Published online 30 June 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 18* ***8065–8077***
doi:10.1093/nar/gkr478

Identification of an hepatitis delta virus-like ribozyme at the mRNA 5′-end of the L1Tc retrotransposon from *Trypanosoma cruzi*

Francisco J. Sánchez-Luque[1], Manuel C. López[1,*], Francisco Macias[1], Carlos Alonso[2] and M. Carmen Thomas[1,*]

[1]Departamento de Biología Molecular, Instituto de Parasitología y Biomedicina López Neyra—CSIC, Parque Tecnológico de Ciencias de la Salud, Granada and [2]Centro de Biología Molecular Severo Ochoa—CSIC, Madrid, Spain

Received January 21, 2011; Revised and Accepted May 24, 2011

ABSTRACT

L1Tc is a non-LTR LINE element from *Trypanosoma cruzi* that encodes its transposition machinery and bears an internal promoter. Herewith, we report the identification of an *in vitro* active hepatitis delta virus-like ribozyme located in the first 77 nt at the 5′-end of the L1Tc mRNA (L1TcRz). The data presented show that L1TcRz has a co-transcriptional function. Using gel-purified uncleaved RNA transcripts, the data presented indicate that the kinetics of the self-cleaving, in a magnesium-dependent reaction, fits to a two-phase decay curve. The cleavage point identified by primer extension takes place at +1 position of the element. The hydroxyl nature of the 5′-end of the 3′-fragment generated by the cleavage activity of L1TcRz was confirmed. Since we have previously described that the 77-nt long fragment located at the 5′-end of L1Tc has promoter activity, the existence of a ribozyme in L1Tc makes this element to be the first described non-LTR retroelement that has an internal promoter–ribozyme dual function. The L1Tc nucleotides located downstream of the ribozyme catalytic motif appear to inhibit its activity. This inhibition may be influenced by the existence of a specific L1Tc RNA conformation that is recognized by RNase P.

INTRODUCTION

Retrotransposons are mobile DNA elements that transpose through an RNA intermediate, which is reverse transcribed and integrated into a new position in the genome. These elements are classified into two major groups: those that are flanked by long terminal repeats (LTR) or LTR retrotransposons and those that lack LTR named non-LTR retrotransposons. Two groups of elements lacking LTR have been described: long interspersed nucleotide elements (LINEs, also known as L1) with coding capacity and short interspersed nucleotide elements (SINEs) without coding capacity. LINEs and SINEs bear a poly-A tail and are flanked by direct target site duplication (TSD) sequences. Some of these elements exhibit site specificity for insertion, while others present a random distribution. Since transcription is the first step in the element mobilization process, some non-LTR elements (fly, humans, mouse or *Trypanosoma cruzi* L1) bear an internal promoter to preserve its autonomous character (1–5). Furthermore, LINEs encode the proteins implicated in their mobilization mechanism, being able to mobilize SINEs in *trans*. The mobilization mechanism has been termed target-primed reverse transcription (TPRT) (6).

The L1Tc element is the most well studied non-LTR retrotransposon of the genome of *T. cruzi*, a protozoan parasite responsible for Chagas' disease, a chronic sickness that affects 10 million people in South America (http://www.who.int/mediacentre/factsheets/fs340/en/index.html). L1Tc is actively transcribed as a polyadenylated mRNA (7). It codes for the enzyme machinery involved in TPRT, including an apurinic/apyrimidinic (AP) endonuclease (7,8), 3′-phosphatase, 3′-phosphodiesterase, a reverse transcriptase (9), RNase H (10) and a nucleic acid chaperone (11,12). L1Tc also bears at its N-terminal end a functional 2A-autoproteolitic sequence (13) similar to that found in small-size RNA viruses. The first 77 nt of L1Tc (Pr77) correspond to an internal promoter that generates abundant and translatable transcripts (5). Run-on analyses employing RNA polymerase-specific inhibitors suggest that L1Tc is

*To whom correspondence should be addressed. Tel: +34 958 181 661; Fax: +34 958 181 632; Email: mclopez@ipb.csic.es
Correspondence may also be addressed to M. Carmen Thomas. Tel: +34 958 181 662; Fax: +34 958 181 632; Email: mcthomas@ipb.csic.es

transcribed by RNA polymerase II. Pr77-derived transcripts are not processed by a *trans*-splicing mechanism and initiate at nucleotide +1 of L1Tc (5). In spite of the large and diverse chromosome distribution of L1Tc in several strains of the parasite (14), there is evidence suggesting that L1Tc may show certain insertion site specificity (15). Pr77 is not restricted to L1Tc from *T. cruzi* as it is also present at the NARTc non-autonomous retrotransposon of the *T. cruzi* genome (16), at the *ingi* and RIME non-LTR retrotransposons of the *T. brucei* genome (17), at short interspersed degenerate retrotransposon in the genomes of *Leishmania* species (18) and also it has been found associated with sequences not related to retroelements at different positions of the *T. brucei* genome (15).

Recently, the presence of an active hepatitis delta virus (HDV)-like ribozyme at the 5′-untranslated region (5′-UTR) of the *Drosophila simulans* R2 element has been described (19). The R2 non-LTR retrotransposon copies from *Drosophila* are specifically integrated into the same position of the 28S ribosomal genes and are co-transcribed with the rRNA. The ribozyme releases the R2 mRNA from the 28S-R2 co-transcript *in vitro* leaving a 5′-end similar to that detected in R2 transcripts *in vivo*. In contrast to L1Tc, the R2 element does not have an internal promoter.

The HDV ribozyme belongs to the group of small autocatalytic RNAs whose members are smaller than 200 nt in length and catalyze a *trans*-esterification that leads to the cleavage of the RNA sugar–phosphate backbone leaving 5′-hydroxyl and 2′,3′-cyclic phosphate ends (20,21). The HDV ribozyme was first described to reside within the HDV circular RNA, where it releases the genomic RNA units from the large concatemers generated by rolling circle replication (22). This ribozyme cleaves at a site located immediately upstream of the minimal catalytic domain that exhibits a compact tertiary structure. New members of the HDV-like ribozyme family have been described to be present in the human genome (23) as well as in other organisms, including insects, plants and fish, where they have shown to be functional (24).

In the present article, we report the identification of an active HDV-like ribozyme located in the first 77 conserved nucleotides of the 5′-end of the mRNA from the L1Tc retrotransposon (L1TcRz). The sequence of the minimal catalytic domain endowed with activity is compatible with an HDV-like ribozyme folding. We have analyzed the cleavage kinetics at different Mg^{2+} concentrations and identified the cleavage site. We also analyzed the nature of the 5′-end generated by the activity of L1TcRz and described the influence of the upstream sequences in the optimal folding required for catalysis. The results are consistent with the idea that L1TcRz belongs to the HDV-like ribozyme type. The L1TcRz is the first HDV-like ribozyme reported in Trypanosomatids, the second reported in a non-LTR retroelement and the fourth characterized in depth. Since we have previously described that Pr77 has promoter activity, the existence of a ribozyme in this region of L1Tc makes this element to be the first described non-LTR retroelement with an internal promoter–ribozyme dual function. It was observed that the region of L1Tc located downstream of L1TcRz negatively affects the ribozyme activity. This region leads the 5′-UTR to adopt a structure that sequesters the L1TcRz into a non-catalytic conformation. Pr77-derived transcripts are translated despite the fact that they are not processed by *trans*-splicing (5), suggesting the existence of a cap-independent translation mechanism similar to the internal ribosome entry site (IRES) described for mouse L1 (25). Since IRESs are generally sensitive to RNase P cleavage *in vitro* (26,27), the data of the RNase P-mediated cleavage presented in this article support the hypothesis that an IRES may be present in L1Tc.

MATERIALS AND METHODS

Construction of DNA templates for transcription

DNA templates were generated for *in vitro* transcription by PCR employing two different templates: L1Tc genomic clone 7134 and L1Tc cDNA clone 55 (7) (accession number X83098). Both clones differ in the composition of the sequence located upstream of the +1 position of the element. Constructs were also generated bearing the pGEM-T easy sequence and the L1Tc sequence of different lengths starting at its +1 position.

The general scheme of PCRs consisted of a single 5′-primer that incorporates the T7 RNA polymerase promoter and anneals several nucleotides upstream of the +1 position of the element or at the pGEMT-easy vector, where L1Tc 1–152 fragment is cloned and different 3′-primers are annealed at different positions within L1Tc sequence. The 5′-primers are common for all constructs of each clone, while 3′-primers are common for the three series and unique for each length product.

PCRs were performed by ReddyMix Kit (Thermo Fisher Scientific—ABgene) and agarose gel purified by phenolic extraction. The 5′-primers were: T7 c55 −100 (5′-TAATACGACTCACTATAGGGCGCTGTA CTA-3′) annealing 100 nt upstream the +1 nt of clone 55; T7-5′-UTR-G3PDHf (5′-TAATACGACTCACTAT AGGGATATTTTTACTTTGAAAGCCA-3′) annealing 171 nt upstream the +1 nt of clone 7134; and M13 universal forward primer (5′-GTTTCCCAGTCACGAC-3′). The 3′-primers were: L1Tc55+47r (5′-GAGTACTAGAC CCTGGCACCA-3′), L1Tc55+59r (5′-CTCTCTAGCAA AGAGTACTAGACCCT-3′), L1Tc55+70r (5′-CGCTTA GCTTCCTCTCTAGC-3′), L1Tc55+77r (5′-CAGCAGG CGCTTAGCTTCCTCT-3′), L1Tc55+126r (5′-CCGACC CGTTTGTGCGGCG-3′) and L1Tc55+152r (5′-TGTAA ATGGCTCCATCT-3′).

Additional substrates were used for cleavage reactions resolved in native gels. In this case, clone 7134 was chosen for these experiments and uncleaved RNAs were designed to contain 10 nt upstream the cleavage position. DNA templates were generated by PCR using T7 L1Tc7134-10f (5′-TAATACGACTCACTATAGGGACATCCTCA GCCCTG-3′) primer in 5′ and L1Tc55+77 r, L1Tc55+ 126r and L1Tc+152r in 3′. NARTc RNAs synthesis and co-transcriptional cleavage assays were performed using DNA templates generated by PCR using pBAC52-EcoR1 digested band 3 (14) as DNA template and the primers T7 NARTc-11f (5′-TAATACGACTCACTATA

GGGTATCTTTGGCCCCTG-3′) in 5′ and L1Tc55+77r, NARTc+126r (5′-AAAACTAAGTAACAACTACTCATC-3′) and NARTc+152r (5′-CTCCAACATCTGCCCTTCC-3′) in 3′.

Co-transcriptional cleavage assays

About 22 ng of PCR templates were transcribed by T7 RNA polymerase kit (PROMEGA) following manufacturer instructions with the following exceptions: reactions were scaled down to 10 µl final volume, 0.5–1 µCi of α^{32}P-UTP was added to each reaction, and UTP final concentration was reduced to 0.8 mM. Reactions were carried out for 2 h and performed at three different temperatures: 42, 37 and 25°C. About 10 µl of formamide buffer 2× were added to each reaction as stop and loading buffer. Samples were resolved in 8% polyacrylamide, 7 M urea and tris–borate–EDTA (TBE) 1× gel electrophoresis.

Cleavage reactions

About 45 ng of the DNA templates were transcribed using the T7 RNA polymerase kit (PROMEGA) following manufacturer instructions in a 100 µl final volume reaction and co-transcriptionally radiolabeled by reducing UTP to 0.4 mM final concentration and adding 40 µCi of α^{32}P-UTP to the reaction. Transcriptions were performed at 37°C for 2 h and stopped by addition of 100 µl 2× formamide buffer. Uncleaved fragments were purified from 8% polyacrylamide, 7 M urea and TBE 1× gel electrophoresis. RNA elution was performed at 4°C, shaking overnight in buffer TEN_{250} (10 mM, pH 7 Tris–HCl, 1 mM EDTA, 250 mM NaCl) followed by phenol extraction.

Cleavage reactions were performed as follows. Trace amounts of endogenously radiolabeled RNA were renatured in 0.5 mM pH 7 Tris–HCl, 0.05 mM EDTA by incubation at 85°C for 5 min followed by 10 min at room temperature. Samples were then tempered at 37°C for 2 min. This point is taken as time 0. Three different $MgCl_2$ concentrations were assayed for cleavage reactions. Cleavage reaction is started by adding reaction buffer until final concentrations of 40 mM pH 7 Tris–HCl, 10 mM NaCl and either 10, 1 and 0.1 mM $MgCl_2$. The final concentration of EDTA in the reactions was 0.02 mM. It was necessary to incorporate an extra series of experiments for pGEM-T easy clone RNA at 0.5 mM $MgCl_2$ final concentration. Different time points were taken and cleavage reactions were stopped by adding one volume of 2× formamide buffer and kept frozen at −80°C. Reactions were performed in triplicate and resolved by electrophoresis in 8% polyacrylamide, 7 M urea and 1× TBE gel. Gels were dried and incubated with phosphor-storage screens for scanning in Typhoon and quantification.

The plotted data was fitted to both double exponential (Equation 1) and hyperbolic (Equation 2) equations by the Prism 5 v.500 software (GraphPad Software, Inc.). These equations are as follows:

$$f_c = A + Be^{-k_1 t} + Ce^{-k_2 t} \quad (1)$$

$$f_c = A/(t + T_{0.5}) \quad (2)$$

where f_c is the cleaved fraction, t is the time, A is the cleavage fraction at infinite times, $-B$ and $-C$ are the amplitudes of the observable phases (Equation 1), k_1 and k_2 are the observed first-order rate constant for the fast and slow phases respectively (Equation 1) and $T_{0.5}$ is the time at $f_c = A/2$ (Equation 2).

Checking for 5′-hydroxyl ends

The 77 nt length 3′-fragments generated by co-transcriptional cleavage reactions of pGEM-T-61/L1Tc+77, 7134-171/L1Tc+77 and 55-171/L1Tc+77 constructs were gel purified as described above for uncleaved RNA. For the cleavage 3′-product of the pGEM-T easy construct, one-half of the product was dephosphorylated at 37°C for 30 min with 1 U of calf intestinal alkaline phosphatase (Roche) and subsequently gel purified (3cGEMT*). The other half was not dephosphorylated. Aliquots of each sample RNA were split into two tubes to insure the same radiolabeling level for samples with and without T4 polynucleotide kinase (T4 PNK, Roche). Control samples were carried out in a 10 µl final volume with nuclease-free water. Experimental samples were 5′-end radiolabeled by phosphorylation reaction in 10 µl final volume for 20 min by adding 30 µCi γ^{32}P-ATP, 10 U T4 PNK enzyme (New England Biolabs) and the appropriate reaction buffer. Reactions were stopped by addition of 10 µl of 2× formamide buffer and resolved by electrophoresis in 8% polyacrylamide, 7 M urea, 1× TBE gel.

Determining the HDV-like ribozyme cleavage site by primer extension

The same 3′-cleavage fragments used for the analysis of the hydroxyl nature of their 5′-ends were used unlabeled as substrate for primer extension analysis. Reaction was performed with L1Tc55+77r primer and M-MuLV reverse transcriptase (Roche). First, L1Tc55+77r primer was radiolabeled by phosphorylation by T4 PNK using γ^{32}P-ATP as phosphate donor. The reaction mix contains 15 pmol of primer, 10 U of enzyme and 30 µCi of γ^{32}P-ATP in a final volume of 10 µl and at 1× reaction buffer. Reaction is performed at 37°C for 30 min. The radiolabeled product was gel purified as described for uncleaved RNAs. Reverse transcription reaction was performed by mixing 200 ng of RNA, 1 pmol of primer and 40 U RNase-inhibitor in a final volume of 6.5 µl. This mix is denatured at 85°C for 5 min and primer annealing is performed at 25°C for 10 min. The reaction was carried in 20 µl final volume in the presence of 40 U enzyme, additional 25 U of RNase-inhibitor (Promega), at final concentrations 10 mM of each dNTPs and at 1× reaction buffer. Reaction was performed for 1 h at 37°C and stopped by adding the same volume of 2× formamide buffer.

Manual sequencing of the c7134-171/L1Tc+126 PCR DNA template was performed with the same primer used for sequence ladder. Sequencing reactions were performed using Thermo Sequenase™ Cycle Sequencing Kit (USB) and following manufacturer's indications.

RNA cleavage by RNase P

The RNA fraction of *Escherichia coli* RNAse P known as M1 RNA and its natural substrate pre-tRNA^{tyr} were *in vitro* transcribed using the T7 RNA polymerase kit (PROMEGA) and DNA templates kindly ceded by Dr Jordi Gomez. Both templates were linearized by FokI digestion prior to transcription. RNAs were gel purified by UV shadowing and eluted as described before for uncleaved RNAs. M1 RNA was resuspended in nuclease-free water and adjusted to a final concentration of 500 nM. The same volume of 2× buffer B (20 mM, pH 7.5 Tris–HCl, 20 mM $MgCl_2$, 200 mM NH_4Cl) was added to the reaction. Then, the sample was denatured at 65°C for 5 min and renatured by slow cooling to reach room temperature. The M1 RNA is aliquoted and preserved at −20°C. The different RNAs tested as substrates for M1 RNA cleavage were endogenously radiolabeled during transcription and gel purified. About 300–500 cps of each RNA are incubated for 1 h at 37°C in 50 mM Tris–HCl, 0.1 mM NH_4Cl, 100 mM $MgCl_2$ and 4% polyethylene glycol with 20 U of RNase-inhibitor and 1.25 pmol of M1 RNA in a final volume of 10 µl. The reactions were stopped by adding 10 µl of 2× formamide buffer and resolved by 8% polyacrylamide, 7 M urea, TBE 1× gel electrophoresis.

RESULTS

Identification of an HDV-like ribozyme folding in the 5′-end of L1Tc

Based on the recent identification of several HDV-like ribozyme candidates located at the 5′-end of retrotransposons (19), we performed as a first attempt a manual structure-based search in the sequence located at the 5′-termini of the L1Tc non-LTR retrotransposon (Figure 1A). Folding and sequence analyses showed specific characteristics at the first 77 nt of L1Tc that are partially compatible with those present in the three HDV-like ribozymes that experimentally have been shown to be active (Figure 1B$_1$) (19,22,23). The three helices of HDV-like ribozymes, known as P1, P2 and P4, are detected in the putative L1Tc ribozyme folding as well

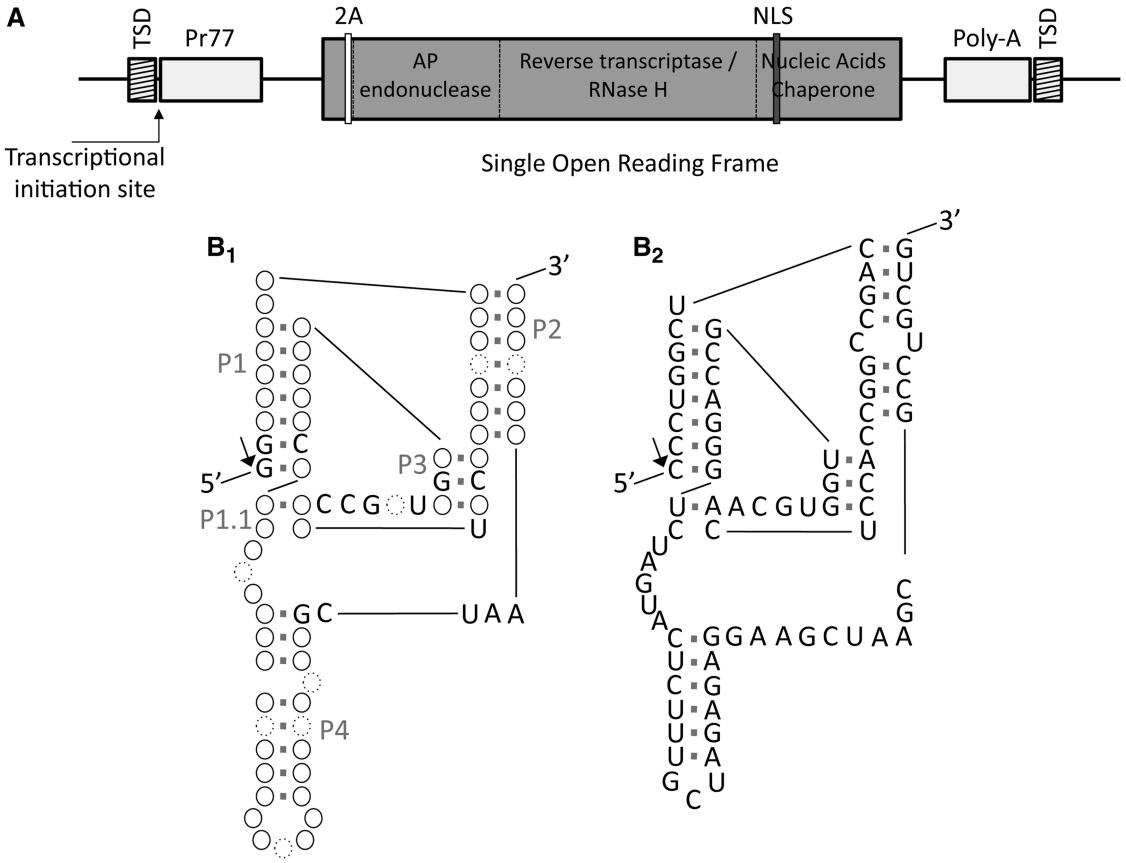

Figure 1. Schematic representation of L1Tc element, HDV-like ribozyme and L1TcRz foldings. (**A**) L1Tc non-LTR retrotransposon present in *T. cruzi* genome. Its sequence is flanked by two direct repetitions called target site duplications (TSDs). The 5′-UTR contains a 77-nt internal promoter (Pr77). Some copies present a single open reading frame that codifies for the following functional domains: AP endonuclease, reverse transcriptase (RT), RNase H and nucleic acids chaperone. A 2A autoproteolitic domain (2A) is present within the N-terminal end preceded by a short peptide. A nuclear localization signal (NLS) is detected within the nucleic acid chaperone domain. The 3′-UTR is composed by a 21 nt fragment and a poly-A track. (**B$_1$**) HDV-like ribozyme and (**B$_2$**) L1TcRz foldings. HDV-like ribozyme scheme is based on the three described ones that have been fully characterized (19,22,23). Solid line circles represent conserved base pairings of variable composition; dotted line circles also represent variability in nucleotide number (from 0 to more than one); the conserved nucleotides in the three characterized ribozymes are indicated. Helix names are shown as P1, P1.1, P2, P3 and P4. The arrow points to the cleavage site.

as the two *pseudo*knots. The P3 *pseudo*knot involves 3 bp, while the P1.1 *pseudo*knot involves 1 or 2 bp (Figure $1B_2$). Furthermore, there is high identity at the nucleotide level of several bases implicated in the helices and *pseudo*knots formation.

Co-transcriptional self-cleaving activity in the L1Tc 5′-termini RNA. Identification of the catalytic domain

To determine the existence of co-transcriptional cleavage activity in the L1Tc 5′-end RNA, a series of DNA templates bearing L1Tc 5′-end sequences were generated for *in vitro* transcription. The cleavage activity associated with the sequence that adopts the HDV-like folding (the 77-nt length sequence of the 5′-end of L1Tc) was evaluated as well as shorter (70, 59 and 47 nt) and longer (126 and 152 nt) sequences (see scheme in Figure $2A_1$–$3C_1$). To evaluate whether or not sequences located upstream of L1Tc had any influence in the activity of the ribozyme, three DNA template groups were constructed. Sequences located upstream of the first nucleotide of two different copies of L1Tc were assayed. One of the sequences corresponds to a genomic clone, called 7134 (Figure $2A_1$), and the other, named clone 55 (Figure $2B_1$), corresponds to a cDNA clone isolated by positive hybridization to an RT probe in the screening of a *T. cruzi* cDNA library (7) (accession number X83098). The third DNA template group was generated bearing a sequence of the pGEM-T easy vector located upstream of L1Tc sequences (Figure $2C_1$).

The analysis of the RNA products revealed co-transcriptional cleavage activity in all constructs containing the 77 nt in length sequence as they gave rise to the expected

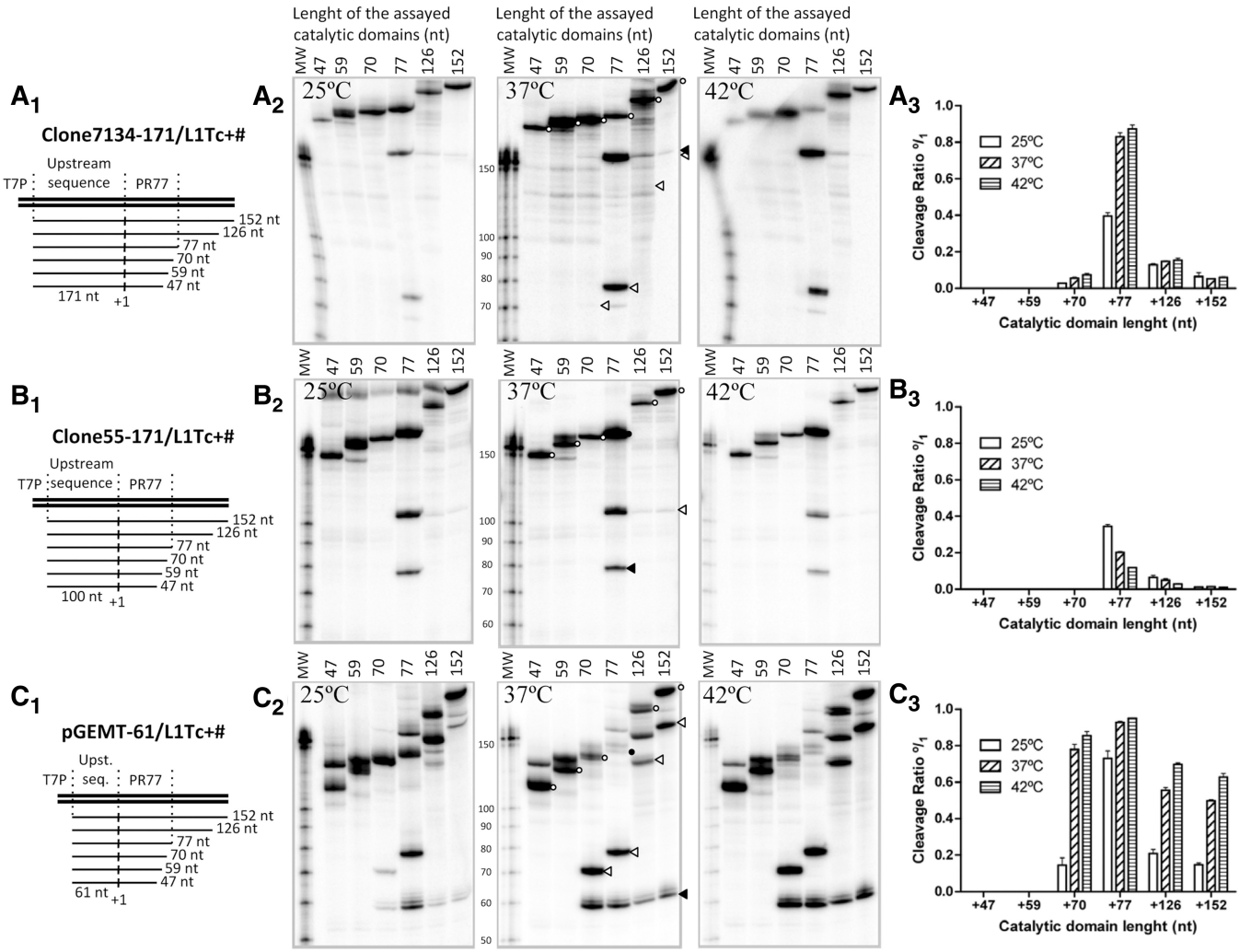

Figure 2. Co-transcriptional cleavage of L1Tc RNAs of different length and at different temperatures. Co-transcriptional cleavage of the L1TcRz was checked in three different insertion environments: genomic clone 7134 (**A**), cDNA clone 55 (**B**) and L1Tc-unrelated pGEM-T easy clone (**C**). The constructs of different length assayed are represented as thin lines in A_1, B_1 and C_1 schemes. The vertical lines labeled as +1 represent the expected cleavage point. Autoradiographs of the electrophoretic analysis of the transcription and co-transcriptional cleavage reactions at different temperatures are represented in A_2, B_2 and C_2. The full-length uncleaved fragment is represented by empty circles; the uncleaved fragment with the complete 77-nt length L1TcRz catalytic domain is represented by a solid circle; the cleavage 5′-product is represented by solid arrowheads; the cleavage 3′-product is represented by empty arrowheads. The average cleavage quantification of three independent triplicates of each reaction is represented on A_3, B_3 and C_3.

full-length RNAs of 248, 177 and 138 nt (Figure $2A_2$–C_2, respectively, lane 77, black circle) in addition to the cleavage 5′-products of the 171, 100 and 61 nt (Figure $2A_2$–C_2, respectively, lanes 77 black arrowheads) and the 77 nt 3′-product (Figure 2A2–C2, lanes 77 white arrowheads). Construct +70 had significant catalytic activity in the pGEM-T easy clone (Figure $2C_2$, lane 70), weak catalytic activity in clone 7134 (Figure $2A_2$, lane 70) and an undetectable one in clone 55 (Figure 2B, lane 70). No co-transcriptional activity was observed when subsequent deletions +59 and +47 of the three clones were assayed (Figure $2A_2$–C_2, lanes 59 and 47). The addition of 49- and 75-nt length sequences located downstream of the L1Tc Pr77 in constructs +126 and +152 did not prevent the co-transcriptional cleavage although it was substantially reduced (Figure $2A_2$–C_2, lanes 126 and 152). Data representing the cleavage levels of all mentioned constructs at different temperatures is shown in Figure 2, panels A_2–C_2 and represented in Figure 2, panels A_3–C_3. The sequences located upstream of L1Tc influenced the cleaving efficiency since maximal cleavage was observed in the pGEM-T easy -61/L1Tc+77 construct. A lower cleaving activity was observed in clones 7134–171/L1Tc+77 and particularly in clones 55–100/L1Tc+77. The highest co-transcriptional cleavage activity was observed when the reactions were carried out at 42°C in 7134 and pGEM-T easy clones. A slight decrease was observed when the reactions were performed at 37°C. However, the co-transcriptional activity drastically dropped when the reactions were performed at 25°C. An opposite behavior was observed when clone 55 was assayed.

Nature of the 5′-ends generated by L1TcRz

To determine the nature of the 5′-ends of the products generated by L1TcRz, we analyzed whether the 3′-products are directly accessible to 5′-end radioactive labeling by T4 PNK using γ^{32}P-ATP as phosphate donor (see scheme of Figure 3A). Thus, co-transcriptional assays using clones 7134–171/L1Tc+77, 55–171/L1Tc+77 and pGEM-T easy -61/L1Tc+77 sequences were performed as before. The endogenously radiolabeled RNA 3′-fragments were gel purified and the same amount of each was both used for the 5′-end phosphorylation assay and kept as control of the initial radiolabeling stage. The results showed that the 5′-ends of the 3′-fragments generated in these reactions were sensitive to phosphorylation (Figure 3B) confirming the hydroxyl nature of the 5′-end of the 3′-fragment generated by the cleavage activity of the L1TcRz. To confirm that the observed increase of labeling of the 3′-products was not due to the presence of a small subpopulation of dephosphorylated molecules, one-half of the 3′-product of the pGEM-T

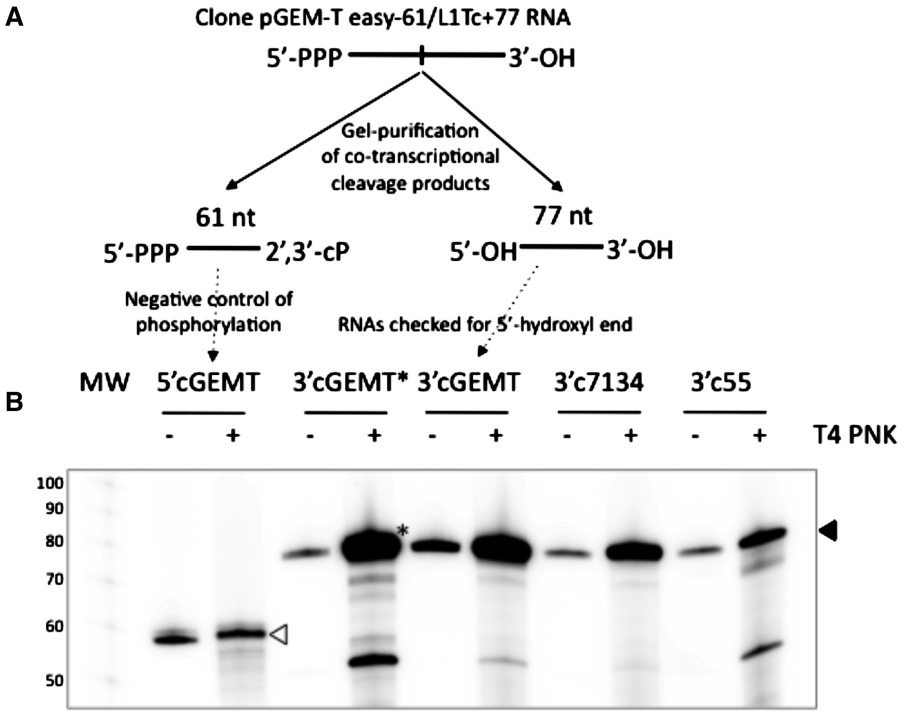

Figure 3. Analysis of the 5′-hydroxyl nature of the ends of the cleavage 3′-products. Schematic cleavage reaction of the clone pGEM-T easy −61/L1Tc+77 RNA is represented in (**A**). The uncleaved RNA is expected to have 5′-triphosphate and 3′-hydroxyl ends. The cleavage 5′- and 3′-products are expected to have 2′,3′-cyclic phosphate and 5′-hydroxyl ends, respectively. The T4 polynucleotide kinase (T4 PNK) challenge is represented in (**B**). 5′-hydroxyl ends, not 5′-phosphate, are sensible to phosphorylation by T4 PNK. Same quantity of endogenously radiolabeled cleavage fragments was both ice preserved in reaction buffer and phosphorylated by T4 PNK using gamma ^{32}P-ATP as phosphate donor. The cleavage 3′-products of clones 7134, 55 and pGEM-T easy RNAs were further radiolabeled confirming the expected 5′-hydroxyl nature of their 5′-ends (solid arrowhead). The 61 nt in length RNA 5′-product of the cleavage of the pGEM-T easy construct is used as negative control in the phosphorylation reaction (the empty arrow indicates the labeled 5′-product). One of the 3′-products is pre-treated with alkaline phosphatase prior to being treated with T4 PNK. (marked with an asterisk).

easy construct was dephosphorylated by alkaline phosphatase treatment, gel-purified and subsequently phosphorylated (substrate 3cGEMT* in Figure 3). The results showed that the level of phosphorylation was the same in the dephosphorylated and in the not dephosphorylated RNA samples as an indication that the whole population of the molecules had a 5′-hydroxyl end.

Cleavage kinetics and Cleavage site of L1TcRz

Clones 7134–171/L1Tc+77, 55–100/L1Tc+77 and pGEM-T easy -61/L1Tc+77 uncleaved products (full-length transcripts) generated by an *in vitro* transcription, shown in Figure $3A_2$–C_2 (lanes 77-black circles), were gel purified and used for the analysis of the cleavage kinetics (schematic representation in Figure $4A_1$–A_3). Cleavage reactions were performed at physiological temperature (37°C) at low ionic strength in the presence of EDTA and at different Mg^{2+} concentrations. As it is shown in Figure $4B_1$–B_3, the cleavage reactions are Mg^{2+} dependent as no cleavage products were observed at 0.1 mM Mg^{2+}. A positive progression in the catalytic rates was detected when the reaction was assayed at higher Mg^{2+} concentrations (1 and 10 mM). The highest cleaving efficiency of the RNA derived from clones 7134–171/L1Tc+77 was reached at a Mg^{2+} concentration of 10 mM (Figure $4B_2$). The highest cleaving efficiency of the RNA derived from clone pGEM-T easy-61/L1Tc+77 was reached at 1 mM Mg^{2+} (Figure $4B_1$). Under the condition used in the present experiments, almost no appreciable cleaving activity was observed for the RNA derived from clones 55–100/L1Tc+77 (Figure $4B_3$). Figure 4, panels C_1–C_3, shows the cleavage reaction kinetics for each assayed clone at 1 mM Mg^{2+}.

The analysis of the data shown in Figures $5B_1$ and B_2 indicates that a high proportion of the cleavage of the pGEM-T easy-61/L1Tc+77 and clones 7134–171/L1Tc+77 transcripts occur in the first 2 min of the reaction suggesting that L1TcRz is a fast-reacting ribozyme. In fact, the maximum cleavage of pGEM-T easy-61/L1Tc+77 RNA was reached at 0.277 ± 0.029 min at 10 mM Mg^{2+}. The cleavage kinetic curve of clone 7134–171/L1Tc+77 RNA at 1 mM Mg^{2+} (Figure $4B_2$) fits better to a two-phase decay equation (solid line) than to a hyperbolic equation (dotted line). This behavior has been previously described for the HDV-like ribozyme of the CPEB3 gene (28). The R^2 coefficient of the cleavage reaction for clones 7134–171/L1Tc+77 RNA at 10 mM Mg^{2+} and pGEM-T easy-61/L1Tc+77 RNA at any Mg^{2+} concentration is higher for a two-phase decay curve than for a hyperbolic one. However, due to the fast reacting activity of the ribozyme the fitting of the equation performed by Prism 5 v.5.00 software (GraphPad Software, Inc.) remains statistically ambiguous. The two-phase decay curve used for the data fitting corresponds to the double exponential equation used by Chadalavada *et al.* (28). The parametric data of both

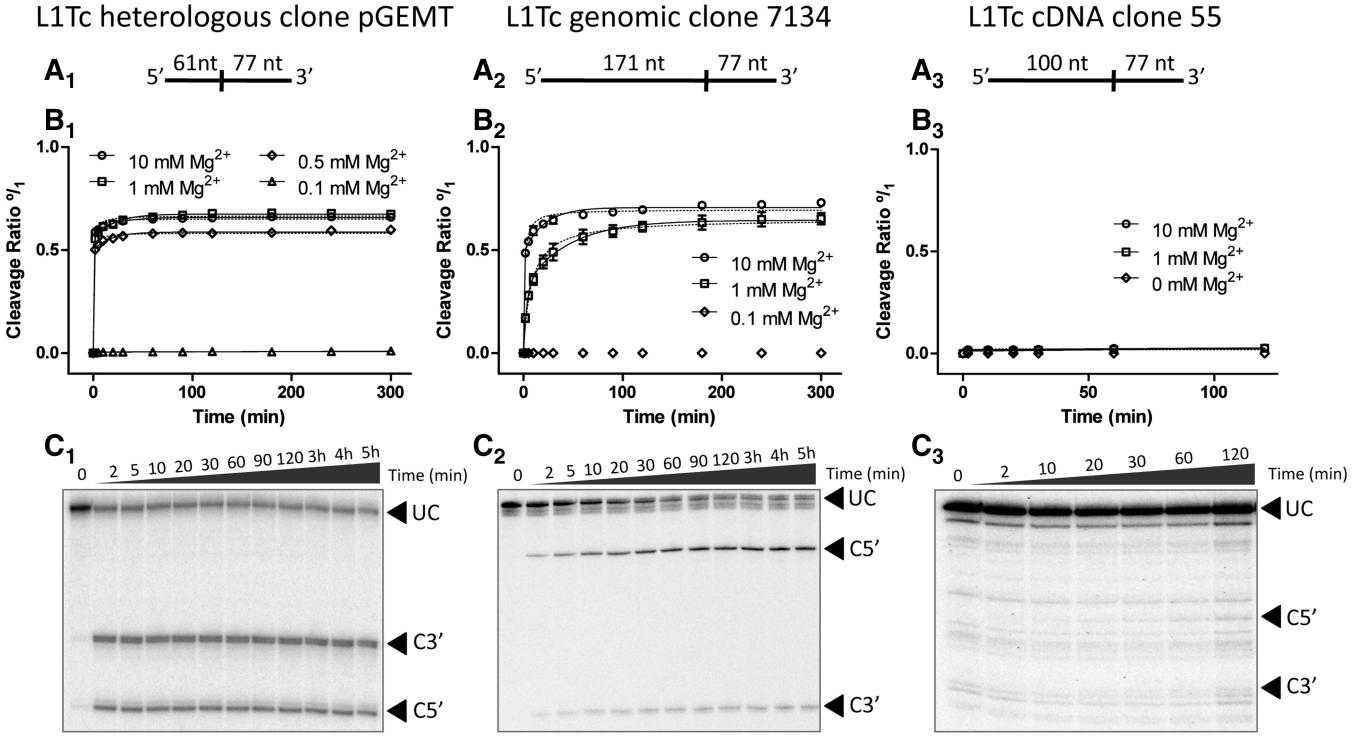

Figure 4. Ribozyme kinetics. pGEM-T easy −61/L1Tc+77 ($\mathbf{A_1}$), clone 7134–171/L1Tc+77 ($\mathbf{B_1}$) and clone 55–100/L1Tc+77 ($\mathbf{C_1}$) uncleaved RNAs were gel purified for cleavage reaction assays. The cleavage point (vertical line) and the expected size of the cleavage fragments are also represented in (**A**). Cleavage reactions were performed at 37°C at different Mg^{2+} concentrations. The result of the quantification of triplicates of each reaction is represented in ($\mathbf{B_1}$), ($\mathbf{B_2}$) and ($\mathbf{B_3}$) and the data are fitted to both a two-phase decay curve (solid line) and a hyperbolic one (dotted line). The cleavage rate is represented from 0 (minimum) to 1 (maximum) indicated as (°/1). Autoradiographs of an example kinetic of each assayed RNA at 1 mM Mg^{2+} is shown in (C). Arrowheads indicate the uncleavaged RNAs (UC), and the 5′- (C5′) and 3′-products (C3′).

hyperbolic and double exponential data fitting are summarized in the Supplementary Table S1.

To analyze the specific cleavage site of L1TcRz, primer extension assays were performed using the 3′-end fragments resulting from the co-transcriptional cleavage reactions of the pGEM-T easy-61/L1Tc+77 and clones 7134–171/L1Tc+77 templates. The gel-purified products were used as template for the extension of the radiolabeled L1Tc55+77r primer and the products resolved in a 6% polyacrylamide, 7 M urea, sequencing gel. Sequencing of the clone 7134–171/+126 DNA template was also carried out using the above-mentioned primer and employed as molecular weight marker. As shown in Figure 5, the full-length extended products correspond to the first nucleotide of the L1Tc element (white arrowhead). The cleavage site coincides with the transcription initiation site of Pr77, previously described by Heras *et al.* (5). It is noteworthy that two strong elongation stops were also observed at G_{+37} and C_{+38} positions (black arrowhead) that may be due to the difficulty of the reverse transcriptase to extend the primer through the triple GC base pairing predicted for the P1 helix of L1TcRz (see folding Figure $1B_2$).

Influence of Pr77 downstream sequence in the required folding for L1TcRz ribozyme activity

As it was shown in Figure 2, the L1Tc sequences located downstream of the first 77 nt of the element partially inhibit the co-transcriptional cleavage activity of L1TcRz. To determine the specificity of the downstream sequence in disabling the ribozyme activity, a series of DNA templates were generated in which a pGEM-T easy sequence of 146 nt in length, named 146xeno, replaced the L1Tc downstream +77 nt sequence in clones 7134, clone 55 and pGEM-T easy. As it is shown in Figure $7A_1$, lanes L1Tc7134, L1Tc55 and L1TcpGEMT, co-transcriptional cleavage was observed in all reactions. Quantification of the cleavage products and the comparison of their abundance with those produced by the RNAs that bear +77, +126 and +152 L1Tc sequences indicate (Figure $6A_2$) that the replacement of the sequences located downstream of Pr77 in L1Tc by those of pGEM-T easy partially restores the co-transcriptional cleavage of clones 7134, 55 and pGEM-T easy (+77) derived RNAs.

We further checked the implication of a sequence different from L1Tc but naturally located downstream of Pr77 in the ribozyme activity of L1TcRz. Thus, the sequence from the NARTc non-autonomous retroelement was chosen as this element shares the first 77 nt with L1Tc with close to 100% identity and 56% of the nucleotide conservation with the Pr77 downstream sequence. The two templates, NARTc −10/+126 and NARTc −10/+152, exhibited cleavage products of a size consistent with cleaving at position +1 of L1Tc (Figure $6B_1$). Quantification of the cleavage products (Figure $6B_2$) in NARTc −10/+126 and −10/+152 RNAs indicates that the NARTc downstream sequences do not negatively influence the cleavage efficiency as much as the corresponding fragments from L1Tc do.

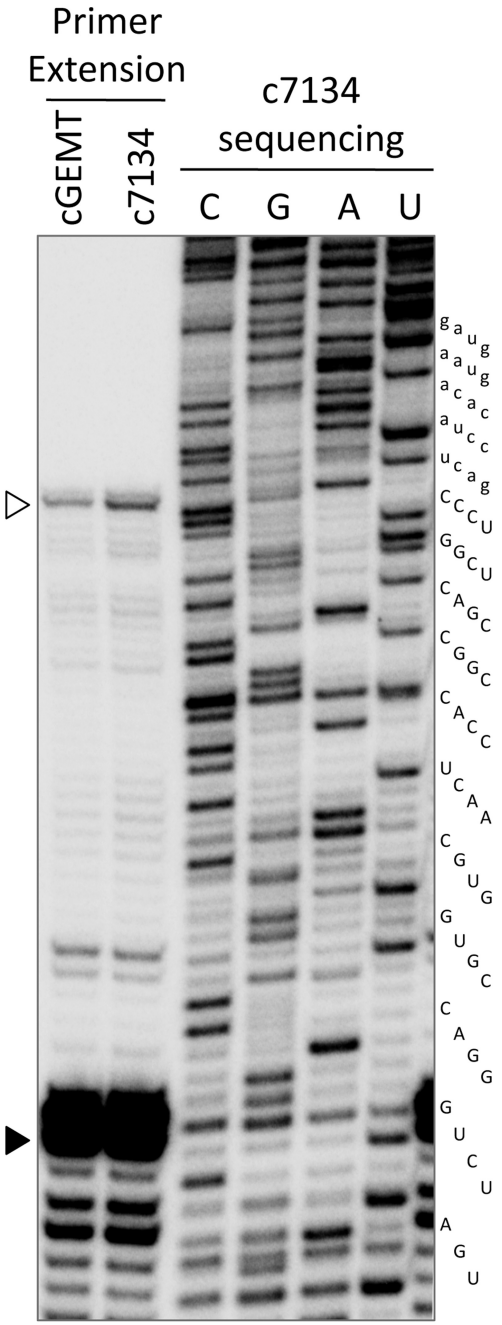

Figure 5. Primer extension of the 3′-products of *in vitro* cleavage. 3′-Products of co-transcriptional cleavage of pGEM-T easy −61 and c7134–171/L1Tc +77 RNAs were used as template for primer L1Tc55+77r extension. Manual sequencing of clones 7134–171/L1Tc+126 PCR template using the same primer was resolved in the same polyacrylamide gel as size standard. The maximum extension corresponding to the 5′-end of the products (empty arrowhead) is localized in the C_1 nucleotide of the element. An intense premature stop of the extension is localized at position G_{37} (solid arrowhead) corresponding to the tight junction (triple GC base pairing) of the P1 helix (Figure 2).

The folding adopted by the L1Tc 5′-UTR sequence is cleaved by RNAse P

We further investigated the implications of L1Tc sequences located downstream of Pr77 in limiting the

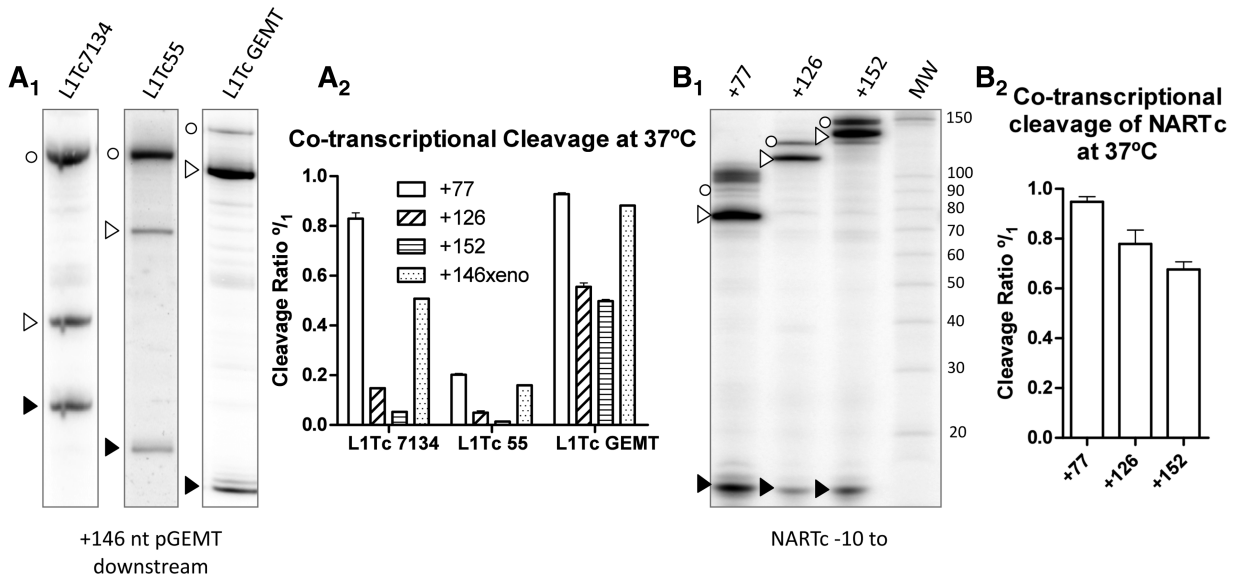

Figure 6. Role of downstream nucleotides in L1TcRz activity. (**A**) Co-transcriptional cleavage analyses of constructs clones 7134–171/L1Tc+77, clone pGEM-T easy −61/L1Tc+77 and clone 55–100/L1Tc+77 having fused the 146 nt of pGEM-T easy to their 3′-ends. Autorradiograph of the analytical gel electrophoresis of the 146xeno constructs co-transcriptional cleavage reactions is shown in (**A_1**). Quantification of triplicates of these reactions and its comparison with different length L1Tc sequences, as shown in Figure 2 lacking the +146Xeno sequence, is indicated in (**A_2**). NARTc −10/+77, +126 and +152 constructs were also analyzed for co-transcriptional cleavage activity (**B**). Autoradiograph of the analytical gel electrophoresis of NARTc constructs is shown in (**B_1**) and quantification of triplicates of these reactions in (**B_2**). Empty circles indicate the uncleaved RNAs, empty arrowheads correspond to the 3′-products and the solid arrowheads correspond to the 5′-products.

proper folding required for L1TcRz activity. Thus, we analyzed the particular electrophoretic mobility of the conformation endowed with ribozyme activity and checked whether the sequences located downstream of the first 77 nt of L1Tc (49 and 75 nt in length, RNAs −10/+126 and −10/+152, respectively) induce a particular folding that may inhibit the one required for L1TcRz ribozyme activity. The results revealed a global switching of the folding of the L1Tc RNA from a catalytic form to a non-catalytic one when sequence downstream the +77 position is included in the molecule (Supplementary Figure S1). Since this inhibition is restricted to L1Tc (not NARTc) and we suspect the existence of an IRES-like structure in the L1Tc 5′-UTR but not in NARTc, we checked the existence of a RNase P recognition motif in these regions. Thus, we analyzed whether the folding adopted by the 3′-products generated by L1Tc ribozyme in the RNAs +77, +126 and +152 is recognized and cleaved by the M1 RNA of *E. coli* RNase P. Thus, *in vitro* transcription was carried out using the constructs pGEM-T easy −61/L1Tc+77, /+126 and /+152. The 3′-products of the L1TcRz cleavage were gel purified for subsequent RNase P digestion. Pre-tRNA^{tyr} was used as a positive control. The results show that the cleavage of the RNA + 126 and of RNA + 152 gave rise to two fragments: one close to 50 nt long in each case and the other close to 75 and 100 nt long, respectively (Figure 7A). Cleavage products were not detected in the control reactions carried out in the absence of RNase P or in the reactions of the RNA + 77. The fact that there are only two fragments after RNase P digestion indicates that there is a single recognition site in the L1Tc 5′-UTR. Since one of the two fragments resulting from the RNase P cleavage of both substrates (Figure 7A, labeled by an asterisk) has the same size (50 nt in length, approximately), it is most likely that the RNase P cleavage site is located in the vicinity of position +50 of the L1Tc RNA (see scheme of Figure 7B).

To confirm that the RNase P-digested structure corresponds to a folding adopted by the mentioned L1Tc sequence, transcripts of the same size were generated bearing NARTc sequences. The data shown in Figure 7C indicate that only the pre-tRNA^{tyr}, used as a control, was digested by RNAse P. Cleavage by RNAse P was not detected when the L1Tc+77/+146xeno RNA was assayed (data not shown). Further experiments will be required to unambiguously show the presence of an IRES at the 5′-end of the L1TcRz.

DISCUSSION

In this work, we describe the presence of an HDV-like ribozyme at the 5′-end of the L1Tc mRNA, named L1TcRz. When the first 77 nt of L1Tc mRNA were analyzed, we manually obtained a compatible HDV-like ribozyme folding structure (19,22,23). The analysis of the transcription of templates from positions upstream of the first nucleotide of the L1Tc element revealed the presence of RNA fragments putatively generated by co-transcriptional cleavage activity. Co-transcriptional cleavage assays carried out on the sequences located at the 5′-end of L1Tc having various lengths (+47, +59, +70 and +77) show that the highest autocatalytic activity of L1TcRz resides within the first 77 nt of L1Tc. Nevertheless, in the pGEMT +70 construct, the L1Tc

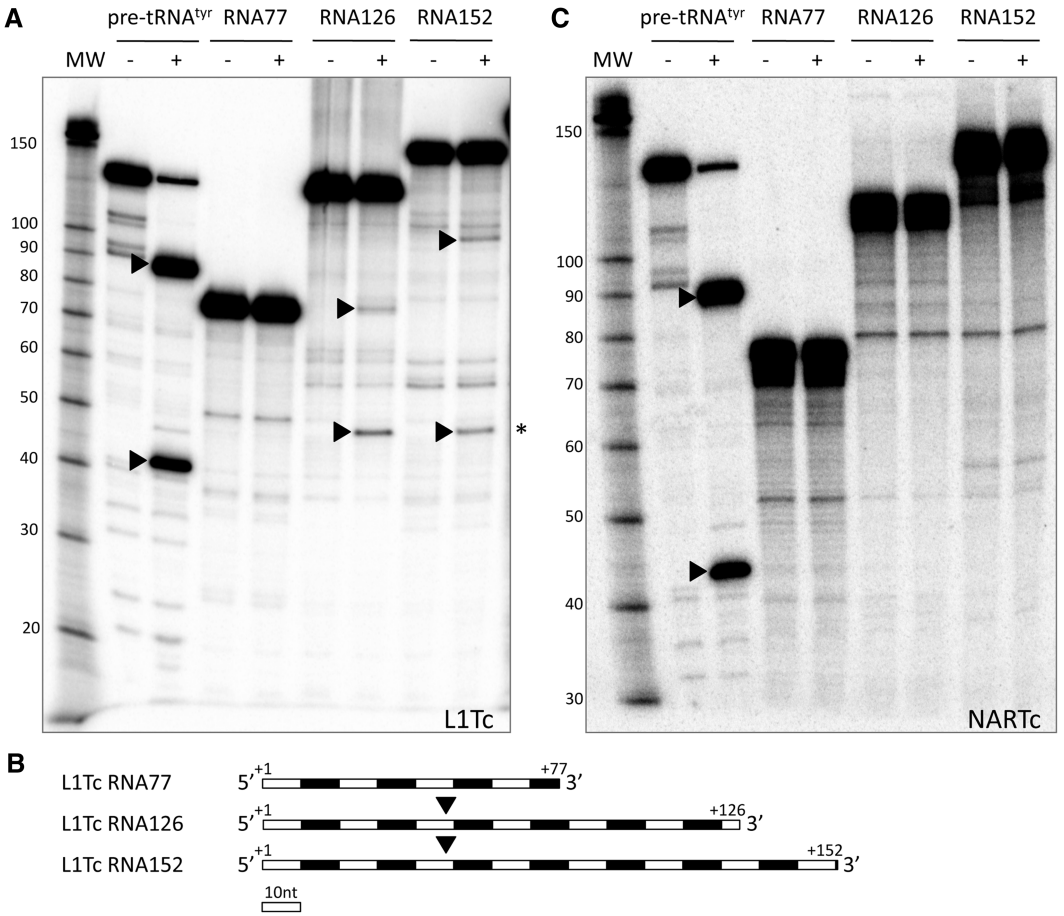

Figure 7. Presence of RNase P cleavage motif within L1Tc 5′-UTR RNA. L1Tc and NARTc +1/+77, +1/+126 and +1/+152 RNAs were subjected to cleavage reaction by *E. coli* RNase P M1 RNA. Autoradiographs of the gel analysis of the reactions are shown in (A) and (C). (+) lines correspond to M1 RNA treatment and (−) to the control buffer-only. Two fragments (pointed by arrowheads) are generated in the L1Tc RNAs +1/+126 and +1/+152 cleavage reactions compared to the control in buffer, denoting the existence of a single cleavage point as in the control pre-tRNA^{tyr} (**A**). The generation of fragments (asterisks) having the same size in the L1Tc +1/+126 and +1/+152 RNAs cleavage reactions indicates that this corresponds to the 5′-product (**B**) and that the cleavage is produced close to the nucleotide +50. No cleavage site is detected in NARTc RNAs (**C**).

ribozyme cleaves at high levels even in the absence of a P2 stem. This length corresponds to the catalytic sequence predicted to appropriately fold into an HDV-like ribozyme conformation. Moreover, the overall size of these putative RNA fragments was consistent with the expected cleavage point according to the manually predicted folding. The catalytic activity was confirmed when gel-purified full-length uncleaved transcripts self-cleaved in a magnesium-dependent reaction, generating fragments with the same size as those co-transcriptionally obtained.

The kinetics of the self-cleaving reaction fit a two-phase decay curve described specifically for the HDV-like ribozyme type. The 5′-hydroxyl end of the co-transcriptionally 3′-products obtained was consistent with the biochemical cleavage reaction typical of HDV-like ribozymes (29). Primer extension analysis with these fragments confirmed the expected cleavage point at the 5′-side of the +1 nt of L1Tc. In previously reported primer extension analyses carried out using the epimastigotes polyadenylated RNA-purified fraction, we described that L1Tc mature messengers started at L1Tc +1 nt (5). It was consequently deduced that Pr77-driven L1Tc mRNAs transcription started at +1 nt (5). However, taking into account the results reported here, we believe that the *in vivo* cleavage activity of the ribozyme may well be responsible of the detected transcripts starting at the L1Tc +1 nt. This fact suggests that L1TcRz is *in vivo* functionally active.

The data shown indicate, therefore, that the 77-nt length sequence of the L1Tc 5′-end has two different functions: the internal promoter (Pr77) previously described as DNA (5) and the HDV-like ribozyme currently described as RNA. The mechanisms to generate mRNAs for non-LTR retrotransposons may imply either the utilization of internal promoters (1–3,5,30) or an HDV-like ribozyme that releases mRNAs from host co-transcripts (19). The HDV-like ribozyme cleaves upstream its catalytic domain and thus both, ribozymes and internal promoters, persist within the mRNAs and their functions are preserved after transposition. The only element, up to date, which has been described to have an HDV-like ribozyme is the R2 retroelement of *Drosophila* that

specifically mobilizes into the 28S rDNA and co-transcribes its mRNAs within the ribosomal cassette (31). L1Tc becomes now the first-described retroelement with a dual internal promoter–ribozyme system.

The presence of the ribozyme in L1Tc could be related to the particular transcriptional system of the Trypanosomatids, the natural host of L1Tc. The genome of these parasites is organized in large gene clusters unidirectionally and polycistronically transcribed, and separated by the called strand switch regions (SSRs) (32). Many L1Tc/NARTc copies are localized in this SSRs and it has been previously speculated with the possible involvement of Pr77 in the transcription promotion of the large clusters (5). However, despite some L1Tc/NARTc sequence specificity for insertion has been reported (15), its distribution is considered random (14), so the ribozyme activity could be yielding the appropriate releasing of L1Tc/NARTc mRNAs from large polycistronic transcripts. We cannot exclude the possibility that Pr77 initiates transcription of some nucleotides upstream of the ribozyme cleavage site. The L1TcRz activity may insure a precise 5′-end identical for Pr77-driven transcribed RNAs or those released from polycistronic transcripts. Since spliced leader RNA is absent in the *in vivo* L1Tc mRNAs and a cap structure has not been detected at their 5′-ends (5), the end generated by the ribozyme could be related to a precise 5′-structure that supplies the protective cap function. L1Tc and NARTc copies have been found in tandem and isolated distributions, in addition to be associated to genomic regions rich in repetitive DNA sequences (14). Considering the faint polyadenylation signals of these elements, the ribozyme could preserve the monocistronic character of the mRNAs.

Our data also indicate that HDV-like ribozyme cleavage activity can be modulated by sequences upstream of the cleavage point. The cleavage ability of the three assayed constructs seems to adequately fit to the predicted upstream sequence-dependent ribozyme-folding interferences. The RNA derived from clones 55 and 7134 includes a G_{-1} capable of extending the P1 helix and disrupting the single base pair of the P1.1 *pseudo*knot (Supplementary Figure S2A and S2B, bi-headed gray arrow). This fact has been also described for the R2 ribozyme in which the nucleotides located upstream of the catalytic folding, able to extend the helix P1, can prevent the P1.1 *pseudo*knot formation and moderately affect the cleavage efficacy (33). The L1Tc ribozyme may be sensitive to P1.1 *pseudo*knot disruptions by helix P1 extension because this *pseudo*knot is composed only by a single base pair. For clone 55, the RNAfold software (34) predicts the existence of an alternative helix between the 7 nt upstream of the L1TcRz cleavage site and 7 nt located within the ribozyme core (Supplementary Figure S2A, dotted gray line). A similar folding disruption has been reported to dramatically affect the cleavage activity of the ribozyme present in the human CPEB3 gene (28). Since we could not find any handicap for pGEM-T easy cloned ribozyme (Supplementary Figure S2C), all these features can explain the progressive decrease of the catalytic activity from clone pGEM-T easy to clone 7134 and to clone 55. It is interesting to note that clone 55 has been obtained from a *T. cruzi* cDNA library and bears 110 nt non-related to L1Tc sequence at its 5′-end. This is consistent with the L1TcRz low cleavage rate detected in the clone 55 construct.

The cleavage reactions of full-length RNAs performed at different Mg^{2+} concentrations showed that the catalytic activity of L1TcRz is magnesium dependent. Mg^{2+} is commonly required for the proper folding of RNAs and, consequently, for the catalytic activity of structure-dependent RNAs, like ribozymes. Moreover, despite the fact that some cleavage has been recorded under high hydrostatic pressure conditions in total absence of Mg^{2+} (35), the HDV ribozyme is the only one of the small catalytic ribozymes probed to coordinate a Mg^{2+} cation at its catalytic center (36,37). Thus, the Mg^{2+} dependence of the L1TcRz is consistent with the ribozyme nature of the reported cleavage.

Our results show that the co-transcriptional cleavage *in vitro* decreases when sequences longer than 77 nt L1Tc are assayed (constructs +126 and +152). This inhibition is neither observed in NARTc constructs nor in 146xeno. Thus, the L1Tc nucleotides located downstream of the ribozyme fragment appears to inhibit its activity. This sequence-dependent attenuator effect may be a consequence of the use of T7 RNA polymerase for *in vitro* transcription and could not be occurring *in vivo*. The premature transcription of downstream attenuator sequences due to the T7 RNA polymerase high speed (~20-fold faster than the eukaryotic enzymes) has shown to prevent the HDV ribozyme folding and to inhibit consequently the co-transcriptional cleavage (38). Since the NARTc RNA downstream sequences do not interfere with the cleavage rate, we believe that these data strongly suggest the existence in L1Tc RNA of a consistent and specific conformation that is incompatible with the cleavage activity.

The data presented here support the co-transcriptional function of the L1TcRz and are consistent with the previously described activity of the wild-type HDV ribozyme whose RNA co-transcriptional folding promotes a catalytically active conformation rather than the post-transcriptional one (38). Thus, the highest cleavage level is always reached in co-transcriptional assays when compared to those performed with full-length gel-purified RNAs. In fact, the L1TcRz is a fast reacting ribozyme [faster than the human CPEB3 (28)]. These data suggest that the RNA cleavage is produced during the first events of the element expression allowing an *in vivo* early co-transcriptional cleavage.

It is worth noting that the 126- and 152-nt length L1Tc RNAs, but neither the same length RNAs of NARTc nor the 146xeno have a RNase P cleavage site at position +50 of Pr77 RNA. The existence of a RNase P-recognition motif together with the absence of a capped spliced leader structure at the L1Tc RNAs 5′-end suggest a cap-independent translation mechanism for L1Tc, such as an IRES. Consisted with this, NARTc does not code for proteins (16) and lacks this IRES-related RNase P motif. L1Tc+77 RNA is not cleaved by the RNase P despite the fact that the cleavage position is included within this sequence. This result shows that the RNase P

recognition motif is not present in the first 77 nt of L1Tc although some of the sequences forming this motif are present in this fragment. An IRES has been shown to precede each one of the two ORFs of the mouse L1 (25). Further studies will be required to evaluate the existence of an IRES also at the 5′-end of L1Tc. Interestedly, both the Pr77 derived and the L1TcRz-released transcripts of NARTc/L1Tc may lack the cap structure and coincidently have the same 5′-end. This fact could be related to the requirement of an appropriate 5′-end structure for efficient protection of the uncapped RNA.

As mentioned above, Pr77-like sequences are restricted neither to L1Tc nor to retrotransposons and to Trypanosomatids. Sequence searches have identified Pr77 homologous sequences in a wide range of organisms from fungi to higher eukaryotes. Furthermore, manual searches have identified a folding structure compatible to HDV-like ribozymes in Pr77-like sequences from *Haliangium ochraceum*, *T. brucei* and *Penicillium chrysogenum* (laboratory data). *In vitro* transcription assays and co-transcriptional cleaving analyses have shown that sequences from *T. brucei* homologous to Pr77 included in the *ingi* and RIME retroelements have also HDV-like ribozyme activity (Carreira, P. *et al*. manuscript in preparation). The origin and acquisition of this type of sequences by a wide range of organisms generate controversy and it is still unknown. The analysis of these and other related sequences will help us to elucidate this question.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We are grateful to Dr Jordi Gomez from IPBLN-CSIC for providing the pre-tRNA^{Tyr} and RNase P M1 RNA templates and for his assistance in RNase P-cleavage assays.

FUNDING

Plan Nacional I+D+I del Ministerio de Ciencia e Innovación (MICINN)—Spain (grant numbers BFU2007/65095/BMC, BFU2007-64999/BMC, BFU2010-16470); Instituto de Salud Carlos III (ISCIII)—Redes Temáticas de Investigación Cooperativa en Salud (RETIC)—Spain (grant numbers RD06/0021/0014, RD06/0021/0008) and Fondo Europeo de Desarrollo Regional (FEDER). Funding for open access charge: Plan Nacional I+D+I del Ministerio de Ciencia e Innovación (MICINN)- Spain (grant number BFU2010-16470)

Conflict of interest statement. None declared.

REFERENCES

1. McLean,C., Bucheton,A. and Finnegan,D.J. (1993) The 5′ untranslated region of the I factor, a long interspersed nuclear element-like retrotransposon of Drosophila melanogaster, contains an internal promoter and sequences that regulate expression. *Mol. Cell. Biol.*, **13**, 1042–1050.
2. Mizrokhi,L.J., Georgieva,S.G. and Ilyin,Y.V. (1988) Jockey, a mobile Drosophila element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. *Cell*, **54**, 685–691.
3. Swergold,G.D. (1990) Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.*, **10**, 6718–6729.
4. DeBerardinis,R.J. and Kazazian,H.H. Jr (1999) Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics*, **56**, 317–323.
5. Heras,S.R., Lopez,M.C., Olivares,M. and Thomas,M.C. (2007) The L1Tc non-LTR retrotransposon of Trypanosoma cruzi contains an internal RNA-pol II-dependent promoter that strongly activates gene transcription and generates unspliced transcripts. *Nucleic Acids Res.*, **35**, 2199–2214.
6. Luan,D.D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, **72**, 595–605.
7. Martin,F., Maranon,C., Olivares,M., Alonso,C. and Lopez,M.C. (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from Trypanosoma cruzi: homology of the first ORF with the ape family of DNA repair enzymes. *J. Mol. Biol.*, **247**, 49–59.
8. Olivares,M., Alonso,C. and Lopez,M.C. (1997) The open reading frame 1 of the L1Tc retrotransposon of Trypanosoma cruzi codes for a protein with apurinic-apyrimidinic nuclease activity. *J. Biol. Chem.*, **272**, 25224–25228.
9. Garcia-Perez,J.L., Gonzalez,C.I., Thomas,M.C., Olivares,M. and Lopez,M.C. (2003) Characterization of reverse transcriptase activity of the L1Tc retroelement from Trypanosoma cruzi. *Cell. Mol. Life Sci.*, **60**, 2692–2701.
10. Olivares,M., Garcia-Perez,J.L., Thomas,M.C., Heras,S.R. and Lopez,M.C. (2002) The non-LTR (long terminal repeat) retrotransposon L1Tc from Trypanosoma cruzi codes for a protein with RNase H activity. *J. Biol. Chem.*, **277**, 28025–28030.
11. Heras,S.R., Lopez,M.C., Garcia-Perez,J.L., Martin,S.L. and Thomas,M.C. (2005) The L1Tc C-terminal domain from Trypanosoma cruzi non-long terminal repeat retrotransposon codes for a protein that bears two C2H2 zinc finger motifs and is endowed with nucleic acid chaperone activity. *Mol. Cell. Biol.*, **25**, 9209–9220.
12. Heras,S.R., Thomas,M.C., Macias,F., Patarroyo,M.E., Alonso,C. and Lopez,M.C. (2009) Nucleic-acid-binding properties of the C2-L1Tc nucleic acid chaperone encoded by L1Tc retrotransposon. *Biochem. J.*, **424**, 479–490.
13. Heras,S.R., Thomas,M.C., Garcia-Canadas,M., de Felipe,P., Garcia-Perez,J.L., Ryan,M.D. and Lopez,M.C. (2006) L1Tc non-LTR retrotransposons from Trypanosoma cruzi contain a functional viral-like self-cleaving 2A sequence in frame with the active proteins they encode. *Cell. Mol. Life Sci.*, **63**, 1449–1460.
14. Olivares,M., del Carmen Thomas,M., Lopez-Barajas,A., Requena,J.M., Garcia-Perez,J.L., Angel,S., Alonso,C. and Lopez,M.C. (2000) Genomic clustering of the Trypanosoma cruzi nonlong terminal L1Tc retrotransposon with defined interspersed repeated DNA elements. *Electrophoresis*, **21**, 2973–2982.
15. Bringaud,F., Bartholomeu,D.C., Blandin,G., Delcher,A., Baltz,T., El-Sayed,N.M. and Ghedin,E. (2006) The Trypanosoma cruzi L1Tc and NARTc non-LTR retrotransposons show relative site specificity for insertion. *Mol. Biol. Evol.*, **23**, 411–420.
16. Bringaud,F., Garcia-Perez,J.L., Heras,S.R., Ghedin,E., El-Sayed,N.M., Andersson,B., Baltz,T. and Lopez,M.C. (2002) Identification of non-autonomous non-LTR retrotransposons in the genome of Trypanosoma cruzi. *Mol. Biochem. Parasitol.*, **124**, 73–78.
17. Bringaud,F., Berriman,M. and Hertz-Fowler,C. (2009) Trypanosomatid genomes contain several subfamilies of ingi-related retroposons. *Eukaryot. Cell*, **8**, 1532–1542.
18. Bringaud,F., Muller,M., Cerqueira,G.C., Smith,M., Rochette,A., El-Sayed,N.M., Papadopoulou,B. and Ghedin,E. (2007) Members of a large retroposon family are determinants of

post-transcriptional gene expression in Leishmania. *PLoS Pathog.*, **3**, 1291–1307.
19. Eickbush,D.G. and Eickbush,T.H. R2 retrotransposons encode a self-cleaving ribozyme for processing from an rRNA cotranscript. *Mol. Cell. Biol.*, **30**, 3142–3150.
20. Symons,R.H. (1992) Small catalytic RNAs. *Annu. Rev. Biochem.*, **61**, 641–671.
21. Cochrane,J.C. and Strobel,S.A. (2008) Catalytic strategies of self-cleaving ribozymes. *Acc. Chem. Res.*, **41**, 1027–1035.
22. Been,M.D. and Wickham,G.S. (1997) Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur. J. Biochem.*, **247**, 741–753.
23. Salehi-Ashtiani,K., Luptak,A., Litovchick,A. and Szostak,J.W. (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science*, **313**, 1788–1792.
24. Webb,C.H., Riccitelli,N.J., Ruminski,D.J. and Luptak,A. (2009) Widespread occurrence of self-cleaving ribozymes. *Science*, **326**, 953.
25. Li,P.W., Li,J., Timmerman,S.L., Krushel,L.A. and Martin,S.L. (2006) The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition. *Nucleic Acids Res.*, **34**, 853–864.
26. Lyons,A.J. and Robertson,H.D. (2003) Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J. Biol. Chem.*, **278**, 26844–26850.
27. Serrano,P., Gomez,J. and Martinez-Salas,E. (2007) Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element. *RNA*, **13**, 849–859.
28. Chadalavada,D.M., Gratton,E.A. and Bevilacqua,P.C. The human HDV-like CPEB3 ribozyme is intrinsically fast-reacting. *Biochemistry*, **49**, 5321–5330.
29. Nakano,S., Chadalavada,D.M. and Bevilacqua,P.C. (2000) General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science*, **287**, 1493–1497.
30. Nur,I., Pascale,E. and Furano,A.V. (1988) The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter. *Nucleic Acids Res.*, **16**, 9233–9251.
31. Zhou,J. and Eickbush,T.H. (2009) The pattern of R2 retrotransposon activity in natural populations of Drosophila simulans reflects the dynamic nature of the rDNA locus. *PLoS Genetics*, **5**, e1000386.
32. Bringaud,F. (2005) [Comparative genomics of trypanosomatid parasitic protozoa]. *Med. Sci.*, **21**, 1027–1028.
33. Ruminski,D.J., Webb,C.H., Riccitelli,N.J. and Luptak,A. (2010) Processing of insect retrotransposons by self-cleaving ribozymes. *Nature Precedings*, Available from Nature Precedings <http://hdl.handle.net/10101/npre.2010.4333.1>.
34. Gruber,A.R., Lorenz,R., Bernhart,S.H., Neubock,R. and Hofacker,I.L. (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36**, W70–W74.
35. Fedoruk-Wyszomirska,A., Giel-Pietraszuk,M., Wyszko,E., Szymanski,M., Ciesiolka,J., Barciszewska,M.Z. and Barciszewski,J. (2009) The mechanism of acidic hydrolysis of esters explains the HDV ribozyme activity. *Mol. Biol. Rep.*, **36**, 1647–1650.
36. Chen,J.H., Gong,B., Bevilacqua,P.C., Carey,P.R. and Golden,B.L. (2009) A catalytic metal ion interacts with the cleavage Site G.U wobble in the HDV ribozyme. *Biochemistry*, **48**, 1498–1507.
37. Chen,J.H., Yajima,R., Chadalavada,D.M., Chase,E., Bevilacqua,P.C. and Golden,B.L. A 1.9 A crystal structure of the HDV ribozyme precleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester cleavage. *Biochemistry*, **49**, 6508–6518.
38. Chadalavada,D.M., Cerrone-Szakal,A.L. and Bevilacqua,P.C. (2007) Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. *RNA*, **13**, 2189–2201.