 \pm 3.1

We amplified and quantified ligated exons by RT-qPCR from an unfractionated splicing reaction after 2 h using the 30P_5exonLM and 6QR_LACZ primers, which span the splicing junction.

^aThe error is the standard deviation over two independent experiments. The results for each experiment are the mean values of two or three trials.

free 5' exon, which can reinitiate the mechanism in another precursor molecule. We hypothesize that circles containing an extra nucleotide may form due to a lack of fidelity at the 3' SS, the residues of which have been shown to affect the second step of splicing (Jacquier and Michel 1990; Chanfreau and Jacquier 1993; Costa et al. 2000). The RmInt1 intron 3' SS consists of a run of four cytidine residues. We therefore created a series of mutant introns to investigate the role of this run of cytidine residues in the fidelity of RmInt1 excision at the 3' SS. These mutants were designed to mimic the residue composition of the 3' SS of high-fidelity introns, such as the ai5 γ intron of the yeast mitochondrial *oxi 3* gene (Arnberg et al. 1980), the PL.LSU/2 intron from mitochondria of the alga *Pylaiella littoralis* (Costa et al. 1997) and the bacterial intron *O.i.II* from *Oceanobacillus iheyensis* (Takaki et al. 2004). We examined the efficiency of the first and second splicing steps by carrying out a kinetic analysis of the rates and products of the reaction and assessing its fidelity by sequencing the splicing products of these mutants (Tables 1–3; Fig. 5). We showed that mutations affecting the γ - γ' interaction (γ - γ' -AU) increased the fidelity of the reaction, as shown by the smaller number of atypical splicing products and an enrichment in lariat forms. However, the efficiency of both the first and second steps of the reaction was diminished, resulting in a low rate of exon ligation. According to the available crystal struc-

tures (Marcia and Pyle 2012; Marcia et al. 2013), the γ (452) position is not located close to the catalytic center. However, modifications to the γ' (1884) position, which is adjacent to the scissile phosphate, may interfere in the general architecture of the intron, affecting the availability of the catalytic elements and, thus, the efficiency of the reaction. Furthermore, mutations affecting the EBS3-IBS3 interaction increased both the efficiency and fidelity of the RmInt1 intron splicing reaction. EBS3-IBS3-CG and EBS3-IBS3-AA mutants displayed no impairment of reaction efficiency, and both produced a higher proportion of lariat forms. The EBS3-IBS3-AA mutant may restore the WT phenotype by means of a non-Watson-Crick interaction between the two adenosine residues, a combination that has not been observed in natural conditions (Simon et al. 2008). According to the available group II intron structures (Marcia and Pyle 2012; Marcia et al. 2013), two adenosine residues in the EBS3 and IBS3 positions would be able to engage in *cis*-Watson-Crick/Watson-Crick or *cis*-Watson-Crick/sugar-edge interactions (Leontis and Westhof 2001; Leontis et al. 2002; Nasalean et al. 2009).

The increase in the proportion of lariat forms was accompanied by high rates of exon ligation for the EBS3-IBS3-CG and EBS3-IBS3-AA mutant introns. In contrast, the γ - γ' -CC mutant yielded only circular forms and hardly any ligated exons. The γ - γ' -CC mutant has been reported to produce lariat-derived products *in vivo* (Molina Sanchez et al. 2011), but the methods used in those assays did not distinguish between intermediate forms linked to the 3' exon and fully excised lariat intron. In addition, the γ - γ' -AA mutant, which also displayed a strong defect of exon ligation, gave rise to far fewer excised products of the size of a lariat/circular intron (Fig. 5). Then, we observe a correlation between lariat intron formation and ligated exons. We suggest that atypical circle formation would be associated to higher rates of free exons, since we were not able to detect incorrect ligated exons.

TABLE 3. Sequence analysis for the clones obtained by RT-PCR from low-mobility splicing products of mutant RmInt1 intron constructs

Construct assayed	Splicing products detected				
	Lariat	Circle	Circle +1 ^a	Circle D6-Pre ^b	Circle D6-D1 ^c
WT	1/11 (9%)	0/11 (0%)	3/11 (27%)	5/11 (45%)	2/11 (18%)
1883A	4/13 (31%)	0/13 (0%)	5/13 (38%)	4/13 (31%)	0/13 (0%)
1883U	3/23 (13%)	2/23 (9%)	9/23 (39%)	7/23 (30%)	2/23 (9%)
γ - γ' -CC	0/13 (0%)	4/13 (31%)	5/13 (38%)	4/13 (31%)	0/13 (0%)
γ - γ' -AU	13/16 (81%)	2/16 (13%)	1/16 (6%)	0/16 (0%)	0/16 (0%)
EBS3-IBS3-CG	8/12 (67%)	0/12 (0%)	2/12 (17%)	1/12 (8%)	1/12 (8%)
EBS3-IBS3-UA	3/13 (23%)	3/13 (23%)	4/13 (31%)	3/13 (23%)	0/13 (0%)
EBS3-IBS3-AA	9/16 (56%)	2/16 (13%)	3/16 (19%)	1/16 (6%)	1/16 (6%)

^aCircular product including an extra nucleotide from the 3' exon.

^bCircular product in which truncated D6 is linked to the 5' end of the transcript.

^cCircular product in which truncated D6 is linked to the 5' end of the intron.

The RmInt1 intron uses various mechanisms to control the dispersion of harmful genetic elements in its bacterial host

In natural conditions, the RmInt1 intron is found in the region encoding the transposase of an *IS_{Rm2011-2}* insertion sequence (Selbitschka et al. 1995; Martínez-Abarca et al. 1998). Several stop codons in the sequence of the intron result in a premature end to the translation of the transposase gene, rendering intron splicing essential for the mobility of the insertion sequence in the host genome (Martínez-Abarca et al. 1998). The location of the RmInt1 intron in a non-essential gene contrasts with the observed situation for mitochondrial group II introns, most of which are inserted into indispensable genes and require efficient splicing to ensure the survival of their hosts (Arnberg et al. 1980; Schmidt et al. 1990; Fontaine et al. 1995; Simon et al. 2008). With the lack of pressure for their efficient splicing, it has been suggested that RmInt1 and other bacterial group II introns behave essentially as retroelements rather than introns (Dai and Zimmerly 2002; Chillón et al. 2011). This study revealed some of the mechanisms used by RmInt1 to regulate its splicing efficiency and fidelity and, hence, the expression of the insertion sequence it interrupts. The first mechanism by which RmInt1 controls splicing concerns the ligation of exons by its transposase. It has already been demonstrated *in vivo* that the efficiency of exon ligation by the RmInt1 intron is extremely low, with <1% of exons correctly spliced (Chillón et al. 2011). It has already been suggested, and we confirm here, that the presence of an alternative IBS1 sequence in the 3' exon leads to hydrolytic cleavage at position +10 or +11 of the 3' exon, leading to the irreversible loss of a fraction of functional transposase mRNA molecules (Costa et al. 2006a). Finally, we suggest here that the existence of a run of four cytidine residues spanning the 3' SS, and probably leading to an increase in the generation of unproductive intron circles, also helps to reduce expression of the transposase. Indeed, similar runs of cytidine residues have been detected in other bacterial group II introns of class D, such as the Sr.md.I1 intron in *Sinorhizobium medicae*, the Pr.ae.I3 intron in *Prosthecochloris aestuarii*, the Pr.vi.I1 intron in *Prosthecochloris vibrioformis* and the Zu.pr.I1 intron in *Zunongwangia profunda* (Candales et al. 2012). Interestingly, all these introns interrupt non-essential genes encoding transposases and integrases, for example, suggesting that the same mechanism may operate in other bacterial group II introns.

CONCLUSION

The experiments reported here highlight various mechanisms by which the RmInt1 group II intron may fine-tune its splicing efficiency and fidelity as a means of controlling the dispersion of mobile genetic elements in the host genome. We confirmed that the predicted alternative IBS1* sequence in the 3' exon was able to decrease the splicing fidelity

of the RmInt1 intron. We also suggest that a run of cytidine residues in the 3' SS may be responsible for the lack of fidelity in recognition of the 3' SS. This may prevent the intron from adopting the three-dimensional conformation necessary for correct 3' splice site selection and exon ligation. We showed that mutations affecting the EBS3–IBS3 interaction improved efficiency and the fidelity at the 3' SS. Together, these mechanisms ensure low levels of expression of the insertion sequence interrupted by RmInt1 in its natural context, helping to ensure the survival of the bacterial host.

MATERIALS AND METHODS

Plasmid constructs

Δ ORF-WT2 was derived from Δ ORF-WT (previously named as Δ I5 or pLM1) (Costa et al. 2006a). Both constructs are derived from the pUC19 plasmid. The 3' end of Δ ORF-WT consists of 146 bp of the RmInt1 3' exon, corresponding to the sequence of *IS_{Rm2011-2}*, resulting in a 911-nt transcript. We created Δ ORF-WT2, by using the 5Q_CAT (5'-TGTACCTATAACCAGACC GTTCAG) and 27PR_lacZtranscript (5'-GGAGATCTCATATG ACTGTTGGGAAGGGCGA) primers to amplify a 1009-bp fragment of pICG- Δ ORF (Chillón et al. 2011), which was then inserted between the ClaI and BamHI sites of pLM1. Mutant Δ ORF-WT2-derivative plasmids were generated by cloning a ClaI–BglII fragment containing the mutated position. Site-directed mutagenesis was performed as described by Higuchi et al. (1988) (Supplemental Table S1). The sequences of the mutant constructs were checked by sequencing. Thus, the 3' region of Δ ORF-WT2 consists of 5 bp of the RmInt1 3' exon from *IS_{Rm2011-2}* and 145 bp of the *E. coli lacZ* gene (positions 42–169) flanked by restriction sites, making it possible to produce two transcripts of 775 and 908 nt, respectively.

In vitro transcription and purification

The Δ ORF-WT template was generated by BamHI digestion. Δ ORF-WT2 templates were generated by NdeI or BamHI digestion to produce the long and short transcripts, respectively. Transcription was performed with 2–4 units of T7 RNA polymerase, 4 μ g of linearized plasmid, 1 \times transcription buffer [10 \times transcription buffer: 150 mM MgCl₂, 400 mM Tris–HCl, pH 7.5, 20 mM Spermidine, and 50 mM DTT], 0.96 mM ATP, 0.96 mM CTP, 0.96 mM UTP, 0.064 mM GTP, 50 μ Ci [α -³²P]GTP (3000 Ci/mmol; 10 mCi/mL; PerkinElmer), 10 mM DTT and 80 units of RNaseOUT (Invitrogen), in a final volume of 50 μ L (Pyle and Green 1994). Transcription was stopped after 1.5 h at 37°C, by adding an equal volume of gel loading dye (10 M urea, 0.1% [w/v] xylene cyanol and bromophenol blue dyes, 40 mM Tris [pH 7.5], 8.3% [w/v] sucrose, and 0.83 mM EDTA) and running directly on a 5% (w/v) denaturing (7 M urea) polyacrylamide gel. After migration, the transcripts were visualized by autoradiography (10-min exposure), excised from the gel, and eluted by incubation for 3 h in elution buffer (0.3 M NaCl, 10 mM MOPS pH 6, 1 mM EDTA). RNA was then isolated by ethanol precipitation of the eluate (three volumes of 100% ethanol and 40 μ g of glycogen [Roche]). The RNA was dried and dissolved in 30 μ L of RNA storage buffer (10 mM MOPS pH 6, 1 mM EDTA).

Self-splicing assays

The precursor RNA was incubated in 80 mM MOPS buffer pH 7.5 at 95°C for 1 min. The RNA sample was then allowed to cool to reaction temperature before being combined with 500 mM (NH₄)₂SO₄ (final concentration). Reactions were initiated by adding 100 mM MgCl₂ (final concentration) and were incubated at 50°C. For time-course analyses, aliquots (1–2 μL) were removed at the indicated time points, added to quenching buffer (1.8% sucrose, 1× TBE [Tris–borate EDTA, pH 8.3], 0.018% xylene cyanol dye, 36% [v/v] formamide, and 25 mM EDTA) and placed on ice to stop the reaction. Samples were then loaded onto a 5% denaturing polyacrylamide gel and the products were resolved by electrophoresis. The gels were dried and placed against phosphorus plates (Imaging Plate 2040, Fujifilm). Images were acquired after 24–48 h of exposure, with the Personal Molecular Imager FX (Bio-Rad) laser scanning system, and quantified with Quantity One software (Bio-Rad).

Identification of in vitro splicing products

Total splicing reactions incubated for 2 h were subjected to electrophoresis in denaturing gels and low-mobility bands were purified. Reverse transcription and PCR were performed to detect the formation of lariat and circle forms, as previously described (Molina-Sanchez et al. 2006). The PCR products were inserted into the pGEM-T Easy Vector (Promega) and after transformation of *E. coli* cells the isolated clones were amplified by colony PCR and sequenced. For the identification of ligated exons, an unfractionated splicing reaction (2 h) mixture was subjected to reverse transcription with the 6QR_LACZ (5′-GATGTGCTGCAAGGCGATT) primer, followed by PCR amplification with this primer and 30P_5exonpLM (5′-GGGAATTTTCATCGATGAGA). The resulting PCR products were cloned and sequenced, as described above.

Kinetic analysis of time courses

Individual bands on autoradiographs were quantified with Quantity One software (Bio-Rad). Counts for a nearby position in the same lane at which there was no band present were subtracted, to provide an internal measurement of background RNA degradation at each time point. The relative fractions of precursor and products were calculated and corrected for the guanosine content of each species (Daniels et al. 1996). We then used Prism 6 (GraphPad Software, Inc.) to fit a single exponential decay equation with endpoint correction to the data for wild-type transcript reactivity. This equation took the form $(y_0 - y_{\text{final}}) \times e^{-k_{\text{obs}}(t)} + y_{\text{final}}$, where y_0 is the starting point, y_{final} is the endpoint, and k_{obs} is the reaction rate. Computer simulations were performed with KinTek Explorer (KinTek Corporation), for the fitting of the data.

SYBR Green quantitative PCR

We amplified and quantified ligated exons, using the 30P_5exonpLM and 6QR_LACZ primers, which span the splicing junction. The RT-qPCR experiments were conducted with a LightCycler 480 System (Roche). Each reaction was run in triplicate and contained 1 μg of cDNA template, 1 pmol of each oligonucleotide primer, 2× Master Mix I (Roche), and water, added to a final volume of 10 μL. The PCR cycling conditions were a 5-min hot start at 95°C, followed

by 45 cycles of denaturation at 95°C for 10 sec, annealing at 63°C for 2 sec, and extension at 72°C for 2 sec. Melting curve analysis was performed for all PCR runs, to check the identity of the PCR product and the specificity of the primers. We assessed mRNA levels by relative quantification, according to the 2^{−ΔΔCT} method (Livak and Schmittgen 2001). The normalized expression ratio was defined as 2^{−ΔΔCT}, where $\Delta\Delta C_T = \Delta C_T$ (calibrator or WT) − ΔC_T (test or mutant), and is expressed as a percentage of the value for the WT construct.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank all members of the Toro and Pyle laboratories, including, in particular, Vanessa Díaz Prado for technical assistance, Dr. Laura Martínez Rodríguez for providing the image of the secondary structure of RmInt1 intron, and Dr. Marco Marcia for valuable discussions and constructive criticism. We thank Anna Marie Pyle for her help and assistance during the development of this work. This project was supported by research grants BIO2011-24401 from the Spanish Ministerio de Ciencia e Innovación, and CSD 2009-0006 from Programme Consolider-Ingenio, which includes ERDF (European Regional Development Funds). I.C. was supported by a predoctoral fellowship from the Consejo Superior de Investigaciones Científicas and she is currently a postdoctoral research associate of the Howard Hughes Medical Institute (USA). O.F. is a research scientist of the Howard Hughes Medical Institute (USA). M.D.M.-S. is a research associate, F.M.G.-R. is a research associate, F.M.-A. is a research scientist, and N.T. is an investigator at the Consejo Superior de Investigaciones Científicas (Spain).

Received July 21, 2014; accepted September 29, 2014.

REFERENCES

- Arnberg AC, Van Ommen GJ, Grivell LA, Van Bruggen EF, Borst P. 1980. Some yeast mitochondrial RNAs are circular. *Cell* **19**: 313–319.
- Candales MA, Duong A, Hood KS, Li T, Neufeld RA, Sun R, McNeil BA, Wu L, Jarding AM, Zimmerly S. 2012. Database for bacterial group II introns. *Nucleic Acids Res* **40** (Database issue): D187–D190.
- Chanfreau G, Jacquier A. 1993. Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. *EMBO J* **12**: 5173–5180.
- Chillón I, Martínez-Abarca F, Toro N. 2011. Splicing of the *Sinorhizobium meliloti* RmInt1 group II intron provides evidence of retroelement behavior. *Nucleic Acids Res* **39**: 1095–1104.
- Costa M, Fontaine JM, Loiseaux-de Goër S, Michel F. 1997. A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J Mol Biol* **274**: 353–364.
- Costa M, Michel F, Westhof E. 2000. A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J* **19**: 5007–5018.
- Costa M, Michel F, Molina-Sánchez MD, Martínez-Abarca F, Toro N. 2006a. An alternative intron–exon pairing scheme implied by unexpected in vitro activities of group II intron RmInt1 from *Sinorhizobium meliloti*. *Biochimie* **88**: 711–717.
- Costa M, Michel F, Toro N. 2006b. Potential for alternative intron–exon pairings in group II intron RmInt1 from *Sinorhizobium meliloti* and its relatives. *RNA* **12**: 338–341.

- Cousineau B, Lawrence S, Smith D, Belfort M. 2000. Retrotransposition of a bacterial group II intron. *Nature* **404**: 1018–1021.
- Dai L, Zimmerly S. 2002. Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res* **30**: 1091–1102.
- Dai L, Toor N, Olson R, Keeping A, Zimmerly S. 2003. Database for mobile group II introns. *Nucleic Acids Res* **31**: 424–426.
- Dai L, Chai D, Gu SQ, Gabel J, Noskov SY, Blocker FJ, Lambowitz AM, Zimmerly S. 2008. A three-dimensional model of a group II intron RNA and its interaction with the intron-encoded reverse transcriptase. *Mol Cell* **30**: 472–485.
- Daniels DL, Michels WJ Jr, Pyle AM. 1996. Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. *J Mol Biol* **256**: 31–49.
- Fedorova O, Zingler N. 2007. Group II introns: structure, folding and splicing mechanism. *Biol Chem* **388**: 665–678.
- Fica SM, Tuttle N, Novak T, Li NS, Lu J, Koodathingal P, Dai Q, Staley JP, Piccirilli JA. 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* **503**: 229–234.
- Fontaine JM, Rousvoal S, Leblanc C, Kloareg B, Loiseaux-de Goër S. 1995. The mitochondrial LSU rDNA of the brown alga *Pylaiella littoralis* reveals α -proteobacterial features and is split by four group IIB introns with an atypical phylogeny. *J Mol Biol* **251**: 378–389.
- Higuchi R, Krummel B, Saiki RK. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* **16**: 7351–7367.
- Jacquier A, Jacquesson-Breuleux N. 1991. Splice site selection and role of the lariat in a group II intron. *J Mol Biol* **219**: 415–428.
- Jacquier A, Michel F. 1990. Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J Mol Biol* **213**: 437–447.
- Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlman PS. 1988. Group II intron self-splicing. Alternative reaction conditions yield novel products. *J Biol Chem* **263**: 3432–3439.
- Lambowitz AM, Zimmerly S. 2010. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol* **3**: a003616.
- Leontis NB, Westhof E. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA* **7**: 499–512.
- Leontis NB, Stombaugh J, Westhof E. 2002. The non-Watson–Crick base pairs and their associated isosteric matrices. *Nucleic Acids Res* **30**: 3497–3531.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25**: 402–408.
- Marcia M, Pyle AM. 2012. Visualizing group II intron catalysis through the stages of splicing. *Cell* **151**: 497–507.
- Marcia M, Somarowthu S, Pyle AM. 2013. Now on display: a gallery of group II intron structures at different stages of catalysis. *Mob DNA* **4**: 14.
- Martinez-Abarca F, Zekri S, Toro N. 1998. Characterization and splicing in vivo of a *Sinorhizobium meliloti* group II intron associated with particular insertion sequences of the IS630-Tc1/IS3 retroposon superfamily. *Mol Microbiol* **28**: 1295–1306.
- Martinez-Abarca F, Barrientos-Durán A, Fernández-López M, Toro N. 2004. The RmInt1 group II intron has two different retrohoming pathways for mobility using predominantly the nascent lagging strand at DNA replication forks for priming. *Nucleic Acids Res* **32**: 2880–2888.
- Molina-Sánchez MD, Martinez-Abarca F, Toro N. 2006. Excision of the *Sinorhizobium meliloti* group II intron RmInt1 as circles in vivo. *J Biol Chem* **281**: 28737–28744.
- Molina-Sánchez MD, Barrientos-Durán A, Toro N. 2011. Relevance of the branch point adenosine, coordination loop and 3' exon binding site for in vivo excision of the *Sinorhizobium meliloti* group II intron RmInt1. *J Biol Chem* **286**: 21154–21163.
- Muñoz-Adelantado E, San Filippo J, Martínez-Abarca F, García-Rodríguez FM, Lambowitz AM, Toro N. 2003. Mobility of the *Sinorhizobium meliloti* group II intron RmInt1 occurs by reverse splicing into DNA, but requires an unknown reverse transcriptase priming mechanism. *J Mol Biol* **327**: 931–943.
- Murray HL, Mikheeva S, Coljee VW, Turczyk BM, Donahue WF, Bar-Shalom A, Jarrell KA. 2001. Excision of group II introns as circles. *Mol Cell* **8**: 201–211.
- Nagy V, Pirakitikulr N, Zhou KI, Chillón I, Luo J, Pyle AM. 2013. Predicted group II intron lineages E and F comprise catalytically active ribozymes. *RNA* **19**: 1266–1278.
- Nasalean L, Stombaugh J, Zirbel CL, Leontis NB. 2009. RNA 3D structural motifs: definition, identification, annotation, and database searching. In *Non-protein coding RNAs* (ed. Nils W, et al.), Vol. 13, pp. 1–26. Springer-Verlag, Berlin, Heidelberg.
- Pyle AM. 2010. The tertiary structure of group II introns: implications for biological function and evolution. *Crit Rev Biochem Mol Biol* **45**: 215–232.
- Pyle AM, Green JB. 1994. Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry* **33**: 2716–2725.
- Pyle AM, Lambowitz AM. 2006. Group II introns: ribozymes that splice RNA and invade DNA. In *The RNA world: The nature of modern RNA suggests a prebiotic RNA world* (ed. Gesteland RF, et al.), pp. 469–506. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schmidt U, Riederer B, Mörl M, Schmelzer C, Stahl U. 1990. Self-splicing of the mobile group II intron of the filamentous fungus *Podospira anserina* (COI I) in vitro. *EMBO J* **9**: 2289–2298.
- Selbitschka W, Arnold W, Jording D, Kosier B, Toro N, Pühler A. 1995. The insertion sequence element ISRm2011-2 belongs to the IS630-Tc1 family of transposable elements and is abundant in *Rhizobium meliloti*. *Gene* **163**: 59–64.
- Sharp PA. 1991. “Five easy pieces”. *Science* **254**: 663.
- Simon DM, Clarke NA, McNeil BA, Johnson J, Pantuso D, Dai L, Chai D, Zimmerly S. 2008. Group II introns in eubacteria and archaea: ORF-less introns and new varieties. *RNA* **14**: 1704–1713.
- Takaki Y, Matsuki A, Chee GJ, Takami H. 2004. Identification and distribution of new insertion sequences in the genome of the extremely halotolerant and alkaliphilic *Oceanobacillus iheyensis* HTE831. *DNA Res* **11**: 233–245.
- Toor N, Robart AR, Christianson J, Zimmerly S. 2006. Self-splicing of a group IIC intron: 5' exon recognition and alternative 5' splicing events implicate the stem-loop motif of a transcriptional terminator. *Nucleic Acids Res* **34**: 6461–6471.
- Toro N, Martínez-Abarca F, Fernández-López M, Muñoz-Adelantado E. 2003. Diversity of group II introns in the genome of *Sinorhizobium meliloti* strain 1021: splicing and mobility of RmInt1. *Mol Genet Genomics* **268**: 628–636.
- Toro N, Jiménez-Zurdo JI, García-Rodríguez FM. 2007. Bacterial group II introns: not just splicing. *FEMS Microbiol Rev* **31**: 342–358.
- Tourasse NJ, Stabell FB, Reiter L, Kolstø AB. 2005. Unusual group II introns in bacteria of the *Bacillus cereus* group. *J Bacteriol* **187**: 5437–5451.
- van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA. 1986. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* **44**: 225–234.
- Zhong J, Lambowitz AM. 2003. Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription. *EMBO J* **22**: 4555–4565.
- Zimmerly S, Hausner G, Wu X. 2001. Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* **29**: 1238–1250.