



HAL
open science

A group II intron encodes a functional LAGLIDADG homing endonuclease and self-splices under moderate temperature and ionic conditions

Sahra-Taylor Mullineux, Maria Costa, Gurminder S Bassi, François Michel, Georg Hausner

► **To cite this version:**

Sahra-Taylor Mullineux, Maria Costa, Gurminder S Bassi, François Michel, Georg Hausner. A group II intron encodes a functional LAGLIDADG homing endonuclease and self-splices under moderate temperature and ionic conditions. *RNA*, 2010, 16, pp.1818 - 1831. 10.1261/rna.2184010. hal-03914629

HAL Id: hal-03914629

<https://hal.science/hal-03914629>

Submitted on 28 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A group II intron encodes a functional LAGLIDADG homing endonuclease and self-splices under moderate temperature and ionic conditions

SAHRA-TAYLOR MULLINEUX,^{1,2} MARIA COSTA,² GURMINDER S. BASSI,^{2,3} FRANÇOIS MICHEL,² and GEORG HAUSNER¹

¹Department of Microbiology, University of Manitoba, Winnipeg R3T 2N2, Canada

²Centre de Génétique Moléculaire du CNRS, 91190 Gif-sur-Yvette, France

ABSTRACT

A group II intron encoding a protein belonging to the LAGLIDADG family of homing endonucleases was identified in the mitochondrial *rns* gene of the filamentous fungus *Leptographium truncatum*, and the catalytic activities of both the intron and its encoded protein were characterized. A model of the RNA secondary structure indicates that the intron is a member of the IIB1 subclass and the open reading frame is inserted in ribozyme domain III. In vitro assays carried out with two versions of the intron, one in which the open reading frame was removed and the other in which it was present, demonstrate that both versions of the intron readily self-splice at 37°C and at a concentration of MgCl₂ as low as 6 mM. The open reading frame encodes a functional LAGLIDADG homing endonuclease that cleaves 2 (top strand) and 6 (bottom strand) nucleotides (nt) upstream of the intron insertion site, generating 4 nt 3' OH overhangs. In vitro splicing assays carried out in the absence and presence of the intron-encoded protein indicate that the protein does not enhance intron splicing, and RNA-binding assays show that the protein does not appear to bind to the intron RNA precursor transcript. These findings raise intriguing questions concerning the functional and evolutionary relationships of the two components of this unique composite element.

Keywords: ribozyme; group II intron; LAGLIDADG homing endonuclease

INTRODUCTION

Many group II introns are mobile retroelements that are capable of site-specific insertion into homologous intronless alleles with the assistance of an intron-encoded protein (IEP) (Moran et al. 1995). These mobile group II introns typically contain an open reading frame (ORF) encoding a multifunctional protein with maturase, DNA endonuclease (ENase), and reverse transcriptase (RT) activities in addition to their ribozyme component, which catalyzes splicing (for review, see Lambowitz and Belfort 1993; Saldanha et al. 1993; Michel and Ferat 1995; Lambowitz et al. 1999; Lambowitz and Zimmerly 2004). Ribozyme-catalyzed splicing via the branching pathway leads to the excision of the intron as a lariat, or branched, molecule with a characteristic 2'-5' phosphodiester bond. More

rarely, self-splicing can occur via the hydrolytic pathway, resulting in the excision of the intron as a linear molecule (Daniels et al. 1996; Vogel and Börner 2002). A number of group II introns in yeast, algae, and bacteria have been shown, in fact, to catalyze their own removal from primary transcripts (Peebles et al. 1986; Schmelzer and Schweyen 1986; van der Veen et al. 1986; Schmidt et al. 1990; Ferat and Michel 1993; Costa et al. 1997; Robart and Zimmerly 2005).

The maturase domain of the IEP facilitates efficient intron splicing in vivo by stabilizing the catalytically active structure, and the ENase and RT domains promote intron mobility. A survey of group II introns and their RT ORFs in bacteria and organelles indicated that the two elements appear to coevolve (Toor et al. 2001), although the results of a recent large-scale phylogenetic analysis of group II intron and intron RT sequences suggested that the coevolution of these elements may be imperfect (Simon et al. 2009). During expression of the host gene the intronic ORF is translated and a ribonucleoprotein particle is formed between the intron lariat and the IEP. In a process referred to as retrohoming, the intron and its ORF colonize typically

³Deceased.

Reprint requests to: Georg Hausner, Department of Microbiology, University of Manitoba, Winnipeg R3T 2N2, Canada; e-mail: hausnerg@cc.umanitoba.ca; fax: (204) 474-7603.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.2184010>.

intron-less cognate alleles through a site-specific DNA integration mechanism known as target DNA-primed reverse transcription (Zimmerly et al. 1995a,b). Retrohoming involves both the intron RNA and the IEP and is completed by host DNA repair mechanisms (Schäfer et al. 2003; Lambowitz and Zimmerly 2004).

In contrast to their bacterial counterparts, group II introns in organellar genomes are generally inserted within conserved genes encoding rRNA, tRNA, and components of the electron transport chain (for review, see Saldanha et al. 1993; Michel and Ferat 1995; Lambowitz and Zimmerly 2004; Robart and Zimmerly 2005). The use of insertion sites (IS) that coincide with highly conserved sequence stretches is important for the persistence of these elements in the genome because their elimination would likely require precise deletion so as to not disrupt the host gene.

Also peculiar to organellar genomes was the identification by sequence analysis of several group II introns with ORFs encoding putative LAGLIDADG homing endonucleases (LHEases) in the mitochondrial (mt) small subunit (SSU) and large subunit (LSU) rRNA genes of fungi belonging to the Ascomycota and Basidiomycota (Michel and Ferat 1995; Toor and Zimmerly 2002; Monteiro-Vitorello et al. 2009). However, neither the introns nor the ORFs were characterized functionally in these studies. Genes coding for LHEases are traditionally associated with group I and archaeal introns or with inteins, or they may be present as free-standing ORFs (Dujon 1980; Dalgaard et al. 1993; Dürrenberger and Rochaix 1993; Belfort and Roberts 1997; Jurica and Stoddard 1999; Chevalier et al. 2005; Gibb and Hausner 2005; Stoddard 2006; Sethuraman et al. 2008, 2009; Bae et al. 2009; Singh et al. 2009). LHEase proteins are named for the LAGLIDADG amino acid α -helical motif that comprises each half of the enzyme's active site. LHEases recognize and bind to long (>20 base pairs [bp]) DNA target sites and exhibit flexibility in sequence recognition (Aagaard et al. 1997; Jurica et al. 1998; Lucas et al. 2001; Moure et al. 2003; Eastberg et al. 2007). LHEases promote homing by generating a double-stranded cut with 4 nucleotide (nt) 3' OH overhangs in DNA; the break is repaired by the host's double-stranded break repair processes using the intron/LHEase gene (LHEG)-containing allele as a template (for review, see Belfort and Roberts 1997; Belfort et al. 2002; Stoddard 2006; Edgell 2009). In meiotic crosses this can lead to super-Mendelian inheritance of the LHEG/intron sequence. Some LHEases have been shown to function as maturases and promote the splicing of their host group I intron and occasionally related introns (Lazowska et al. 1989; Ho et al. 1997; Ho and Waring 1999; Bassi et al. 2002; Bassi and Weeks 2003; Belfort 2003; Longo et al. 2005).

During a recent study of the arrangement of the mt SSU (*rns*) gene within *Leptographium* species we uncovered several closely related composite group II introns with putative LHEGs. Comparative sequence analysis and RNA

modeling indicate that the LHEase ORF is situated in domain (D) III, a ribozyme component that acts as a catalytic effector in intron splicing (Lehmann and Schmidt 2003; Fedorova and Zingler 2007). To the best of our knowledge this study constitutes the first biochemical analysis of a group II intron that encodes an HEase of the LAGLIDADG family rather than the RT family. The specific goals of this study were (1) to determine if ORF-less and ORF-containing versions of the intron are self-splicing in vitro; (2) to determine if the LHEase enhances intron-catalyzed splicing in vitro; and (3) to assess if the ORF encodes an active LHEase and to map its cleavage site.

RESULTS

The mt *rns* gene contains a group IIB1 intron with an LHEG

The mt *rns* gene of 47 strains belonging to the fungal genus *Leptographium* was screened for the presence of introns using PCR. Nineteen strains were found to contain insertions between 1.8 and 2.8 kb in size; phylogenetic analyses of these insertions will be described in a future study. Amplification of the mt *rns* gene yielded an amplicon of either 1.2 kb (corresponding to the intron-minus allele) or 3 kb (representing the intron-plus allele). Sequence analysis revealed the presence of a single intron-belonging to the group II intron family inserted at position 952 (based on the numbering of positions in the 16S rRNA gene of *Escherichia coli* strain J01695). The intron IS is conserved in members of *Leptographium* and corresponds to intron 3 (I3) of the mt *rns* gene of *Cryphonectria parasitica* (Fig. 1A). The Lt.SSU/1 intron in *Leptographium truncatum* strain CBS929.85 is 1840 nt in length and encodes an ORF of 915 nt that is inserted after intron position 685. The PCR screening of transformants that were obtained from cloning the amplicon of the mt *rns* gene revealed the presence of both intron-minus (1.2 kb) and intron-plus (3 kb) alleles in *L. truncatum* strain CBS929.85, and both versions of the *rns* gene were sequenced. Heteroplasmy in the mt *rns* gene was detected as well in other strains of *L. truncatum* (Fig. 1A). Consequently, RT-PCR-based approaches or Northern analysis could not be used to determine the intron IS or demonstrate intron splicing in vivo, for it would not be possible to unambiguously distinguish between ligated exons and intron-minus alleles. Therefore, the intron IS was delimited by comparison of intron-plus and intron-minus alleles.

A model of the RNA secondary structure showed that the intron belongs to the IIB1 subclass (Michel et al. 1989) and the ORF is inserted in DIII (Fig. 1B). Exon binding site 1 (EBS1), which defines the 5' splice site by pairing with intron binding site 1 (IBS1), spans a segment of 7 nt. The EBS3-IBS3 base-pairing interaction, involved in 3' splice site recognition, tertiary interactions (α - α' , γ - γ' ,

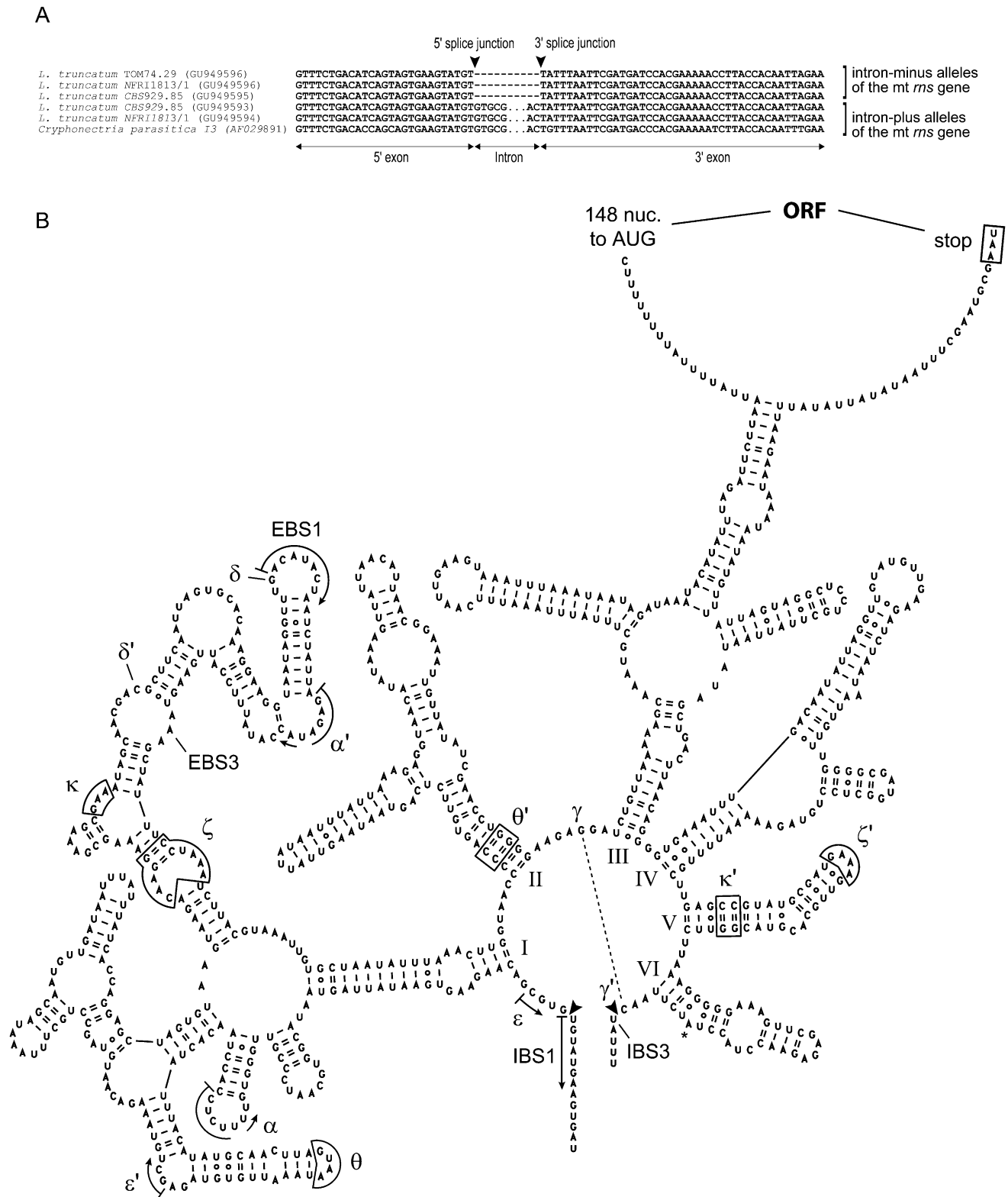


FIGURE 1. (A) Sequence alignment showing the IS of group II introns encoding putative LHEases in the mt *rns* gene of *Leptographium* spp. The intron IS was delimited by comparison of intron-minus and intron-plus alleles of the mt *rns* gene of *C. parasitica*, which contains a related intron (I3), is included. Both the 5' and 3' exon sequences that flank the intron IS and the intron sequences at the 5' and 3' termini, which follow the group II consensus GUGYG and AY, respectively, are shown. The rest of the intervening sequence is represented by dots. The 5' and 3' splice junctions are designated by black arrows. Heteroplasmy (intron-plus and intron-minus alleles) has been observed in *L. truncatum* strains CBS929.85 and NFR11813/1. (B) Secondary structure model of the group IIB1 intron, Lt.SSU/1, of *L. truncatum* showing the insertion of the ORF in a looped-out region in DIII. The 5' and 3' splice sites are marked by black arrowheads and the bulged adenosine residue in DVI is indicated by the asterisk. EBS and IBS nucleotides are shown, as well as nucleotides/segments involved in tertiary interactions (α - α' , γ - γ' , δ - δ' , ϵ - ϵ' , ζ - ζ' , κ - κ' , and θ - θ').

δ - δ' , ε - ε' , ζ - ζ' , κ - κ' , and θ - θ') that are involved in ribozyme folding, and the bulged adenosine residue in DVI (involved in the branch reaction) were identified using previously published structural models (Costa et al. 1997, 1998, 2000; Pyle et al. 2007).

Analysis of the ORF sequence using ORF Finder (National Center for Biotechnology Information) and the ProtParam (Gasteiger et al. 2005) programs identified a putative gene encoding an LHEase of 304 amino acids that contains two LAGLIDADG motifs: ICGLVDAEG and LAGFIEGEA. Analysis using the BLASTP program (National Center for Biotechnology Information) showed that this particular LHEase is related to that encoded within the mS952 group II intron present in the mt *rns* gene of *C. parasitica* (Toor and Zimmerly 2002; Monteiro-Vitorello et al. 2009). The LHEase that was expressed in *E. coli* was 35 kDa, as determined by SDS-PAGE analysis (data not shown). Based on the nomenclature of homing endonucleases proposed by Roberts et al. (2003), the LHEase has been designated I-LtrII, as it represents the second intron-encoded HEase described from *L. truncatum* (Sethuraman et al. 2009).

The Lt.SSU/1 intron self-splices in vitro in the absence of protein factors

The ability of Lt.SSU/1 to self-splice in the absence of protein factors was examined for both ORF-less and ORF-containing versions of the intron RNA. The in vitro-transcribed precursor contained 98 nt of the 5' exon and 262 nt of the 3' exon. The ORF-less intron sequence (Ltr Δ ORF.3) was 819 nt in length, while the full-length, ORF-containing intron (19U/Ltr.2) was 1840 nt. Self-splicing of the in vitro-generated RNA transcripts was tested under low (0.1 M) or high (1 M) concentrations of either NH₄Cl or KCl and 10–100 mM MgCl₂ (Fig. 2). The Ltr Δ ORF.3 readily self-splices during the 20-min reaction period at 37°C and 45°C, and both intron lariat and linear molecules were observed. Self-splicing in the presence of KCl or at an elevated temperature (45°C) resulted in the enhancement of splicing via hydrolysis at the expense of the branching mechanism, although both pathways operate simultaneously under the conditions tested (Fig. 2A). Bands corresponding to additional products are visible in potassium-containing reactions, as well as in some reactions carried out at elevated magnesium concentrations. The former conditions are known to promote hydrolysis at sites that can pair with the EBS1 sequence: we note that sequences AAUAUGU and UGUAUGU, which resemble the IBS1 sequence, are present at intron positions 615 and 688, respectively, of the ORF-less construct.

We also examined self-splicing of the full-length (ORF-plus) intron sequence to determine if the presence of the ORF sequence in DIII affected ribozyme activity; the DIII region is involved in promoting the catalytic efficiency of

the active ribozyme (Lehmann and Schmidt 2003; Fedorova and Zingler 2007). Self-splicing of the ORF-containing precursor RNA was examined within a 15-min reaction period, and the production of lariat and linear intron molecules was observed at 37°C and 42°C, 0.1–1 M of monovalent cation, and 10–50 mM MgCl₂. As was observed with the ORF-less version of the intron RNA, splicing occurs via the branching and hydrolytic pathways (Fig. 2B).

The I-LtrII protein does not bind RNA or enhance the efficiency of intron splicing

Potential maturase activity of the I-LtrII protein was examined for both the N-terminal His₆-tagged form of the purified protein and a near-native version with an additional glycine residue at the C terminus. N-terminal His₆-tagged I-LtrII (at a concentration of 0–80 nM) was added to 20 nM of internally labeled Ltr Δ ORF.3 precursor RNA in a splicing buffer that contained 150 mM KCl and either 6 or 12 mM MgCl₂. The reaction was incubated at 37°C and terminated after 10 and 60 min. Under the conditions tested, the Ltr Δ ORF.3 intron was capable of self-splicing; however, we did not observe an increase in the production of linear and lariat intron RNAs under increasing concentrations of I-LtrII, suggesting that it did not enhance the efficiency of intron splicing (Fig. 3A).

The binding affinity of the I-LtrII protein for the intron RNA precursor was examined using RNA filter-binding assays. Addition of N-terminal His₆-tagged I-LtrII (at a concentration of 0–512 nM) to 5 nM of internally labeled Ltr Δ ORF.3 RNA precursor in the presence of 6 or 12 mM MgCl₂ showed negligible RNA binding (Fig. 3B). The apparent binding of 10% (in the presence of 6 mM MgCl₂) and 14% (in the presence of 12 mM MgCl₂) of Ltr Δ ORF.3 RNA precursor by I-LtrII (at a concentration of 1024 nM) is a reflection of aggregation. The low affinity of the I-LtrII protein for intron RNA indicates that it likely does not have RNA-binding activity.

Since removal of the ORF sequence could have eliminated intron sequences involved in RNA binding by the LHEase, self-splicing of the ORF-containing RNA precursor (19U/Ltr.2) in the presence of near-native I-LtrII was examined over the course of 60 min. Self-splicing of 100 nM of internally labeled precursor transcript in the presence of 6 mM MgCl₂ was compared for samples to which near-native I-LtrII was added (at a concentration of 19.5 nM) or excluded. The results show that addition of near-native I-LtrII does not significantly enhance the formation of ligated exons and release of the full-length (ORF-plus) intron RNA (Fig. 4A). The progress of the reaction was quantitated by determining the ratio of the molar concentration of ligated exons to the sum of the precursor and ligated exons (Fig. 4B). While near-native I-LtrII appears to have a minor effect on the extent and rate of the

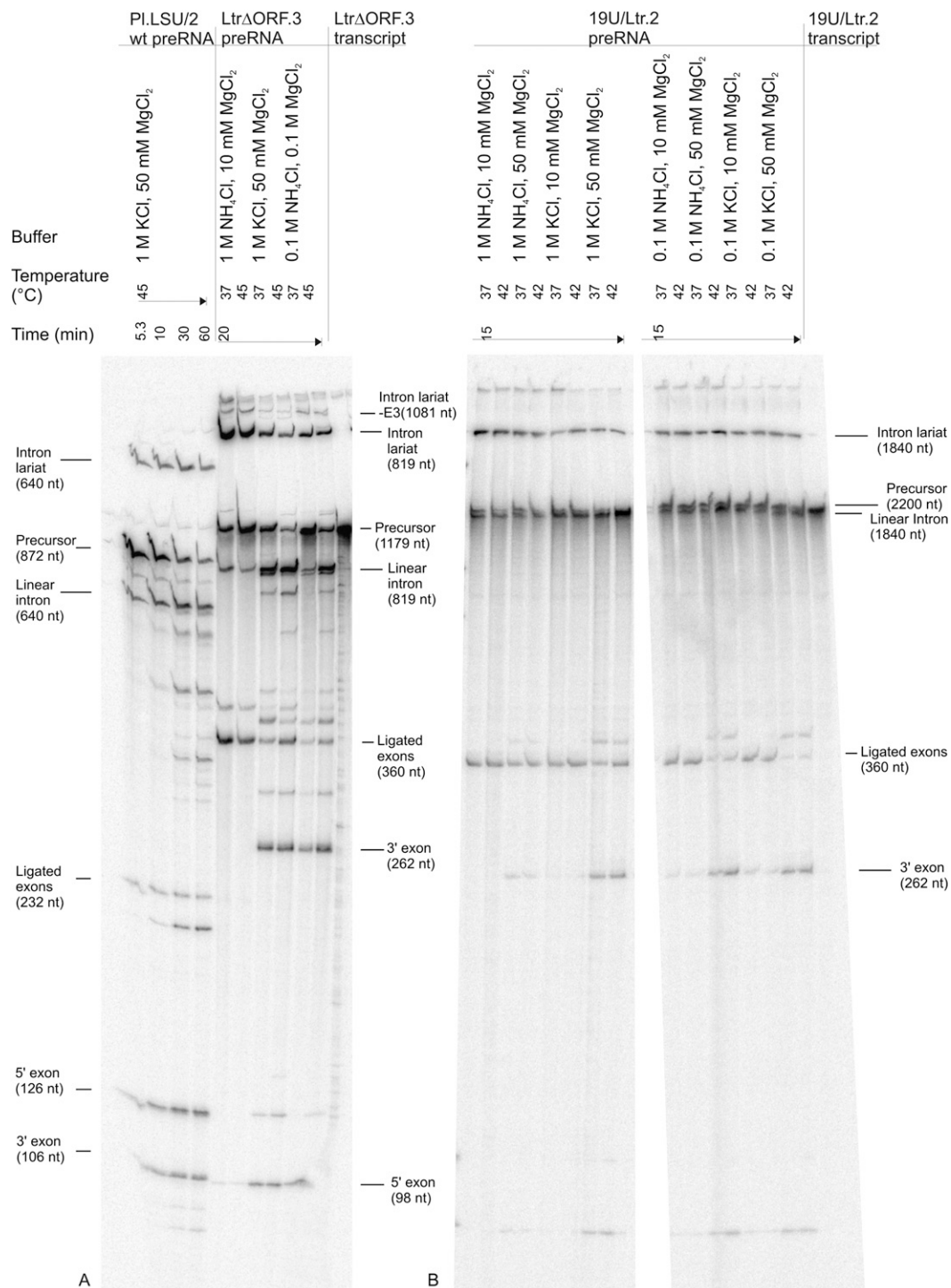


FIGURE 2. Gel electrophoresis of in vitro splicing reactions by Lt.SSU/1 intron under various temperature and ionic conditions. Internally labeled precursor transcripts were incubated for the length of time indicated prior to loading onto a 4% polyacrylamide/8 M urea gel. (A) The Ltr Δ ORF.3 version of the intron, from which the ORF sequence is absent, comprises 819 nt. Products from a self-splicing reaction of the PI.LSU/2 intron from *P. littoralis* (Costa et al. 1997) were used to generate a molecular weight calibration. (B) The 19U/Ltr.2 version of the intron (1840 nt) contains the full-length ORF sequence. “19U/-Ltr.2 transcript” designates unincubated precursor transcript.

formation of ligated exons by the ORF-containing intron precursor, the difference is far too small to be considered significant. Finally, we also attempted to observe maturase

activity at a magnesium concentration (3 mM) that does not support in vitro self-splicing but failed to restore splicing with the addition of the I-LtrII protein (data not shown).

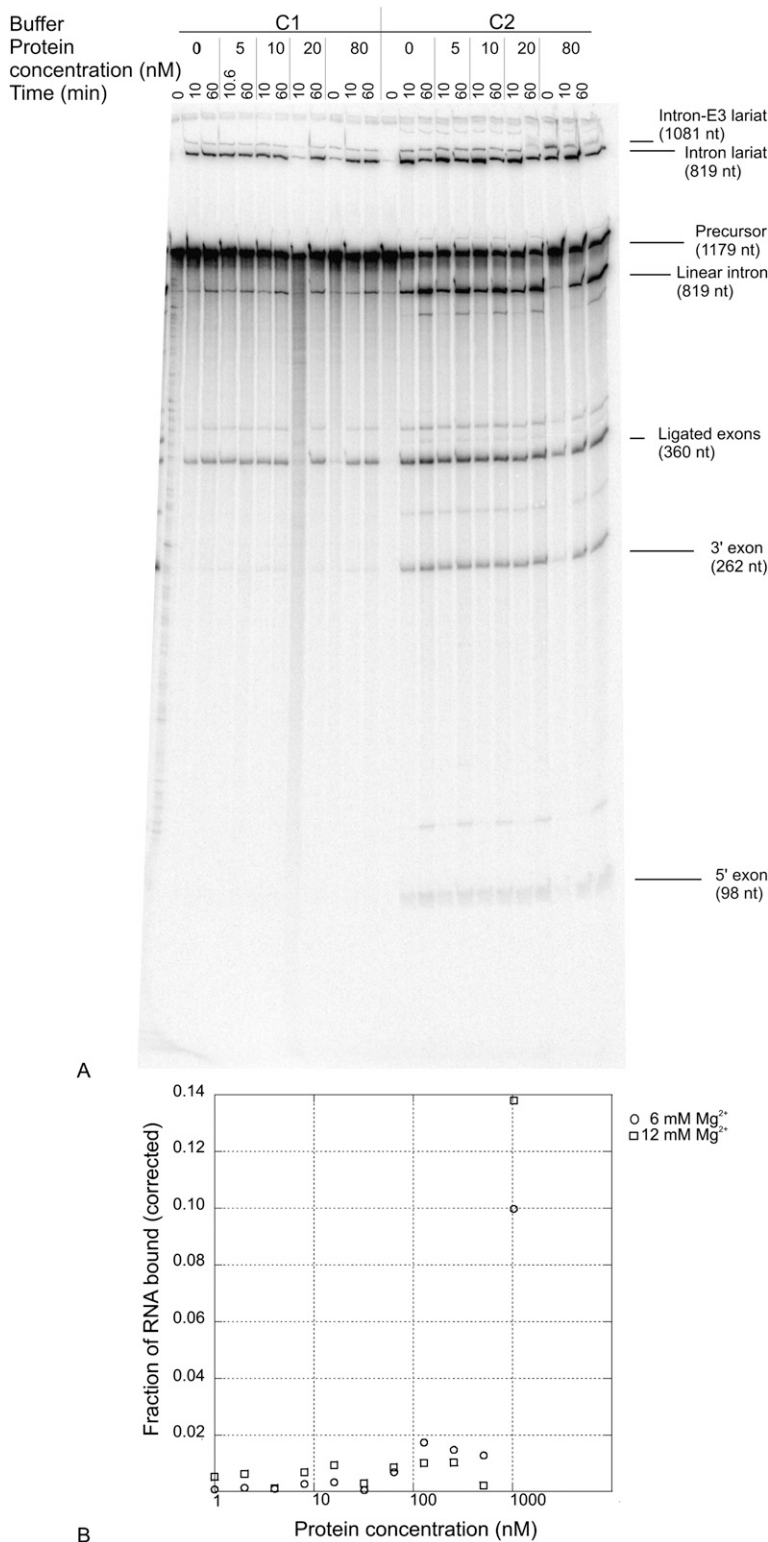


FIGURE 3. (A) Gel electrophoresis of in vitro splicing reactions of the Ltr Δ ORF.3 precursor transcript in the presence of N-terminal His₆-tagged I-LtrII. Internally labeled precursor transcripts were incubated with increasing concentrations of I-LtrII at 37°C for the length of time indicated prior to loading onto a 4% polyacrylamide/8 M urea gel. The composition of buffers C1 and C2 are described in Materials and Methods. (B) RNA-binding assays of N-terminal His₆-tagged I-LtrII. Values were corrected for binding of the eluate to an RNA-binding Hybond N+ filter, and also for the residual binding observed in the absence of protein.

I-LtrII cleaves the mt *rns* gene upstream of the intron IS

Cleavage activity of I-LtrII was characterized using two potential substrates: pCR4mtrnsEx, consisting of the intron-minus allele of the mt *rns* gene (1.2 kb), and pSKmtrnsORFis, which is a synthetic 100-bp fragment composed of intronic sequences that flank the ORF (50 bp were included on each side of the ORF). The cleavage assays were carried out using nonlabeled plasmid DNA that was first linearized with either NcoI (pCR4mtrnsEx) or NotI (pSKmtrnsORFis). Negative controls were NcoI-linearized plasmid DNA containing the intron-plus allele of the mt *rns* gene (pCR4mtrnsExIn) or the mt cytochrome b (*cob*) gene (construct pCR4mtcob); the latter sequence is an unrelated gene cloned into the pCR4-TOPO vector. The substrates and controls were incubated with near-native I-LtrII at 37°C for 45 min, with time points taken every 15 min. The DNA samples were separated by agarose gel electrophoresis and the results show that I-LtrII efficiently cleaves only the linear plasmid containing the exon sequences; that is, I-LtrII cleaves the mt *rns* gene (Fig. 5).

To map the cleavage site, a 248-bp PCR product of the substrate was amplified with the same primers that were used to generate the sequencing ladder. Cleavage products were 5' end-labeled on either the top (sense) or bottom (antisense) strand, and these products were denatured and separated alongside the sequencing ladder. The top strand comigrated with the T at position 950 (based on the numbering of positions in the *E. coli* 16S rRNA gene) in the 5' exon (Fig. 6A), while the bottom strand comigrated with the C at position 946 (Fig. 6B), which correspond to 2 and 6 nt, respectively, upstream of the intron IS (Fig. 6C). I-LtrII binds and cleaves double-stranded DNA within the mt *rns* gene, generating the characteristic 4 nt 3' OH overhangs expected for LAGLIDADG-type HEases.

DISCUSSION

Group II introns with LHEGs

The discovery of novel group II introns that encode ORFs usually associated with group I introns (Michel and Ferat 1995; Toor and Zimmerly 2002) was intriguing; the presence of LHEGs within group II introns could represent the evolution of a new type of composite mobile element or, less interestingly, might simply reflect the existence of yet another genomic niche within which the promiscuous LHEGs may persist as phenotypically silent insertions. In this study, we purified the LAGLIDADG protein encoded by a group II intron and submitted it to biochemical assays in order to determine (1) if it is required to assist the group II ribozyme in splicing; that is, it acts as a maturase, as has been seen for several LHEases associated with group I introns (Ho and Waring 1999; Bassi et al. 2002; Longo et al. 2005); and (2) if it has DNA cleavage activity, and, if that is the case, whether its cleavage site is situated at the proximity of the intron IS, in keeping with the possibility that the intron is actually mobilized by the LHEG.

The *Lt.SSU/1* intron self-splices *in vitro*

One of the primary goals of this study was to functionally characterize the group II intron in the mt *rns* gene of *L. truncatum* strain CBS929.85. Ribozyme activity was examined using two versions of the intron, one in which the ORF sequence had been deleted and the other in which the ORF sequence was present. Several organellar group II introns have been shown to self-splice *in vitro*, including *Sc.cob/1* in the *cob* gene and *Sc.cox1/1* and *Sc.cox1/5 γ* in the gene encoding cytochrome oxidase subunit 1 in *Saccharomyces cerevisiae* (for review, see Michel and Ferat 1995). For these introns, however, optimal ribozyme activity was observed under elevated temperature and ionic conditions; for example, *Sc.cox1/5 γ* self-splices at 45°C in the presence of 0.5 M (NH₄)₂SO₄ and 0.1 M Mg²⁺ (Jarrell

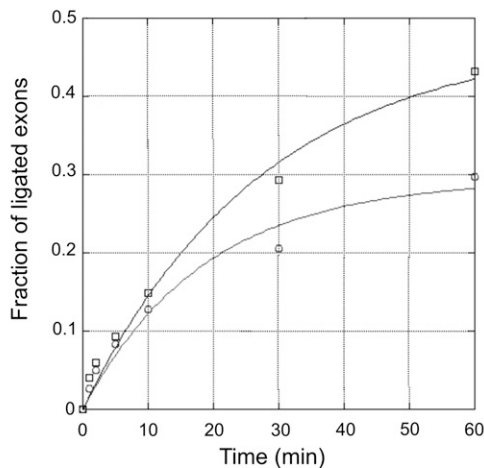
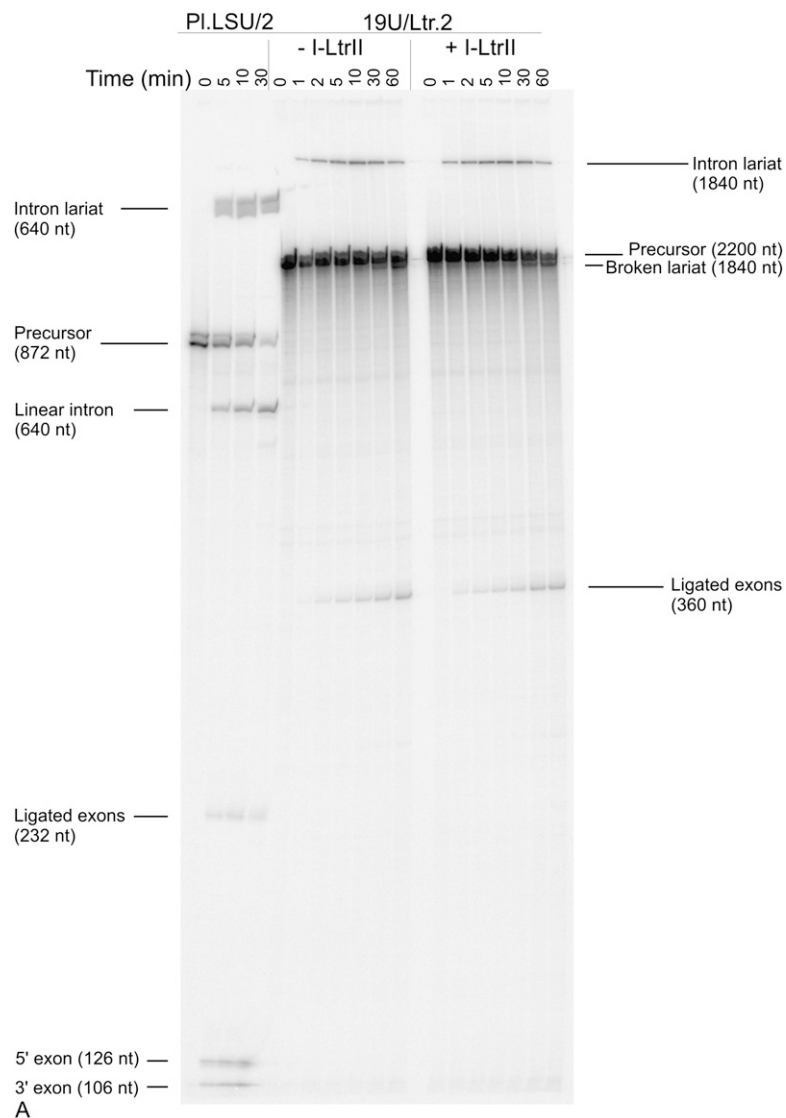


FIGURE 4. (Legend on next page)

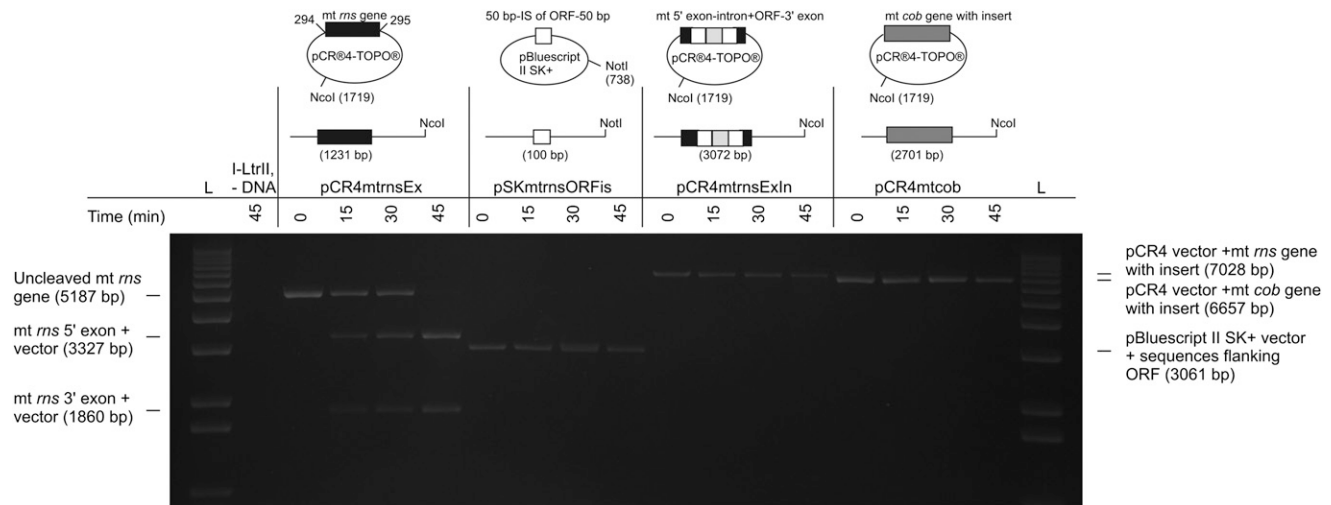


FIGURE 5. Cleavage activity of near-native I-LtrII protein. The ability of I-LtrII to cleave linearized plasmid DNA containing exon sequences (pCR4mtrnsEX) and the intron sequences flanking the LHEase ORF (pSKmtrnsORFis) was tested. The cleavage assays were carried out using nonlabeled plasmid DNA that was first linearized with either NcoI (pCR4mtrnsEx) or NotI (pSKmtrnsORFis). Negative controls were NcoI-linearized plasmid DNA containing the intron-plus allele of the *mt mrs* gene (pCR4mtrnsExIn), in which both potential homing sites are disrupted, or the *mt* cytochrome *b* (*cob*) gene (construct pCR4mtcob), an unrelated gene, that was cloned into the pCR4-TOPO vector. The molecular weight marker, L, used was the 1Kb Plus DNA ladder (Life Technologies).

et al. 1988). In contrast, intron 2 in the LSU rRNA gene of the brown alga *Pylaiella littoralis*, PI.LSU/2, was shown to have an unusually low requirement for Mg^{2+} , as it retains activity at concentrations as low as 0.1 mM (Costa et al. 1997). Both versions of Lt.SSU/1 readily self-splice under moderate temperature (37°C) and ionic (as low as 6 mM Mg^{2+}) conditions, and the presence of the ORF sequence in DIII did not appear to interfere with the proper folding of the group II ribozyme. In most group II introns, the ORF is located in a looped-out region in DIV; this region is dispensable for self-splicing. In contrast, the LHEase ORF has invaded a different helical region, DIII, and genetic and biochemical studies previously indicated that this region enhances splicing efficiency (for review, see Lehmann and Schmidt 2003; Fedorova and Zingler 2007).

Intron-encoded I-LtrII does not enhance intron splicing

Proteins encoded by group I and group II introns have been shown to assist in the self-splicing reactions of the

ribozymes so that the resulting mature RNAs can proceed with their cellular functions either as components of the ribosome or as mRNAs. This “maturase” activity is probably a derived feature that has contributed to the evolutionary success of IEPs by making them essential to the host organism (unless the latter manages to eliminate the entire intron). It may also be interpreted as an advanced stage of symbiosis in which interdependence of the ribozyme and IEP has developed beyond cooperation in mobility. For members of group I it has been suggested that upon the intron being invaded by an HEG, the encoded protein, which initially functions as an ENase to ensure the mobility of the HEG itself, then gains the ability to promote homing of the intron/HEG composite element and that eventually, the LHEase may acquire maturase activity to ensure the long-term survival of the composite intron (Belfort 2003; Gogarten and Hilario 2006). In our study we did not detect evidence for maturase activity in the group II intron-encoded LAGLIDADG protein. N-terminal His₆-tagged I-LtrII failed to bind ORF-less intron RNA precursor transcripts and near-native I-LtrII neither

enhanced splicing of full-length intron RNA precursor transcripts nor rescued it under a prohibitively low (3 mM) Mg^{2+} concentration. This could suggest that the association between the intron and protein results from a relatively recent event or there was little opportunity for selection of maturase activity since, as demonstrated in this study, the group II intron appears to self-splice efficiently in the absence of protein factors.

FIGURE 4. (A) Gel electrophoresis of in vitro splicing reactions with the 19U/Ltr.2 precursor transcript in the presence of near-native I-LtrII. Internally labeled precursor transcripts were incubated with near native I-LtrII at 37°C for the length of time indicated prior to loading onto a 4% polyacrylamide/8 M urea gel. Products from the self-splicing reaction of the PI.LSU/2 intron from *P. littoralis* were used to generate a molecular weight calibration. (B) Quantitation of the progress of the splicing reaction based on the molar ratio of [ligated exons] to ([precursor] + [ligated exons]). The calculation did not take into account the relatively greater vulnerability of the lengthy precursor molecules to degradation over extended reaction times. Curves were fitted to the equation $m1 * [1 - \exp(-m2 * t)]$ using the KaleidaGraph software. Circles ($m1 = 0.294 \pm 0.025$; $m2 = 0.054 \pm 0.012$; Pearson's $R = 0.986$) and squares ($m1 = 0.477 \pm 0.043$; $m2 = 0.036 \pm 0.007$; Pearson's $R = 0.993$) correspond to reactions in the absence and presence, respectively, of the I-LtrII protein.

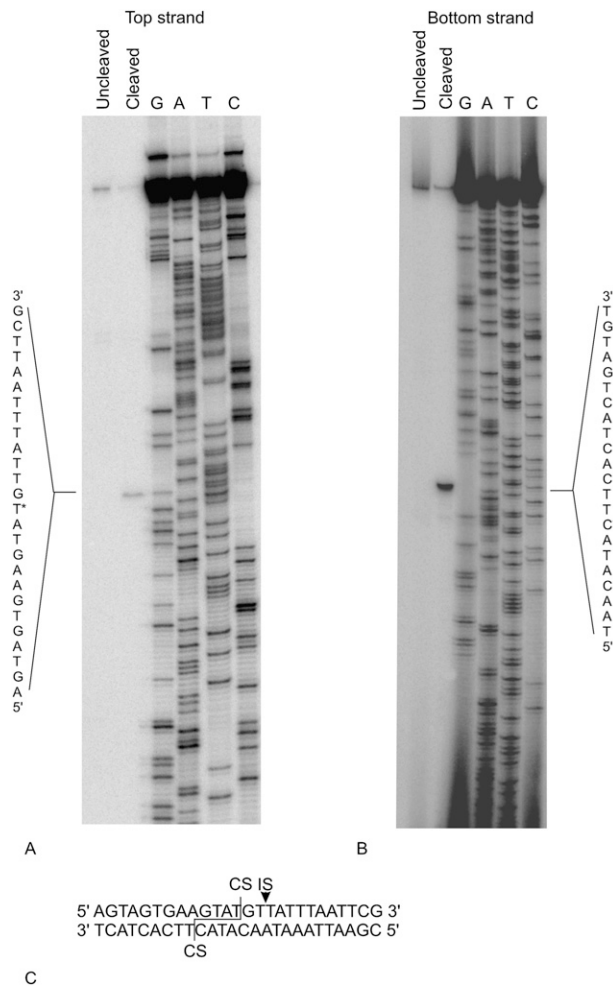


FIGURE 6. Mapping of the cleavage site of the I-LtrII LHEase in the mt *rns* gene. Shown is a representative sequencing ladder generated for the *top* (A) and *bottom* (B) DNA strands that flank the intron IS. The uncleaved product represents the 248-bp PCR amplicon that served as the template for the sequencing reaction. The cleaved product was generated by incubating this amplicon with near-native I-LtrII. The asterisk marks the nt that is immediately 5' to the cut site. Schematic of the region of the mt *rns* gene that flanks the intron IS and includes the cleavage site (CS) of I-LtrII (C).

I-LtrII targets the mt *rns* gene and is capable of promoting intron mobility

Determining the cleavage site of I-LtrII was important as it was supposed to provide a clue as to the possible genetic target(s) of the LHEase, namely, the mt *rns* gene or the intron sequences that flank the ORF. If the target had been the latter, then it would have suggested that while the LHEG had invaded the group II sequence, it probably did not benefit the host intron. In this case, the LHEG likely would act as a truly selfish element that fortuitously found a neutral site to invade and promotes its own mobility. However, not only is the *rns* gene a target, but I-LtrII cleaves specifically the intron-less version of the gene in the immediate vicinity (2 and 6 nt upstream) of the Lt.SSU/1 intron IS,

as observed for other LHEases (Jurica and Stoddard 1999; Sethuraman *et al.* 2009). This leaves little doubt that the protein is capable of promoting the mobility of the group II intron. The LHEG and group II ribozyme may be regarded as having a mutually beneficial relationship in which the intron provides the LHEG with a nondeleterious IS and the LHEG, by virtue of its ENase activity, ensures the mobility of the intron and LHEG. Therefore, this novel arrangement of a LHEG within a group II intron must represent the evolution of a new category of mobile composite element.

Speculations on the origins of the group II ribozyme/LHEG composite element

The generation and persistence of mobile introns are described by the HEG invasion (Belfort 2003) and life cycle (Goddard and Burt 1999; Burt and Koufopanou 2004) hypotheses. HEGs invade DNA sequences encoding self-splicing introns, and since the intron is removed from the mature transcript, it represents a phenotypically silent site for HEG invasion. The HEG spreads in the population through the cleavage activity of the HEase, and, once all possible homing sites have been invaded, the HEG is no longer under purifying selection pressure and begins to accumulate mutations and degenerate. For noncoding intron sequences, there may be little pressure for sequence conservation and the sequences that flank the HEG may begin to accumulate mutations (Haugen *et al.* 2005). However, if the HEG happens to evolve such that the HEase targets a new site, such as intron-less versions of the host gene, the HEG then benefits the intron by mobilizing it and allowing it to spread in the genome or to be transferred horizontally along with the HEG (Loizos *et al.* 1994; Zeng *et al.* 2009). While this theoretical framework was developed specifically to account for mobile, endonuclease-encoding group I and archaeal introns, our demonstration that the I-LtrII protein cleaves next to the intron IS shows that at least some of the group II introns containing LHEGs could also take advantage of the same mechanisms of mobility from which group I and archaeal introns benefit.

Admittedly, few group II intron/LHEG composite elements have been identified, in stark contrast to mobile group I introns, which constitute a significant component of mt and bacteriophage genomes. Moreover, group II introns with LHEGs are so far confined to a single ribozyme structural subgroup (IIB1) and just three sites in mt rRNA genes of fungi (Toor and Zimmerly 2002; Monteiro-Vitorello *et al.* 2009).

One possible argument for the rarity of LHEase-encoding group II introns is that there may be few viable insertion sites for an ORF in group II ribozymes compared with group I ribozymes in which nearly all secondary-structure terminal loops may accommodate the insertion of an ORF (Michel and Westhof 1990). In RT-encoding group II introns, either the entire ORF, or at least the section of it

that specifies the RT, maturase, and ENase domains of the protein, is nearly always located in DIV. However, the LHEG described in this study has invaded the intron in DIII, and while that region enhances catalytic activity (Lehmann and Schmidt 2003; Fedorova and Zingler 2007) the inserted segment did not compromise the intron's potential for self-splicing. This observation can be accounted for by the fact that the ORF is inserted within the DIIIIB terminal loop, a region which tends to be poorly conserved in length and sequence, and the basal, well-conserved segment of DIII is unaffected (Pyle 2010).

For group II introns to evolve from retroelements to elements that move exclusively at the DNA level, they must first be invaded by an HEG. This event is certainly possible in fungal mt genomes, which can harbor large numbers of mobile, HEase-encoding group I introns. However, for group II introns and LHEGs to function effectively as a composite genetic element, the group II intron should be free of an RT-type ORF (and associated maturase activity). In this context, it is an interesting observation that only a handful of subgroup IIB1 introns in mitochondria encode RTs and also that those mt IIB1 ribozymes that were examined proved to self-splice *in vitro* without the assistance of an IEP. In this latter respect, the *Lt.SSU/1* intron, which we have shown to self-splice efficiently in the absence of protein factors under moderate temperature and ionic conditions, is fairly typical of its subclass. It remains now to be established whether or not there exists LHEG-containing group II introns whose ribozyme and protein components cooperate in the splicing process, rather than functioning independently, as appears to be the case for *Lt.SSU/1*.

CONCLUSION

Here we describe a self-splicing group II intron that has been invaded by an LHEG, and together they form a composite element that has the potential to target the mt *rns* gene, thus behaving like group I/HEG composite elements. Therefore, we may be witnessing the evolution of a potentially new combination of mobile element, even though based on currently available data, the LAGLIDADG-type ORFs encoded by group II introns have only a limited distribution, being mostly restricted to mt rDNA of fungi (Toor and Zimmerly 2002; Monteiro-Vitorello et al. 2009). This work demonstrates how novel mobile elements can evolve by shuffling components of unrelated, pre-existing mobile genetic units.

MATERIALS AND METHODS

Amplification, cloning, and sequencing of the mt *rns* gene

The maintenance of fungal cultures and DNA extraction protocols employed in the present study are described by Hausner et al.

(1992). Forty-seven strains belonging to the fungal genus *Lep-tographium* were screened by PCR for the presence of insertions in the mt *rns* gene. The mt *rns* gene was amplified by PCR in a reaction mix containing ($\mu\text{L}/\text{reaction}$): $10\times$ Taq DNA polymerase buffer (Stratagene, Agilent Technologies), containing 15 mM MgCl_2 (5); 50 mM MgCl_2 (0.5); 2.5 mM dNTP (4); 40 μM each forward and reverse primer (0.5 + 0.5); H_2O (38.25); genomic DNA template (1); and Taq polymerase (0.25). The oligonucleotide primers used for amplification were mtsr1 (5'-AG TGGTGTACAGGTGAG-3') and mtsr2 (5'-CGAGTGGTTAGTA CCAATCC-3'). The amplification conditions were initial denaturation (3 min, 93°C); 25 cycles of denaturation (1 min, 93°C), primer annealing (1 min 30 sec, 52.9°C), extension (4 min, 70°C), and a final extension (10 min, 70°C). PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega Corp.). Amplicons were cloned into the pCR4-TOPO vector system using the manufacturer's recommended procedure in the TOPO-TA Cloning Kit for Sequencing (Life Technologies). Plasmid DNA from clones mtsr1435-3 (intron-plus allele) and mtsr1435-16 (intron-minus allele) was harvested and purified using the Wizard *Plus* Minipreps DNA purification system (Promega).

Cycle-sequencing of plasmid DNA was carried out using the Big Dye (version 3.1) protocols supplied by the manufacturer (Applied Biosystems). Sequence data were aligned using GeneDoc V2.7.000 (Nicholas et al. 1997). The program ORF Finder (National Center for Biotechnology Information) was used to identify putative ORF sequences. Parameters of the putative protein, including molecular weight and extinction coefficient, were determined using the ProtParam program available on the ExPASy proteomics server at the Swiss Institute of Bioinformatics (Gasteiger et al. 2005).

Design of precursor constructs and synthesis and purification of transcripts

The *Ltr* Δ ORF.3 precursor was generated in three PCRs from the mtsr1435-3 clone by truncating the 5' exon at position 726 (based on the numbering of positions in the sequence of clone mtsr1435-3) and deleting the ORF sequence. The region between the truncated 5' exon and upstream of the ORF in DIII was amplified using primer pairs Ltr1 (5'-CCCAAACCCTGCAGAAACACCTC AAGAGTAACGTG-3') and Ltr2 (5'-AAAGCTTACGCTTAAAT TACTTCAGAATCAAGTCGTAGC-3'), while the segment between the 3' exon and sequences downstream from the ORF in DIII was amplified using primer pairs Ltr3 (5'-ATTTAAGCG TAAGCTTTAATATATT-3') and mtsr2. The reaction mix contained ($\mu\text{L}/\text{reaction}$): $5\times$ Phusion HF buffer (10); 2.5 mM dNTPs (0.4); 10 μM each forward and reverse primer (5 + 5); H_2O (28.1 μL); template (1); Phusion High-Fidelity DNA polymerase (2 units/ μL , New England Biolabs) (0.5). The amplification conditions were initial denaturation (3 min, 92°C); 15 cycles of denaturation (15 sec, 92°C), annealing (45 sec, 47°C), extension (3 min, 72°C); ramp of 8.0; 25 cycles of denaturation (15 sec, 92°C), annealing (45 sec, 56°C), extension (3 min, 72°C); followed by a final extension (5 min, 72°C). The two segments were joined using self-priming PCR with Ltr1 and mtsr2 primers. The amplification conditions were initial denaturation (3 min, 92°C); 10 cycles of denaturation (15 sec, 92°C), annealing (45 sec, 42°C), extension (3 min, 72°C); ramp of 8.0; five cycles of denaturation

(15 sec, 92°C); annealing (45 sec, 56°C), extension (3 min, 72°C); followed by a final extension (5 min, 72°C). The amplicon was purified with an illustra Sephadex G-50 DNA grade column (GE Healthcare), visualized on agarose gel, and gel extracted using the QIAquick Gel extraction kit (Qiagen). The purified amplicon, 1342 bp in size, was digested with PstI and BmgBI and then ligated with the two fragments of plasmid pTZ19U, which had been triple digested with Ecl136III, AlwNI, and PstI. The ligated product was amplified using the Ant7 primer (5'-AACAGCTATGACCAT GATTACG-3') and 24-mer primer (5'-CGCCAGGGTTTCCCA GTCACGAC-3'). The final construct comprised 98 nt of the 5' exon, 819 nt of the ORF-less intron, and 262 nt of the 3' exon.

The LtrΔORF.3 precursor was transcribed in a reaction mix containing (μL/reaction): 10× transcription buffer (400 mM Tris-HCl [pH 7.5], 120 mM MgCl₂, 100 mM DTT, 20 mM spermidine) (3); rNTPs (25 mM rATP, rGTP, rCTP, and 12.5 mM UTP) (6); 0.5 M DTT (0.6); 1 M MgCl₂ (0.6); α-³²P-UTP (10 μCi/μL) (8); 0.29 μM PCR template (12); T7 RNA polymerase (0.6), and the transcription was run for 2.5 h at 37°C. Three microliters of RQ1 RNase-free DNase (Promega) was then added to the reaction and incubated for 30 min at 37°C, and the reaction was stopped with 3 μL of 500 mM Na₂EDTA (pH 8.0).

19U/Ltr.2 comprises the full-length, ORF-containing, intron sequence (1840 nt), 98 nt of the 5' exon (exon 1), and 262 nt of the 3' exon (exon 2). Transcription of 0.5 μg/μL of plasmid DNA, digested with EcoRI, was carried out in a volume of 50 μL under the same conditions as for LtrΔORF.3, except that the DNase digestion step was omitted. Transcripts were purified as previously described (Michel et al. 1992). Self-splicing assays were carried out as described in Costa et al. (1997).

Expression and purification of N-terminal His₆ tagged I-LtrII

Because of differences in codon use and bias in fungal mt protein-coding sequences, the ORF sequence was optimized for expression in *E. coli* and synthesized commercially (Bio S & T). The optimized ORF sequence was subcloned into the pET200/D-TOPO vector supplied in the Champion pET Directional TOPO expression kit (Life Technologies), following the manufacturer's recommended protocol. The orientation and appropriate reading frame were confirmed by sequencing. For expression, the pET200/Dmtrs1435HEG construct was transformed into *E. coli* strain RosettaBlue (DE3) pLysS (Novagen, Merck KGaA), and 2 mL of an overnight culture was inoculated into 400 mL of 2× YT medium, containing (g/L) tryptone (16), yeast extract (10), NaCl (5), and supplemented with 50 μg/mL of kanamycin. Protein expression was induced at an OD₆₀₀ of 0.8–1.0 with 100 μM of isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3–4 h at 25°C.

The protein was harvested by resuspending the cell pellet in 15 mL of cell lysis (CL) buffer (40 mM HEPES [pH 7.5], 800 mM NaCl, 10% [w/v] glycerol, 6 mM β-mercaptoethanol, and one-half tablet of the complete EDTA-free protease inhibitor cocktail [Roche Diagnostics GmbH]). The N-terminal His₆-tagged protein was purified using Ni-NTA resin (Qiagen S.A.). The sample was washed using CL buffer supplemented with 2–15 mM imidazole. The protein was eluted in buffers containing 40 mM HEPES, 800 mM NaCl, 20% (w/v) glycerol, and either 120 mM or 200 mM imidazole (pH 8). The sample was collected in three fractions (0.75–1 mL) with 120 mM imidazole and then in a final fraction

with 200 mM imidazole (1 mL). Excess imidazole was removed by dialysis in buffer A (40 mM HEPES [pH 7.5], 200 mM NaCl, and 3 mM β-mercaptoethanol) using a slide-a-lyzer dialysis cassette with a 7-kDa molecular weight cutoff (Thermo Fisher Scientific) according to the manufacturer's suggested protocol.

A second purification step was carried out using a HiTrap heparin HP column (GE Healthcare Europe). The sample was washed with one column volume of wash buffer over a range of 200 mM to 1.5 M NaCl. In each wash fraction the NaCl concentration was increased by 100 mM; NaCl was adjusted by mixing the appropriate volumes of buffers A and B (40 mM HEPES [pH 7.5], 1.5 M NaCl, and 3 mM β-mercaptoethanol). The fractions recovered with 600 and 700 mM NaCl were combined and the sample was then concentrated using a Centricon Ultracel YM-30 centrifugation device (Millipore), following the manufacturer's guidelines. The sample was diluted to a final volume of 9 mL in a storage buffer (40 mM HEPES [pH 7.5], 400 mM NaCl, 0.5 mM DTT, and 10% [w/v] glycerol) and centrifuged at 4000g at 4°C until the sample was concentrated in a final volume of 500 μL.

Expression and purification of near-native I-LtrII

The mtrs1435HEG construct was subcloned into the pTYB4 vector (New England Biolabs) for purification using the “intein-mediated purification with an affinity chitin-binding tag” (IMPACT) system (New England Biolabs). For expression, pTYB4-LtrORF was transformed into 50 μL of BL21 Star (DE3) One Shot Chemically competent *E. coli* (Life Technologies). Four hundred milliliters of LB supplemented with 100 μg/mL of ampicillin was inoculated with 1.5 mL of the overnight culture, expression was induced at an OD₆₀₀ of 0.8–1.0 with 250 μM of IPTG, and the culture was incubated for 3 h at 30°C.

The protein was harvested and purified according to the manufacturer's recommended protocol (New England Biolabs). Pellets were resuspended and lysed in buffer CBD 1 (30 mM HEPES [pH 7.5], 1 M NaCl, 0.4 mM EDTA•Na₂•2H₂O [pH 8.0]) to which one-half tablet of the complete EDTA-free protease inhibitor cocktail (Roche Diagnostics) had been added. The sample was washed with 100 mL of CBD2 (30 mM HEPES [pH 7.5] and 700 mM NaCl) and cleaved in CBD2 buffer that was supplemented with 2% (w/v) glycerol and 35 mM DTT; the cleavage reaction was carried out overnight (16–20 h) at 4°C. The protein was eluted in 3 mL of CBD2 supplemented in 4% glycerol and was desalted in an exchange buffer (40 mM HEPES [pH 7.5], 300 mM NaCl, 0.5 mM DTT, and 10% [w/v] glycerol) using an Amicon Ultra 3K NMWL centriprep (Millipore). The I-LtrII protein expressed and purified with this system is near-native, containing an additional glycine residue at its C terminus. The sample was quantitated by measuring the absorbencies at 260 and 280 nm, and the concentration was determined using the formula protein concentration (mg/mL) = 1.55*Abs₂₈₀ - 0.76*Abs₂₆₀.

Maturase and filter-binding assays

Maturase activity of the N-terminal His₆-tagged I-LtrII was assayed using 20 nM of internally labeled LtrΔORF.3 RNA incubated at 37°C with increasing concentrations (up to 80 nM) of protein in splicing buffer C1 (40 mM Na₂ HEPES [pH 7.6], 150 mM KCl, 6 mM MgCl₂) or C2 (40 mM Na₂HEPES [pH 7.6], 150 mM KCl, 12 mM MgCl₂) over a period of 60 min. Maturase activity of

near-native I-LtrII was also tested with the full-length precursor, 19U/Ltr.2. One hundred nanomolars of internally labeled RNA was incubated with 19.5 nM of I-LtrII at 37°C in splicing buffer (40 mM Tris-Cl [pH 7.6 at 37°C], 100 mM NH₄Cl, 6 mM MgCl₂) over a period of 60 min. A molecular weight calibration was generated using the products of the splicing reaction catalyzed by PLLSU/2 (Costa et al. 1997) at 45°C in splicing buffer (40 mM Tris-Cl [pH 7.6 at 37°C], 1 M KCl, 50 mM MgCl₂). Filter-binding assays were carried out with 5 nM of internally labeled LtrΔORF.3 RNA incubated with increasing concentrations of N-terminal His₆-tagged I-LtrII in buffers C1 and C2 at 37°C based on the procedure of Bassi et al. (2002).

Endonuclease assays

To generate the substrates used for the cleavage assays, plasmid DNA, harvested and purified using the Wizard Plus Miniprep DNA purification system (Promega), was linearized with either NcoI (constructs cloned into the pCR4-TOPO vector) or NotI (construct cloned into the pBluescript II SK+ vector). DNA was linearized in a 50 μL reaction volume, containing (μL/reaction): 10× NEBuffer 3 (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, final concentration) (5); plasmid DNA (10); 10 units/μL NcoI (New England Biolabs) (2); H₂O (33). The pBluescript II SK+ vector was linearized as above, except that 15 units/μL of NotI (Life Technologies) was used in REact Buffer 3 (Life Technologies) (50 mM Tris-HCl [pH 8.0], 10 mM MgCl₂, 100 mM NaCl). DNA was purified using the Wizard SV Gel and PCR Clean-Up System (Promega Corp.) and quantified by measuring the absorbance spectrum over 220–330 nm and then using the absorbance maximum in the following formula: concentration (ng/μL) = Abs_{max} * 50 μg/mL. Schematics of the substrates are shown in Figure 5. The potential cleavage substrates were proposed to be either the exon sequences flanking the intron insertion site or the intron sequences in DIII flanking the ORF. Construct pCR4mtrnsEx was used as the intron-minus allele (generated from clone mtrsr1435-16). The second potential substrates comprised 50 bp of the intron sequence upstream of and downstream from the ORF: 5'-TTTTATATTATAAAGGATTTTTCAAGACA TATAAATTAATCAGTTAAATGCGTAAGCTTAATATATTAT ATTTAAGAATAAATAATATGTTTATTAGT-3'. This sequence was synthesized commercially (Bio S & T) and cloned into a pBluescript II SK+ vector to create the construct pSKmtrnsORFis. One control substrate consisted of the intron-plus allele (clone mtrsr1435-3), cloned into the pCR4-TOPO vector as described above, generating the pCR4mtrnsExIn construct; in this construct both potential cleavage substrates are interrupted by the intron and ORF sequences. To assess random cleavage of the vector itself, an unrelated gene, the *cob* gene from *Leptographium lundbergii* strain DAOM60397, which was cloned into the pCR4-TOPO vector, was included as an additional control.

The cleavage reaction mix contained (μL/reaction): Invitrogen's 10× REact Buffer 3 supplemented with 1 mM DTT (2), nuclease-free H₂O (Promega) (14), 35 ng/μL of linearized plasmid DNA (6), and 27 μM I-LtrII (3). Cleavage reactions were incubated at 37°C for the indicated length of time and were stopped by the addition of 3 μL of 200 mM EDTA•Na₂•2H₂O (pH 8.0) and 1 μL of proteinase K (1 mg/mL), with subsequent incubation for 20 min at 37°C. For the zero time sample, I-LtrII was excluded and the sample was incubated for 45 min and the

samples were then stored at –20°C. Samples were separated by 0.8% (w/v) agarose gel electrophoresis.

Mapping of the cleavage site

To map the cleavage site a 248-bp region of the pCR4mtrnsEX construct was PCR amplified and was labeled on either the top or bottom strand. The amplification mixture contained (μL/reaction): 10× ThermoPol reaction buffer (New England Biolabs) (5); 10 mM dNTPs (1); 5 pmol each of γ-³²P-labeled and unlabeled primers (5 + 1.67); 10 ng plasmid DNA (1); *Taq* DNA polymerase (New England Biolabs) (0.5); H₂O (35.83). The amplification was carried out under the following conditions: initial denaturation (2 min, 95°C); 30 cycles of denaturation (45 sec, 93°C), annealing (30 sec, 50°C), extension (30 sec, 70°C); and a final extension (2 min, 70°C). The amplicon was purified using the GenElute PCR clean-up kit (Sigma-Aldrich Chimie S.a.r.l.). Sequencing ladders of the top and bottom strands of the cleavage substrate were generated using the *fmol* DNA sequencing system (Promega). We carried out 5' end-labeling of the primers with γ-³²P-ATP and generation of the sequencing ladders according to the manufacturer's suggested protocol, with the following modification: prior to 5' end-labeling the primer with T4 polynucleotide kinase, the oligonucleotide primer was denatured in water for 2 min at 100°C and then snap-cooled on ice. The cleaved product was generated in a reaction mix containing (μL/reaction): 10× NEBuffer 3 (2); PCR substrate (6); 80.4 μM I-LtrII (1); H₂O (11), and was incubated for 60 min at 37°C. The products of the sequencing reaction were resolved alongside cleaved and uncleaved products on a 5% (w/v) polyacrylamide/8 M urea gel. The salt concentration was normalized for all samples in order to reduce aberrations in migration due to salt differences.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for this research in part through an operating grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Manitoba's University Research Grants Program. Graduate studentships/scholarships to S.-T.M (NSERC Postgraduate Scholarship M award, 2004–2006; University of Manitoba Graduate Fellowship, 2006–2007; Manitoba Health Research Council, 2007–2008) and financial support from the Faculty of Science and Faculty of Graduate Studies (University of Manitoba) are also gratefully acknowledged. We express sincere gratitude to Dr. James Reid (Department of Microbiology, University of Manitoba) for generously supplying strains for this study and to Drs. Jean-Luc Ferat, Mireille Bétermier, Bernard Guiard and Cheng-Fang Li (CGM du CNRS), and William Summers (Department of Microbiology, University of Manitoba) for helpful guidance and discussions. We also gratefully acknowledge the three anonymous reviewers whose insightful comments greatly improved this manuscript.

Received March 18, 2010; accepted June 16, 2010.

REFERENCES

- Aagaard C, Awayez MJ, Garret RA. 1997. Profile of the DNA recognition site of the archaeal homing endonuclease I-Dmol. *Nucleic Acids Res* 25: 1523–1530.

- Bae H, Kim KP, Song JM, Kim JH, Yang JS, Kwon ST. 2009. Characterization of intein homing endonuclease encoded in the DNA polymerase gene of *Thermococcus marinus*. *FEMS Microbiol Lett* **297**: 180–188.
- Bassi GS, Weeks KM. 2003. Kinetic and thermodynamic framework for assembly of the six-component bI3 group I ribonucleoprotein catalyst. *Biochemistry* **42**: 9980–9988.
- Bassi GS, de Oliveira DM, White MF, Weeks KM. 2002. Recruitment of intron-encoded and co-opted proteins in splicing of the bI3 group I intron RNA. *Proc Natl Acad Sci* **99**: 128–133.
- Belfort M. 2003. Two for the price of one: a bifunctional intron-encoded DNA endonuclease-RNA maturase. *Genes Dev* **17**: 2860–2863.
- Belfort M, Roberts RJ. 1997. Homing endonucleases: Keeping the house in order. *Nucleic Acids Res* **25**: 3379–3388.
- Belfort M, Derbyshire V, Parker MM, Cousineau B, Lambowitz AM. 2002. Mobile introns: Pathways and proteins. In *Mobile DNA II*. (ed. NL Craig et al.), pp. 761–783. ASM Press, Washington, DC.
- Burt A, Koufopanou V. 2004. Homing endonuclease genes: The rise and fall and rise again of a selfish element. *Curr Opin Genet Dev* **14**: 609–615.
- Chevalier B, Monnat RJ Jr, Stoddard BL. 2005. The LAGLIDADG homing endonuclease family. In *Homing endonucleases and inteins*. (ed. M Belfort et al.), pp. 33–47. Springer, New York.
- 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, Christian EL, Michel F. 1998. Differential chemical probing of a group II self-splicing intron identifies bases involved in tertiary interactions and supports an alternative secondary structure model of domain V. *RNA* **4**: 1055–1068.
- 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.
- Dalgaard JZ, Garrett RA, Belfort M. 1993. A site-specific endonuclease encoded by a typical archaeal intron. *Proc Natl Acad Sci* **90**: 5414–5417.
- 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.
- Dujon B. 1980. Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at the omega and rib-1 loci. *Cell* **20**: 185–197.
- Dürrenberger F, Rochaix JD. 1993. Characterization of the cleavage site and the recognition sequence of the I-CreI DNA endonuclease encoded by the chloroplast ribosomal intron of *Chlamydomonas reinhardtii*. *Mol Gen Genet* **236**: 409–414.
- Eastberg JH, McConnell Smith A, Zhao L, Ashworth J, Shen BW, Stoddard BL. 2007. Thermodynamics of DNA target site recognition by homing endonucleases. *Nucleic Acids Res* **35**: 7209–7221.
- Edgell DR. 2009. Selfish DNA: Homing endonucleases find a home. *Curr Biol* **19**: R115–R117.
- Fedorova O, Zingler N. 2007. Group II introns: Structure, folding, and splicing mechanism. *Biol Chem* **388**: 665–678.
- Ferat JL, Michel F. 1993. Group II self-splicing introns in bacteria. *Nature* **364**: 358–361.
- Gasteiger E, Hoogland C, Gattiker A, Duvar S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (ed. JM Walker), pp. 571–607. Springer, New York.
- Gibb EA, Hausner G. 2005. Optional mitochondrial introns and evidence for a homing-endonuclease gene in the mtDNA *rnl* gene in *Ophiostoma ulmi* s.lat. *Mycol Res* **109**: 1112–1126.
- Goddard MR, Burt A. 1999. Recurrent invasion and extinction of a selfish gene. *Proc Natl Acad Sci* **96**: 13880–13885.
- Gogarten JP, Hilario E. 2006. Inteins, introns, and homing endonucleases: Recent revelations about the life cycle of parasitic genetic elements. *BMC Evol Biol* **6**: 94.
- Haugen P, Wikmark OG, Vader A, Coucheron DH, Sjøttem E, Johansen SD. 2005. The recent transfer of a homing endonuclease gene. *Nucleic Acids Res* **33**: 2734–2741.
- Hausner G, Reid J, Klassen GR. 1992. Do galeate-ascospore members of the *Cephalosporaceae*, *Endomycetaceae* and *Ophiostomataceae* share a common phylogeny? *Mycologia* **84**: 870–881.
- Ho Y, Waring RB. 1999. The maturase encoded by a group I intron from *Aspergillus nidulans* stabilizes RNA tertiary structure and promotes rapid splicing. *J Mol Biol* **292**: 987–1001.
- Ho Y, Kim SJ, Waring RB. 1997. A protein encoded by a group I intron in *Aspergillus nidulans* directly assists RNA splicing and is a DNA endonuclease. *Proc Natl Acad Sci* **94**: 8994–8999.
- 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.
- Jurica MS, Stoddard BL. 1999. Homing endonucleases: Structure, function, and evolution. *Cell Mol Life Sci* **55**: 1304–1326.
- Jurica MS, Monnat RJ Jr, Stoddard BL. 1998. DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol Cell* **2**: 469–476.
- Lambowitz AM, Belfort M. 1993. Introns as mobile genetic elements. *Annu Rev Biochem* **62**: 587–622.
- Lambowitz AM, Zimmerly S. 2004. Mobile group II introns. *Annu Rev Genet* **38**: 1–35.
- Lambowitz AM, Caprara MG, Zimmerly S, Perlman PS. 1999. Group I and group II ribozymes as RNPs: Clues to the past and guides to the future. In *The RNA world* (ed. RF Gesteland et al.), pp. 451–485. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Lazowska J, Claisse M, Gargouri A, Kotylak Z, Spyridakis A, Slonimski PP. 1989. Protein encoded by the third intron of *cytochrome b* gene in *Saccharomyces cerevisiae* is an mRNA maturase. Analysis of mitochondrial mutants, RNA transcripts, proteins and evolutionary relationships. *J Mol Biol* **205**: 275–289.
- Lehmann K, Schmidt U. 2003. Group II introns: structural and catalytic versatility of large natural ribozymes. *Crit Rev Biochem Mol Biol* **38**: 249–303.
- Loizos N, Tillier ER, Belfort M. 1994. Evolution of mobile group I introns: recognition of intron sequences by an intron-encoded endonuclease. *Proc Natl Acad Sci* **91**: 11983–11987.
- Longo A, Leonard CW, Bassi GS, Berndt D, Krahn JM, Hall TM, Weeks KM. 2005. Evolution from DNA to RNA recognition by the bI3 LAGLIDADG maturase. *Nat Struct Mol Biol* **12**: 779–787.
- Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C. 2001. Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. *Nucleic Acids Res* **29**: 960–969.
- Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Annu Rev Biochem* **64**: 435–461.
- Michel F, Westhof E. 1990. Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* **216**: 585–610.
- Michel F, Umesono K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns—a review. *Gene* **82**: 5–30.
- Michel F, Jaeger L, Westhof E, Kuras R, Tihy F, Xu MQ, Shub DA. 1992. Activation of the catalytic core of a group I intron by a remote 3' splice junction. *Genes Dev* **6**: 1373–1385.
- Monteiro-Vitorello CB, Hausner G, Searles DB, Gibb EA, Fulbright DW, Bertrand H. 2009. The *Cryphonectria parasitica* mitochondrial *rns* gene: Plasmid-like elements, introns, and homing endonucleases. *Fungal Genet Biol* **11**: 837–848.
- 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.

- Moire CM, Gimble FS, Quioco FA. 2003. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J Mol Biol* **334**: 685–695.
- Nicholas KB, Nicholas HB Jr, Deerfield DW II. 1997. GeneDoc: Analysis and visualization of genetic variation. *EMBnet.News* **4**: 1–4.
- 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.
- 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, Fedorova O, Waldsich C. 2007. Folding of group II introns: A model system for large, multidomain RNAs? *Trends Biochem Sci* **32**: 138–145.
- Robart AR, Zimmerly S. 2005. Group II intron retroelements: Function and diversity. *Cytogenet Genome Res* **110**: 589–597.
- Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SKh, Dryden DT, Dybvig K, et al. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases, and their genes. *Nucleic Acids Res* **31**: 1805–1812.
- Saldanha R, Mohr G, Belfort M, Lambowitz AM. 1993. Group I and group II introns. *FASEB J* **7**: 15–24.
- Schäfer B, Gan L, Perlman PS. 2003. Reverse transcriptase and reverse splicing activities encoded by the mobile group II intron cobI1 of fission yeast mitochondrial DNA. *J Mol Biol* **329**: 191–206.
- 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.
- Schmidt U, Riederer B, Mörl M, Schmelzer C, Stahl U. 1990. Self-splicing of the mobile group II intron of the filamentous fungus *Podospora anserina* (CO1 I1) in vitro. *EMBO J* **9**: 2289–2298.
- Sethuraman J, Okoli CV, Majer A, Corkery TLC, Hausner G. 2008. The sporadic occurrence of a group I intron-like element in the mtDNA *rnl* gene of *Ophiostoma novo-ulmi* subsp. *americana*. *Mycol Res* **112**: 564–582.
- Sethuraman J, Majer A, Friedrich NC, Edgell DR, Hausner G. 2009. Genes within genes: Multiple LADLIDADG homing endonucleases target the ribosomal protein S3 gene encoded within an *rnl* group I intron of *Ophiostoma* and related taxa. *Mol Biol Evol* **26**: 2299–2315.
- Simon DM, Kelchner SA, Zimmerly S. 2009. A broadscale phylogenetic analysis of group II intron RNAs and intron-encoded reverse transcriptases. *Mol Biol Evol* **26**: 2795–2808.
- Singh P, Tripathi P, Silva GA, Pingoud A, Muniyappa K. 2009. Characterization of *Mycobacterium leprae* RecA intein, a LAGLIDADG homing endonuclease, reveals a unique mode of DNA binding, helical distortion, and cleavage compared with a canonical LAGLIDADG homing endonuclease. *J Biol Chem* **284**: 25912–25928.
- Stoddard BL. 2006. Homing endonuclease structure and function. *Q Rev Biophys* **38**: 49–95.
- Toor N, Zimmerly S. 2002. Identification of a family of group II introns encoding LAGLIDADG ORFs typical of group I introns. *RNA* **8**: 1373–1377.
- Toor N, Hausner G, Zimmerly S. 2001. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7**: 1141–1152.
- 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.
- Vogel J, Börner T. 2002. Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J* **21**: 3794–3803.
- Zeng Q, Bonocora RP, Shub DA. 2009. A free-standing homing endonuclease targets an intron insertion site in the *psbA* gene of cyanophages. *Curr Biol* **19**: 218–222.
- Zimmerly S, Guo H, Perlman PS, Lambowitz AM. 1995a. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**: 545–554.
- Zimmerly S, Guo H, Eskes R, Yang J, Perlman PS, Lambowitz AM. 1995b. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83**: 529–538.



RNA

A PUBLICATION OF THE RNA SOCIETY

A group II intron encodes a functional LAGLIDADG homing endonuclease and self-splices under moderate temperature and ionic conditions

Sahra-Taylor Mullineux, Maria Costa, Gurminder S. Bassi, et al.

RNA 2010 16: 1818-1831 originally published online July 23, 2010
Access the most recent version at doi:[10.1261/rna.2184010](https://doi.org/10.1261/rna.2184010)

References This article cites 64 articles, 14 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/16/9/1818.full.html#ref-list-1>

License

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](#).

Dharmacon[™] Reagents
Custom synthesis, RNAi, and CRISPR solutions

Infinite Reliability
for Functional Genomics [More](#)

horizon
a PerkinElmer company

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