Transitions between the steps of forward and reverse splicing of group IIC introns

CLAIRE M. SMATHERS and AARON R. ROBERT

Department of Biochemistry, West Virginia University, Morgantown, West Virginia 20506, USA

ABSTRACT

Group II introns are mobile genetic elements that perform both self-splicing and intron mobility reactions. These ribozymes are comprised of a catalytic RNA core that binds to an intron-encoded protein (IEP) to form a ribonucleoprotein (RNP) complex. Splicing proceeds through two competing reactions: hydrolysis or branching. Group IIC intron ribozymes have a minimal RNA architecture, and splice almost exclusively through hydrolysis in ribozyme reactions. Addition of the IEP allows the splicing reaction to form branched lariat RNPs capable of intron mobility. Here we examine ribozyme splicing, IEP-dependent splicing, and mobility reactions of a group IIC intron from the thermophilic bacterium *Thermoanaerobacter italicus* (*Ta.it.I1*). We show that *Ta.it.I1* is highly active for ribozyme activity, forming linear hydrolytic intron products. Addition of purified IEP switches activity to the canonical lariat forming splicing reaction. We demonstrate that the *Ta.it.I1* group IIC intron coordinates the progression of the forward splicing reaction through a π - π' interaction between intron domains II and VI. We further show that branched splicing is supported in the absence of the IEP when the π - π' interaction is mutated. We also investigated the regulation of the two steps of reverse splicing during intron mobility into DNA substrates. Using a fluorescent mobility assay that simultaneously visualizes all steps of intron integration into DNA, we show that completion of reverse splicing is tightly coupled to cDNA synthesis regardless of mutation of the π - π' interaction.

Keywords: intron; ribozyme; splicing; retrotransposition

INTRODUCTION

Group II introns are self-splicing ribozymes that reside within the genomes of bacteria (Ferat and Michel 1993; Toro 2003) and the organelles of yeast, fungi, algae, and plants (Michel et al. 1989; Fedorova and Zingler 2007; Bonen 2008). Although these RNAs are divergent in primary sequence, they fold into highly conserved structures that catalyze self-excision from precursor transcripts (Qin and Pyle 1998; Toor et al. 2001, 2009; Michel et al. 2009). In addition to their highly structured RNA machinery, many group II introns also contain an open reading frame encoding an intron-encoded protein (IEP) (Michel and Lang 1985; Mohr et al. 1993; Zimmerly et al. 2001). This multidomain protein assists splicing activity, and remains bound after splicing to form intron lariat-IEP ribonucleoproteins (RNPs). The predominant domain of the IEP is a reverse transcriptase (RT), which allows these RNPs to act as retroelements that invade new DNA locations (for reviews, see Lambowitz and Zimmerly 2004; Lambowitz and Belfort 2015; Zimmerly and Semper 2015).

Group II introns are divided into three major classes based on IEP sequence and RNA secondary structural features: IIA, IIB, and IIC (Toor et al. 2001). While all three classes share a general six domain secondary structure radiating from a central wheel (Michel et al. 1989), they have significant structural and functional differences (Fig. 1). Group IIC introns are the smallest and most divergent class of group II introns, missing significant portions of RNA secondary structure compared to IIA and IIB introns (Granlund et al. 2001; Toor et al. 2001). IIC introns also have relaxed sequence requirements for splice site recognition, and utilize a stem-loop structure to decode the 5' splice site instead of the extensive base-pairing interactions observed in other intron classes (Fig. 1; Toor et al. 2006; Robert et al. 2007; Mohr et al. 2018). Normally, splicing initiates by a bulged adenosine nucleophile in domain VI (DVI) attacking the 5' splice site, resulting in the formation of branched lariat product (Peebles et al. 1986; Schmelzer and Schweyen 1986; van der Veen et al. 1986). Curiously, almost all IIC introns splice exclusively through hydrolysis in ribozyme reactions to form linear intron products (Toor et al. 2006, 2008, 2010). Biochemical

Corresponding author: aaron.robart@hsc.wvu.edu

Article is online at <http://www.rnajournal.org/cgi/doi/10.1261/rna.075044.120>. Freely available online through the RNA Open Access option.

© 2020 Smathers and Robert 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. Secondary structure of the *Ta.it.11* group IIC intron. Six RNA secondary structure domains are shown radiating from a central hub. Colors of domains match previously published group II intron crystal structures. Domain IV encodes an ORF that was removed for in vitro ribozyme splicing assays; points of deletion are shown. Greek letters indicate RNA tertiary interactions. Mutation of the putative π tetraloop is shown. Exons are recognized by IBS-EBS elements, γ - γ' , and a terminator stem-loop element.

studies have shown that group IIC introns are dependent upon their IEP to switch from hydrolysis to the canonical branched splicing pathway (Robart et al. 2007; Mohr et al. 2018). This suggests that reduction in RNA secondary structure complexity in IIC introns coincides with greater dependence on the IEP for proper positioning of the DVI bulged adenosine in the ribozyme catalytic core.

Structural studies of group IIB introns have revealed that both protein and RNA tertiary interactions play pivotal roles in positioning DVI between the two steps of splicing. In the pre-second step of reverse splicing (analogous to the conformation expected in the first step of forward splicing) DVI is positioned by RNA-RNA interactions with DIC, and protein-RNA interactions with the IEP (Haack et al. 2019). In this conformation, the bulged adenosine is positioned in close proximity to the ribozyme active site. The transition between the first and second step of forward splicing is mediated by two RNA tetraloop interactions between DII and DVI in IIB introns, namely π - π' and η - η' (Chanfreau and Jacquier 1996; Robart et al. 2014). This DII-DVI interface has also been observed in cryo-EM studies of a group IIA RNP (Qu et al. 2016). Together these studies support a mechanism by which DVI undergoes a large conformational change. Before the first step of splicing DVI is positioned by both RNA-RNA and IEP-RNA interactions. These interactions hold DVI in close association with DIC in an “up” conformation, positioning the bulged adenosine for attack in the ribozyme active site. During the transition to the second step of splicing, DVI swings to the “down” conformation by engaging the DII-DVI π - π' and η - η' tertiary interaction interface. This elegant swinging action removes the first-step splicing branch point from the active site allowing the second step of splicing to proceed. Many group IIA and IIB introns form branched splicing products in the

absence of their IEPs; thus, RNA tertiary interactions alone are sufficient to position DVI. In contrast, group IIC introns lack the η - η' tertiary interaction and strictly require their IEP for these reactions, suggesting that group IIC introns utilize different structural features to coordinate DVI positioning. The overall secondary structure of DII is similar between group IIB and group IIC introns (Toor et al. 2008); however, it is currently unclear if the similar DII secondary structural elements of group IIB and group IIC introns correlate to a common mechanism of splicing.

In this study, we show that a group IIC intron from the thermophilic bacterium *Thermoanaerobacter italicus* (*Ta.it.11*) and its IEP can be readily reconstituted in vitro, and use this system to investigate the transition between the steps of both forward and reverse splicing. We demonstrate that the *Ta.it.11* group IIC intron requires a basic amino acid patch within the IEP to promote the formation of branched splicing products, and uses the π - π' interaction to transition between the steps of the forward splicing reaction. We also find that mutation of the π - π' interaction allows lariat formation independent of the IEP. We developed a fluorescent assay to investigate *Ta.it.11* retro-mobility into DNA. Using this assay, we show that the engagement of the RT domain of the IEP in cDNA synthesis mediates the transition between the steps of the reverse splicing reaction during *Ta.it.11* integration into DNA substrates.

RESULTS

The *Ta.it.11* ribozyme and IEP are highly active in vitro

Mobile group II introns use an intron-encoded protein (IEP) to assist ribozyme splicing and facilitate intron mobility

into DNA targets. To reconstitute IEP-dependent splicing in vitro, an amino-terminal fusion of His₈-SUMO-IEP was expressed in *E. coli* by auto-induction (Studier 2005), followed by affinity purification on Ni-agarose resin (Fig. 2A). The eluted protein was highly active for reverse transcriptase activity in assays with poly(rA)-oligo(dT)₁₈ substrate using PicoGreen fluorescent incorporation to monitor RT activity (Fig. 2B). As a control, we expressed and purified a mutant version of the IEP where the conserved YADD catalytic motif of the RT domain was mutated to YAAA. Mutation of the tandem aspartic acid residues to alanine perturbs binding of catalytic Mg²⁺ metal ions to inhibit RT activity. The mutant IEP's catalytic activity was significantly reduced, but not completely inactive compared to WT (Fig. 2B). The purified IEP retained activity after removal of the His-SUMO purification tag through proteolytic cleavage by Ulp-1 (SUMO protease).

To reconstitute group II intron in vitro self-splicing reactions, *Ta.it.11* intron RNA was prepared by in vitro transcription using T7 RNA polymerase (Wiryaman and Toor 2017). Under standard in vitro transcription conditions, with buffers containing 25 mM MgCl₂, we observed multiple RNA species migrating at lower molecular weights than the expected precursor product (Fig. 2C, lane 9). These findings demonstrate that the *Ta.it.11* intron performs hydrolytic cotranscriptional splicing where a first step water nucleophile produces linear intron. This activity is consistent with previous observations of the ability

of group IIC introns to perform hydrolytic splicing in vitro. The identity of the hydrolytic splicing products was confirmed by repeating the transcription with templates varying the length of the 3' exon (data not shown). Many group II introns have limited activity in vitro, most likely due to RNA misfolding. Thus, the almost complete (88%, Fig. 2C, lane 9) cotranscriptional ribozyme splicing demonstrates that *Ta.it.11* RNA efficiently folds into a highly active ribozyme. Development of a custom transcription protocol was necessary for production of precatalytic RNA. We performed transcription under a range of MgCl₂ concentrations (Fig. 2C). The *Ta.it.11* intron spliced to near completion during transcription under conditions containing transcription buffers formulated with >13 mM Mg²⁺ (Fig. 2C, lane 5–9). In vitro transcription conditions of 5 mM MgCl₂ and 2.5 mM rNTPs were sufficient to prevent cotranscriptional splicing and support activity of T7 RNA polymerase to yield the precatalytic intron population necessary to test self-splicing activity (Fig. 2C, lane 1).

Reconstitution of IEP-dependent splicing

Ta.it.11 is a group IIC intron, which are highly streamlined introns lacking many hallmark RNA secondary structural features observed in other group II intron families (Fig. 1). To investigate if the reduction in RNA secondary structure correlated to a greater dependence on the IEP to assist lariat formation, we compared splicing in the presence or absence of purified IEP. To examine competition between hydrolytic and branched splicing pathways, we performed a titration of Mg²⁺ ions both in the presence and absence of IEP. In the absence of the IEP, the intron behaves similar to other group IIC introns, exclusively utilizing the hydrolytic splicing pathway resulting in linear intron formation (Fig. 3A, even lanes). Hydrolytic splicing activity was dependent on Mg²⁺ concentration, with maximum activity occurring at [Mg²⁺] above 50 mM. Splicing assays performed in the absence of IEP at 5 mM Mg²⁺ showed no ribozyme activity (Fig. 3A, lane 2); however, upon addition of the IEP, 65% of the intron reacted predominantly through the lariat forming pathway (Fig. 3A, lane 3). Thus, formation of an intron-IEP RNP complex switches the splicing reaction from hydrolysis to the canonical lariat forming pathway classically associated with intron splicing reactions. Assays with varying MgCl₂ concentrations demonstrated that lariat forming efficiency is maximal at 10–25 mM, above which the hydrolytic pathway dominates over intron lariat formation. Splicing was further assayed in both the presence and absence of IEP under a range of 100–500 mM NH₄Cl at 55°C (Fig. 3B). At 100 mM NH₄Cl the IEP promoted lariat formation, albeit at lower levels compared to higher salt concentrations (Fig. 3B, lanes 4,7,10). Structural analysis of the *Oceanobacillus iheyensis* (*O.i.*) group IIC intron demonstrated that NH₄Cl

FIGURE 2. Reconstitution of *Ta.it.11* IEP and precatalytic intron RNA. (A) Schematic representation of the *Ta.it.11* IEP purification construct and SDS-PAGE of purified IEP. Mutations of the IEP used in this study are shown. (B) Fluorescent reverse transcription assay of WT and YAAA mutant IEP. (C) 4% PAGE analysis of *Ta.it.11* intron in vitro transcription products under various [MgCl₂].

FIGURE 3. *Ta.it.1* in vitro splicing activity. (A) *Ta.it.1* splicing assay under buffer conditions of 0.5 M NH_4Cl and $[\text{MgCl}_2]$ ranging from 5–100 mM, with and without addition of IEP at 55°C. (B) Comparison of *Ta.it.1* splicing in 100–500 mM NH_4Cl and KCl buffers. (C) *Ta.it.1* intron splicing assay under 0.5 M NH_4Cl and 10 mM MgCl_2 buffer at incubation temperatures from 45°C–75°C. (D) *Ta.it.1* intron splicing assay with WT or YAAA mutant IEP.

and KCl monovalent ions retain the same ribozyme active site architecture (Marcia and Pyle 2012), and it is known that biochemically KCl can promote hydrolytic reactions (Peebles et al. 1987; Jarrell et al. 1988a,b; Daniels et al. 1996). We observed that assays containing KCl promoted a higher level of complete hydrolytic splicing in low $MgCl_2$ conditions compared to NH_4Cl buffers (Fig. 3B, compare lanes 3,6,9 to 13,16,19). The ability of the IEP to stimulate lariat formation over the competing hydrolysis reaction was maintained over temperature ranges from 30°C–55°C when splicing buffer and IEP were preincubated for 5 min before addition of the intron RNA (Fig. 3C). Lariat formation was also supported by the YAAA mutant IEP, demonstrating that a functional RT domain is not required to support group IIC splicing activity (Fig. 3D). Mutation of

the RT catalytic domain has been shown previously to support splicing of other group II intron classes (Moran et al. 1995; Cui et al. 2004; Mastroianni et al. 2008)

IEP and π - π' interactions are required for branched IIC splicing

In an effort to examine if a DVI–IEP interaction is required to promote branched IIC intron splicing, we performed least square fit analysis of a IIB RNP cryo-EM structure to an active IIC IEP crystal structure (Emsley et al. 2010; Stamos et al. 2017; Haack et al. 2019). The superimposed structures predict a basic patch of amino acids in group IIC intron IEPs in close proximity to DVI that may function as a DVI binding face (Fig. 4A). Alignments of primary amino acid sequence showed strong conservation between IIB and IIC introns in this predicted DVI–IEP interaction face (Fig. 4B). To test the role of these IEP positions in lariat formation, basic amino acids were mutated to acidic amino acids. Two of these positions were previously shown by similar charge flip mutations to inhibit IIB intron splicing but still support RNP formation (Fig. 4B, *T.el.* underlined amino acids; Haack et al. 2019). Mutant IEPs with the sequence changed from KKIKTK to either EEIKTK or KKIETE were expressed and purified. The mutant IEPs retained ~50% RT activity (Fig. 4C). To assess the ability of the mutant IEPs to form RNPs, we performed a pull-down assay using the His affinity tag in the IEP against Ni-agarose beads blocked with BSA and a non-complementary DNA oligonucleotide. No *Ta.it.1* RNA binding was observed in the absence of IEP (Fig. 4D, lane 1), and robust RNP formation was observed with WT IEP (Fig. 4D, lane 2). Similar to the observed reduction in RT activity, the mutant IEPs displayed lower affinity for *Ta.it.1* RNA. The pulled-down mutant RNP complexes contained hydrolytic RNA splicing products. This demonstrates that RNA coordinated to the mutant IEPs can no longer be stimulated to perform branched splicing (Fig. 4D, lanes 3–4). Although we cannot exclude the possibility that the RNA undergoes hydrolytic splicing prior to RNP assembly, the absence of branched products supports that the proposed DVI–IEP interaction is required for lariat formation. Similar results were observed in bulk IEP splicing assays. WT RNA spliced with WT IEP promoted lariat formation, with ~4× more lariat produced than lariat-3' exon splicing intermediate (Fig. 4E, lane 3). IEPs containing K to E charge flip mutations were unable to form lariats, resulting in the accumulation of hydrolytic splicing products (Fig. 4E, lanes 4–5).

To further define the IIC branched splicing reaction, we examined the role of the DII-DVI π - π' interaction by mutating the DII π GNRA tetraloop to a UUCG loop that is unable to participate in tetraloop docking interactions (Fig. 1). Mutation of the π tetraloop caused accumulation of the lariat-3' exon splicing intermediate in IEP-dependent splicing

FIGURE 5. *Ta.it.11* in vitro retromobility. (A) Cartoon representation of the fluorescently labeled DNA substrate engineered for fluorescent in vitro retromobility assays. (B) Schematic diagram of reverse splicing and cDNA synthesis. Each expected intermediate and product are numbered in the order they appear in the reaction. The dashed line box insert shows a close-up view of the short intervening sequence containing a polydT stretch engineered to allow limited cDNA synthesis with dATP addition. (C) WT and YAAA mutant *Ta.it.11* RNP retromobility into DNA substrate; assay was performed with or without addition of dATP/dNTPs. Pseudocoloring of product bands on the gel correspond to which fluorophore was detected: Cy5 = red, FAM = green, both = yellow. (D) RNase H digestion of mobility products confirms the presence of extensive cDNA synthesis. (E) Time-course of WT and π mutant *Ta.it.11* RNP retromobility into DNA substrate.

dNTPs, the reactions were digested with RNase H, which degraded the nascent cDNA-intron RNA hybrid (Fig. 5D, lane 5). We further tested the mobility of RT active site mutant (YADD to YAAA) RNP. The RT mutant displayed WT activity for RNP formation and first step reverse splicing activity. However, addition of dNTPs was not capable of stimulating complete reverse splicing due to inhibition of cDNA synthesis (Fig. 5C, lanes 5–6). Together, our data demonstrate that productive engagement of the RT domain of the IEP in cDNA synthesis is closely linked with the second step of *Ta.it.11* reverse splicing.

We demonstrated that the *Ta.it.11* group IIC intron uses a π - π' interaction to position the branch point during the forward intron splicing reaction (see above). We next investigated if the π - π' interaction is also important for the progression of the reverse splicing reaction into DNA substrates. We did not observe any differences in reverse

splicing activity of mutant RNPs compared to WT (Fig. 5E). There were no significant differences in the accumulation of full reverse splicing intron integration after dATP stimulation or cDNA synthesis after dNTP stimulation (Fig. 5E, lanes 6–7 and 12–13). Thus, mutation of the π - π' contact with the branch point creates a stall between the two steps of forward splicing but does not significantly impair reverse splicing or cDNA synthesis during *Ta.it.11* integration into DNA substrates.

DISCUSSION

In this study, we show that the *Ta.it.11* group IIC intron from the thermophilic bacterium *Thermoanaerobacter italicus* is dependent upon its encoded IEP for lariat formation. Competition between the hydrolytic and lariat forming splicing pathways is starkly different between group II intron structural classes. While many IIA and IIB intron ribozymes can readily position the bulged adenosine to promote branched splicing, many IIC introns appear to have lost this RNA mediated coordination interface, and default to hydrolysis (Granlund et al. 2001; Toor et al. 2006, 2010; Mohr et al. 2018). Interestingly, chimeric IIC introns combining the *O.i.* intron with features from one of the few reported group IIC introns capable of catalyzing branched splicing reactions

(A.v.12 from *Azotobacter vinelandii*) could switch splicing preference from hydrolysis to branching (Monachello et al. 2016). In these IIC chimeras, lariat formation was favored over hydrolysis by adjusting the spacing of the bulged adenosine in DVI, restoration of the DII–DVI interface, and substitutions in the DIC stem. Cumulatively, these chimeras reverse-engineered a IIB-like DVI interaction network. Group IIB intron cryo-EM studies have expanded the DVI coordination network to include a combination of DIC–DVI RNA–RNA interactions (ι - ι') and IEP–DVI protein–RNA interactions (Haack et al. 2019). Together, these studies support a model where lariat forming group IIA and IIB ribozymes position the bulged adenosine first step nucleophile through an RNA mediated DIC–DVI interface via the ι - ι' interaction in the absence of IEP (Li et al. 2011; Monachello et al. 2016). Group IIC introns generally lack this interface and cannot position the

bulged adenosine nucleophile in ribozyme reactions alone, causing the hydrolytic pathway to prevail. To compensate for this missing interaction our data, along with other group IIC reconstitution studies, show that the IEP is chiefly responsible for coordination of DVI through IEP–RNA interactions. We show that a patch of basic amino acids within the maturase domain of *Ta.it.11* IEP are required for lariat formation, and propose that these positions form an IEP–DVI binding interface in group IIC introns similar to structural observations in IIB introns.

Structural studies of IIB introns revealed that DII plays an essential role in the transition between the two steps of forward splicing mediated by the π – π' and η – η' RNA tetraloop interactions between DII and DVI (Chanfreau and Jacquier 1996; Robart et al. 2014). Although biochemical studies of chimeric group IIC introns tentatively inferred a π – π' interaction, structural studies used a truncated DII that facilitated crystallization but prevented direct visualization of this interaction (Toor et al. 2008; Costa et al. 2016; Monachello et al. 2016). Here, we show that IIC introns coordinate the lariat bond during the transition between two steps of IEP-dependent splicing through the π – π' interaction (Fig. 6A). Mutation of the DII π GNRA tetraloop stalls the progression of IEP dependent splicing reactions, resulting in the accumulation of lariat–3' exon splicing intermediate. Impaired splicing was also observed in hydrolytic reactions in the absence of the IEP, demonstrating that the π – π' interaction enhances the efficiency of the second step of IIC intron splicing.

FIGURE 6. Transitions between the steps of group IIC intron splicing and mobility. (A) Forward splicing. Pre-first step, hydrolytic reactions are promoted by sequestration of DVI by π – π' . RNP assembly with the IEP dissociates π – π' , allowing DVI coordinating through IEP interactions (IIB intron: PDB 6MEC). Transition between the steps of forward splicing reforms π – π' (PDBs 4R0D, 6ME0, 5G2X). This removes the branchpoint from the active site (gray box) allowing the second step to proceed. (B) Reverse splicing during mobility. The first mobility step accumulates intron lariat–DNA 3' exon product. Stimulation of cDNA synthesis by dNTP addition promotes the second step of reverse splicing through IEP conformational changes. Predicted DVI dynamics are panel A in reverse.

Disruption of the π – π' interaction resulted in low levels of lariat formation in the absence of the IEP. This indicated that the function of the π – π' interaction in IIC introns is two-fold: promoting the progression through the steps of forward splicing, and inhibiting branching in the absence of the IEP. The π – π' interaction anchors DVI to DII; thus, mutation of this interface would be predicted to allow DVI to explore multiple conformations including transient coordination to the active site. Additionally, incubation of the π mutant with WT IEP stimulated the splicing reaction, illustrating that the IEP–DVI interface is required for efficient initiation of the forward branching reaction. The dynamics of DVI between its π – π' and IEP-bound conformations provides a framework to mechanistically understand group IIC RNP assembly. We speculate that in the absence of the IEP, DVI may dock into DII via the π – π' interaction *immediately* after transcription, sequestering the bulged adenosine far from the catalytic site (Fig. 6A). This scenario provides a rationale for why group IIC introns default to hydrolytic reactions in the absence of their IEP. Upon assembly with the IEP, the ribozyme would be activated for lariat formation by undocking π – π' , and handing DVI off to the IEP in a large conformational change. This RNA-to-RNP DVI transition would allow the bulged adenosine access to the active site and trigger the forward branched splicing reaction (Fig. 6A). We attempted to validate this model through FRET assays but were unsuccessful due to loss of activity upon integration of bulky dye adjuncts near the bulged adenosine.

We extended our splicing studies to examine the transition between the two steps of reverse splicing during intron mobility into fluorescently labeled DNA substrates. The *Ta.it.11* IIC intron is a highly active retroelement that rapidly integrates into DNA by completion of the first step of reverse splicing. The π mutant RNA showed an impairment between the two steps of forward splicing; however, we observe no effects of the π mutation in mobility reactions. During the transition between the steps of forward splicing, the active site is reorganized through exchange of catalytic triplex conformations (Chan et al. 2018). Although lariat formation is impaired by the π mutation, ribozymes that successfully complete the forward splicing reaction would be preconfigured in a catalytic triplex arrangement to immediately perform the analogous reverse splicing step.

We, as well as others, observed that in the absence of dNTPs intron mobility accumulates the lariat–3' DNA exon mobility intermediate. Upon addition of dNTPs to the mobility reaction, the mobility intermediate completes the second step of reverse splicing, fully integrating the intron into DNA substrates. Thus, engagement of the RT domain of the IEP is the major mechanistic control point for the transition between the two steps of reverse splicing. This finding mirrors mobility observations for the *LtrA/L.l.* *LtrB* group IIA intron RNP (Aizawa et al. 2003; Mastroianni

et al. 2008). Priming of cDNA synthesis varies between different group II introns. Introns whose IEP contains an endonuclease domain nick the opposite strand of DNA targets to create a primer in the target site, which is used to integrate the inserted intron by cDNA synthesis (Zimmerly et al. 1995). Group IIC introns, as well as the Rmlnt 1 group IIB intron, lack this endonuclease activity, and instead rely upon nascent replication forks to prime cDNA synthesis (Zhong and Lambowitz 2003; Martínez-Abarca et al. 2004; García-Rodríguez et al. 2019). At this time, it is unclear what conformational changes in the RNP facilitate the completion of the second step of reverse splicing; however, it is very likely that priming cDNA synthesis is closely coupled to the 5' DNA exon attack of the lariat bond.

MATERIALS AND METHODS

Plasmids

Ta.it.11 IEP expression plasmid and mutations were synthesized by Genscript. A codon optimized fusion of 8xHis-SUMO-Ta.it. RT was cloned into pET11a by NdeI and BamHI to produce pET11a-His-SUMO-TaitRT. Ta.it.11 in vitro transcription templates were also made by Genscript. Plasmids used in the study are freely available from MolecularCloud.

Expression and purification of Ta.it.11 IEPs

pET11a His-SUMO-Ta.it. RT and its mutants were expressed in Rosetta (DE3) *E. coli* cells. Expression was performed with 1.5 L of Terrific broth auto-induction media (Studier 2005) containing 100 µg/mL carbenicillin and 20 µg/mL chloramphenicol grown at 25°C for 24 h. Cells were harvested at 5000g and resuspended in a final volume of 40 mL of lysis buffer (20 mM Tris-HCl [pH 7.5], 300 mM NaCl, 5 mM β-mercaptoethanol, 10 mM imidazole) and treated with PMSF protease inhibitor at a 1/1000 dilution. The cells were lysed on ice by sonication: 15 bursts of 8 sec separated by 1-min rests. The cell lysate was cleared at 25,000g at 4°C, added to 2 mL Ni-IDA agarose beads (GoldBio) and allowed to bind in a conical tube while turning end-over-end for 30 min at 4°C. The resin was collected at 500g at 4°C and washed with 30 mL of lysis buffer. This process was repeated three times to pre-clean the resin prior to column loading, which was observed to significantly reduce RNase contamination of the IEP preparation. After prewashing, the resin was loaded onto a 20 mL column, washed with 30 mL of wash buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, and 20 mM imidazole), followed by a wash with 2 M NaCl, 20 mM Tris-HCl (pH 7.5), 5 mM β-mercaptoethanol to remove copurifying nucleic acids bound to the IEP. The YAAA mutant IEP was particularly prone to copurification with non-specific nucleic acids judged by OD 260 nm, which may have caused higher background in PicoGreen RT assays (see below). After rinsing the resin with lysis buffer, the protein was eluted with ~15 mL of elution buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM β-mercaptoethanol, and 300 mM imidazole). The IEP was desalted/buffer exchanged in desalting buffer (0.2 M NH₄Cl, 10 mM MgCl₂, 50 mM Tris-HCl pH 7.5, 5 mM β-mercap-

toethanol) on a HiPrep Desalting FPLC column (AKTA). IEP was further purified by size exclusion chromatography using a Sephacryl S-200HR column. The final protein was concentrated at 4000g followed by addition of 5 mM DTT and glycerol to a final concentration of 50%. The IEP was stored at -20°C.

Fluorescent reverse transcriptase (RT) activity assays

Ta.it.11 IEP was diluted 1:5, 1:10, 1:20, 1:50, and 1:100 in dilution buffer (50 mM Tris-HCl pH 7.5, 2 mM DTT, 20% glycerol). An amount 5 µL of each RT dilution was added to 20 µL of poly(rA) oligo(dT)₁₈ substrate and incubated at 25°C for 30 min before quenching with 2 µL of 220 mM EDTA. An amount of 12.5 µL of each reaction was then transferred to one well of a 96 well plate where 86.5 µL of PicoGreen reagent was added. An amount of 12.5 µL of dilution buffer was used as a negative control. Fluorescence was measured with a Synergy H4 Hybrid plate reader (BioTek). To normalize mutant Ta.it. RT activity to WT Ta.it. RT activity, the concentration of each RT was determined by a Bradford assay and each were diluted to 800 µg/mL. RT reactions were set up with 5 µL of 800 µg/mL RT (WT or mutant), to 20 µL of poly(rA) oligo(dT)₁₈ substrate, and 2 µL of RNaseA and performed as described above.

In vitro transcription of Ta.it.11 RNA

RNA transcription was performed with sequence confirmed plasmid DNA and T7 RNA polymerase (Wiryanan and Toor 2017). During transcription optimization, the 10× T7 buffer was replaced with a Ta.it.11 specific 10× T7 buffer containing 5 mM MgCl₂. After transcription and subsequent DNase I and Proteinase K digestion, the transcripts were filtered, concentrated, and buffer exchanged by several rounds of centrifugation at 4000g using an Amicon 100 kDa cutoff ultrafiltration centrifugal filter unit.

In vitro splicing assays

Standard splicing assays were performed with 1 µg of Ta.it.11 WT RNA, 5 mM fresh DTT, 50 µL of splicing buffer (10 mM MgCl₂, 50 mM Tris-HCl [pH 7.5], 0.5 M NH₄Cl final concentration) and made up to 100 µL volume with Millipore water. Splicing reactions were performed by preincubation of the splicing buffer with or without IEP for 5 min at the assay temperature and initiated by addition of the RNA. Hydrolytic splicing assays were performed without IEP under 100 mM MgCl₂ buffer conditions. Maturase splicing assays were performed with ~1.5× molar excess of Ta.it.11 IEP. Reactions were phenol-CIA extracted and ethanol precipitated with 0.3 M NaOAc, pH 5.2 with a linear polyacrylamide carrier. Extracted RNA was resuspended in water and 2× formamide/2 mM EDTA loading buffer before heating at 95°C for 3 min. Samples were resolved on 4% polyacrylamide (19:1 acrylamide: bisacrylamide)/7 M urea gels, stained with ethidium bromide, and imaged on the Syngene G:Box imaging system.

IEP pull-down assay

Ni-IDA-agarose beads (GoldBio) were blocked with 100 ng/µL BSA and 1 µM pMALc5x-Rev oligonucleotide in 10 mM MgCl₂,

250 mM NH_4Cl , 50 mM Tris-HCl (pH 7.5) buffer. A 10 μg Ta.it.11 RNA IEP-assisted splicing reaction in pull-down buffer: 250 mM NH_4Cl , 10 mM MgCl_2 , 1 mM β -mercaptoethanol, 50 mM Tris-HCl (pH 7.5). The reaction was added to the blocked beads and allowed to bind end-over-end at room temperature for 30 min. The resin was collected by centrifugation at 500g for 2 min and washed with 500 μL of pull-down buffer at room temperature for 15 min three times. Pulled-down RNPs were eluted with pull-down buffer supplemented with 300 mM imidazole. Eluted RNPs were phenol-CIA extracted, ethanol precipitated with 0.3 M NaOAc, pH 5.2 with a linear polyacrylamide carrier, resolved on 4% polyacrylamide/7 M urea gel, and stained with ethidium bromide.

Fluorescent oligonucleotide substrate

Mobility assays were performed with multiple different oligonucleotides labeled with fluorescent probes to eliminate the need for radioactivity. DNA substrate was designed with a 5' end DNA stem-loop labeled with Cy5 and a 3' end hairpin labeled with FAM. The sequence of the substrate, synthesized by IDT, was: Cy5GCAGTCTAAAGTAATTTTAGACTGCTTTTTATTTCCTCGCTTCGGCGCGG, underlined position indicates the internal FAM label position.

Fluorescent in vitro mobility assays

RNPs for mobility reactions were assembled by IEP-dependent splicing reactions under standard maturase conditions stated above at 37°C for 15 min. Mobility reactions were performed with 500 nM substrate with or without 1 mM dNTPs or dATP. Controls included were substrate alone and reactions treated with 50 U RNase H. All reactions and controls were incubated at 37°C for 15 min before phenol-CIA extraction and ethanol precipitation with 0.3 M NaOAc (pH 5.2). Reactions at higher temperatures did not support mobility, likely due to melting of the stem-loops in the DNA substrate. Time course assay reactions were stopped using stop solution (0.3 M NaOAc, 20 mM EDTA, 200 μL phenol-CIA). Extracted RNA was resuspended in water and 2 $\times$ formamide/2 mM EDTA loading buffer before heating at 95°C for 3 min. Samples were then resolved on 4% polyacrylamide/7 M urea gels and scanned for fluorescence on a Typhoon imaging system (GE Amersham).

ACKNOWLEDGMENTS

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (NIH) under award number R01GM133857 and start-up funds from West Virginia University.

Received February 3, 2020; accepted February 27, 2020.

REFERENCES

Aizawa Y, Xiang Q, Lambowitz AM, Pyle AM. 2003. The pathway for DNA recognition and RNA integration by a group II intron retrotransposon. *Mol Cell* **11**: 795–805. doi:10.1016/S1097-2765(03)00069-8

Bonen L. 2008. Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* **8**: 26–34. doi:10.1016/j.mito.2007.09.005

Chan RT, Peters JK, Robart AR, Wiryaman T, Rajashankar KR, Toor N. 2018. Structural basis for the second step of group II intron splicing. *Nat Commun* **9**: 4676. doi:10.1038/s41467-018-06678-0

Chanfreau G, Jacquier A. 1996. An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J* **15**: 3466–3476. doi:10.1002/j.1460-2075.1996.tb00713.x

Costa M, Walbott H, Monachello D, Westhof E, Michel F. 2016. Crystal structures of a group II intron lariat primed for reverse splicing. *Science* **354**: aaf9258. doi:10.1126/science.aaf9258

Cui X, Matsuura M, Wang Q, Ma H, Lambowitz AM. 2004. A group II intron-encoded maturase functions preferentially in cis and requires both the reverse transcriptase and X domains to promote RNA splicing. *J Mol Biol* **340**: 211–231. doi:10.1016/j.jmb.2004.05.004

Daniels DL, Michels WJ, 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. doi:10.1006/jmbi.1996.0066

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501. doi:10.1107/S0907444910007493

Fedorova O, Zingler N. 2007. Group II introns: structure, folding and splicing mechanism. *Biol Chem* **388**: 665–678. doi:10.1515/BC.2007.090

Ferat JL, Michel F. 1993. Group II self-splicing introns in bacteria. *Nature* **364**: 358–361. doi:10.1038/364358a0

García-Rodríguez FM, Neira JL, Marcia M, Molina-Sánchez MD, Toro N. 2019. A group II intron-encoded protein interacts with the cellular replicative machinery through the β -sliding clamp. *Nucleic Acids Res* **47**: 7605–7617. doi:10.1093/nar/gkz468

Granlund M, Michel F, Norgren M. 2001. Mutually exclusive distribution of IS1548 and GBSi1, an active group II intron identified in human isolates of group B streptococci. *J Bacteriol* **183**: 2560–2569. doi:10.1128/JB.183.8.2560-2569.2001

Haack DB, Yan X, Zhang C, Hingey J, Lyumkis D, Baker TS, Toor N. 2019. Cryo-EM structures of a group II intron reverse splicing into DNA. *Cell* **178**: 612–623.e12. doi:10.1016/j.cell.2019.06.035

Jarrell KA, Dietrich RC, Perlman PS. 1988a. Group II intron domain 5 facilitates a trans-splicing reaction. *Mol Cell Biol* **8**: 2361–2366. doi:10.1128/MCB.8.6.2361

Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlman PS. 1988b. Group II intron self-splicing. Alternative reaction conditions yield novel products. *J Biol Chem* **263**: 3432–3439.

Lambowitz AM, Belfort M. 2015. Mobile bacterial group II introns at the crux of eukaryotic evolution. *Microbial Spectr* **3**: MDNA3-0050-2014. doi:10.1128/microbiolspec.MDNA3-0050-2014

Lambowitz AM, Zimmerly S. 2004. Mobile group II introns. *Annu Rev Genet* **38**: 1–35. doi:10.1146/annurev.genet.38.072902.091600

Li C-F, Costa M, Michel F. 2011. Linking the branchpoint helix to a newly found receptor allows lariat formation by a group II intron. *EMBO J* **30**: 3040–3051. doi:10.1038/emboj.2011.214

Marcia M, Pyle AM. 2012. Visualizing group II intron catalysis through the stages of splicing. *Cell* **151**: 497–507. doi:10.1016/j.cell.2012.09.033

Martínez-Abarca F, Barrientos-Durán A, Fernández-López M, Toro N. 2004. The Rmlnt1 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. doi:10.1093/nar/gkh616

Mastroianni M, Watanabe K, White TB, Zhuang F, Vernon J, Matsuura M, Wallingford J, Lambowitz AM. 2008. Group II

- intron-based gene targeting reactions in eukaryotes. *PLoS ONE* **3**: e3121. doi:10.1371/journal.pone.0003121
- Matsuura M, Noah JW, Lambowitz AM. 2001. Mechanism of maturase-promoted group II intron splicing. *EMBO J* **20**: 7259–7270. doi:10.1093/emboj/20.24.7259
- Michel F, Lang BF. 1985. Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature* **316**: 641–643. doi:10.1038/316641a0
- Michel F, Umesono K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns—a review. *Gene* **82**: 5–30. doi:10.1016/0378-1119(89)90026-7
- Michel F, Costa M, Westhof E. 2009. The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem Sci* **34**: 189–199. doi:10.1016/j.tibs.2008.12.007
- Mohr G, Perlman PS, Lambowitz AM. 1993. Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. *Nucleic Acids Res* **21**: 4991–4997. doi:10.1093/nar/21.22.4991
- Mohr G, Kang SY-S, Park SK, Qin Y, Grohman J, Yao J, Stamos JL, Lambowitz AM. 2018. A highly proliferative group IIC intron from *Geobacillus stearothermophilus* reveals new features of group II intron mobility and splicing. *J Mol Biol* **430**: 2760–2783. doi:10.1016/j.jmb.2018.06.019
- Monachello D, Michel F, Costa M. 2016. Activating the branch-forming splicing pathway by reengineering the ribozyme component of a natural group II intron. *RNA* **22**: 443–455. doi:10.1261/rna.054643.115
- Moran JV, Zimmerly S, Eskes R, Kennell JC, Lambowitz AM, Butow RA, Perlman PS. 1995. Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol Cell Biol* **15**: 2828–2838. doi:10.1128/MCB.15.5.2828
- Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Cheng HL. 1986. A self-splicing RNA excises an intron lariat. *Cell* **44**: 213–223. doi:10.1016/0092-8674(86)90755-5
- Peebles CL, Benatan EJ, Jarrell KA, Perlman PS. 1987. Group II intron self-splicing: development of alternative reaction conditions and identification of a predicted intermediate. *Cold Spring Harb Symp Quant Biol* **52**: 223–232. doi:10.1101/SQB.1987.052.01.027
- Qin PZ, Pyle AM. 1998. The architectural organization and mechanistic function of group II intron structural elements. *Curr Opin Struct Biol* **8**: 301–308. doi:10.1016/S0959-440X(98)80062-6
- Qu G, Kaushal PS, Wang J, Shigematsu H, Piazza CL, Agrawal RK, Belfort M, Wang H-W. 2016. Structure of a group II intron in complex with its reverse transcriptase. *Nat Struct Mol Biol* **23**: 549–557. doi:10.1038/nsmb.3220
- Robart AR, Seo W, Zimmerly S. 2007. Insertion of group II intron retroelements after intrinsic transcriptional terminators. *Proc Natl Acad Sci* **104**: 6620–6625. doi:10.1073/pnas.0700561104
- Robart AR, Chan RT, Peters JK, Rajashankar KR, Toor N. 2014. Crystal structure of a eukaryotic group II intron lariat. *Nature* **514**: 193–197. doi:10.1038/nature13790
- Saldanha R, Chen B, Wank H, Matsuura M, Edwards J, Lambowitz AM. 1999. RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry* **38**: 9069–9083. doi:10.1021/bi982799l
- Schmelzer C, Schweyen RJ. 1986. Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. *Cell* **46**: 557–565. doi:10.1016/0092-8674(86)90881-0
- Stamos JL, Lentzsch AM, Lambowitz AM. 2017. Structure of a thermostable group II intron reverse transcriptase with template-primer and its functional and evolutionary implications. *Mol Cell* **68**: 926–939.e4. doi:10.1016/j.molcel.2017.10.024
- Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* **41**: 207–234. doi:10.1016/j.pep.2005.01.016
- Toro N. 2003. Bacteria and archaea group II introns: additional mobile genetic elements in the environment. *Environ Microbiol* **5**: 143–151. doi:10.1046/j.1462-2920.2003.00398.x
- Toor N, Hausner G, Zimmerly S. 2001. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7**: 1142–1152. doi:10.1017/S1355838201010251
- 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. doi:10.1093/nar/gkl820
- Toor N, Keating KS, Taylor SD, Pyle AM. 2008. Crystal structure of a self-spliced group II intron. *Science* **320**: 77–82. doi:10.1126/science.1153803
- Toor N, Keating KS, Pyle AM. 2009. Structural insights into RNA splicing. *Curr Opin Struct Biol* **19**: 260–266. doi:10.1016/j.sbi.2009.04.002
- Toor N, Keating KS, Fedorova O, Rajashankar K, Wang J, Pyle AM. 2010. Tertiary architecture of the *Oceanobacillus iheyensis* group II intron. *RNA* **16**: 57–69. doi:10.1261/rna.1844010
- van der Veen R, Amberg 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. doi:10.1016/0092-8674(86)90756-7
- Wiryaman T, Toor N. 2017. Structure determination of group II introns. *Methods* **125**: 10–15. doi:10.1016/j.ymeth.2017.06.020
- 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. doi:10.1093/emboj/cdg433
- Zimmerly S, Semper C. 2015. Evolution of group II introns. *Mob DNA* **6**: 7. doi:10.1186/s13100-015-0037-5
- Zimmerly S, Guo H, Perlman PS, Lambowitz AM. 1995. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**: 545–554. doi:10.1016/0092-8674(95)90027-6
- Zimmerly S, Hausner G, Xc W. 2001. Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* **29**: 1238–1250. doi:10.1093/nar/29.5.1238

RNA

A PUBLICATION OF THE RNA SOCIETY

Transitions between the steps of forward and reverse splicing of group IIC introns

Claire M. Smathers and Aaron R. Robart

RNA 2020 26: 664-673 originally published online March 3, 2020

Access the most recent version at doi:[10.1261/ma.075044.120](https://doi.org/10.1261/ma.075044.120)

References This article cites 51 articles, 13 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/26/5/664.full.html#ref-list-1>

Open Access Freely available online through the *RNA* Open Access option.

Creative Commons License 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/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>