Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopeda cave crickets

Andres A. Rojas, Alejandro Vazquez-Tello*, Gerardo Ferbeyre, Federica Venanzetti1, Lutz Bachmann2, Bruno Paquin, Valerio Sbordoni1 and Robert Cedergren

Département de Biochimie, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Quebec H3C 3J7, Canada, 1Dipartimento di Biologia, Universita di Roma Tor Vergata, Via Orazio Raimondo, 00173 Roma, Italy and 2Pritzker Laboratory for Molecular Systematics and Evolution, The Field Museum, 1400 South Lake Shore Drive, Chicago, IL 60605-2496, USA

Received May 9, 2000; Revised August 10, 2000; Accepted August 20, 2000

ABSTRACT

This work reports the discovery and functional characterization of catalytically active hammerhead motifs within satellite DNA of the pDo500 family from several Dolichopeda cave cricket species. We show that in vitro transcribed RNA of some members of this satellite DNA family do self-cleave in vitro. This self-cleavage activity is correlated with the efficient in vivo processing of long primary transcripts into monomer-sized RNA. The high sequence conservation of the satellite pDo500 DNA family among genetically isolated Dolichopeda schiavazzii populations, as well as other Dolichopeda species, along with the fact that satellite members are actively transcribed in vivo suggests that the hammerhead-encoding satellite transcripts are under selective pressure, perhaps because they fulfill an important physiological role or function. Remarkably, this is the third example of hammerhead ribozyme structures associated with transcribed repetitive DNA sequences from animals. The possibility that such an association may not be purely coincidental is discussed.

INTRODUCTION

The autocatalytic activity of the RNA structure later known as the ‘hammerhead’ was discovered fortuitously in the small pathogenic satellite RNA of tobacco ringspot virus (sTobRV RNA) (1). The detection of monomeric, both linear and circular, and of multimeric satellite RNAs of both polarities in infected tissues suggested that the satellite RNAs replicate by a rolling circle mechanism and that a highly specific RNase probably catalyzed the processing of long multimers into monomers (2,3). However, the RNase hypothesis was

excluded when Prody et al. showed for the first time that multimeric satellite RNAs actually self-process into monomers in the absence of proteins (1).

These catalytic RNAs adopt a hammerhead-like secondary structure, consisting of three stems of variable sequence and length linked by an asymmetrical single-stranded central core. In contrast to the peripheral stem–loops, all the nucleotides in the core are invariant except the one at position 7 (4–6). Naturally occurring hammerhead ribozymes with demonstrated autocatalytic activity have been identified within a dozen small plant pathogenic RNAs (7–16) and in the transcripts from repetitive Sat2 DNA of salamanders (17) and SMα of schistosomes (18).

We report here the discovery of putative hammerhead ribozyme motifs in the satellite pDo500 DNA family of Dolichopeda cave crickets. In vitro transcribed RNA from satellite members displayed self-cleavage activity, which is correlated with the in vivo processing of satellite multimer transcripts into monomers. This is the third example of catalytically active hammerhead ribozymes associated with transcribed repetitive DNA sequences from animals.

MATERIALS AND METHODS

Searching GenBank for hammerhead ribozyme structural motifs

The computer program RNAMOT (19) was used to search for sequences in GenBank (release of 1997) that could fold into the hammerhead ribozyme structure, as previously described (18). Briefly, the descriptor defining the secondary structure of the hammerhead motif used in this search comprised the 17 nt catalytic core domain, with a variable cleavable triplet (NUH, where H is any nucleotide except guanosine), and the stem–loops I, II and III with unrestricted nucleotide composition. This descriptor should allow the identification of single cis-cleaving hammerhead ribozymes.

*To whom correspondence should be addressed at present address: Hôpital Sainte-Justine, Centre de Recherche, 3175 Cote Sainte-Catherine ave., Montréal, Quebec H3T 1C5, Canada. Tel: +1 514 345 4931; Fax: +1 514 345 4801; Email: vazquez@magellan.umontreal.ca

Present addresses:
Andres A. Rojas, Mycota Biosciences Inc., 225 President-Kennedy Avenue West, Montréal Quebec H2Z 3Y8, Canada
Gerardo Ferbeyre, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
Bruno Paquin, Galileo Genomics, 6700 du Parc avenue, Suite 200, Montréal, Quebec H2V 4H9, Canada

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

Dedicated to the late Robert Cedergren
In vitro transcription

Putative hammerhead motifs were identified in sequences from members of the repetitive satellite pDo500 DNA family from different species of Dolichopoda cave crickets (20). The monomeric units of four members, VetPCR3 and VetPCR4, both isolated from Dolichopoda schiavazzii, Forb-2 from Dolichopoda laetitia and Pst3 from Dolichopoda baccettii, were subcloned into pBluescriptII (SK+) (Stratagene) as XbaI/HindIII fragments. Plasmids were linearized with an appropriate restriction enzyme prior to in vitro transcription (18). Transcripts from VetPCR4 and Pst3 (both 558 nt in length) were synthesized by T7 RNA polymerase in the presence of [γ-32P]dCTP and α-restriction enzyme prior to HinIII were subcloned into pBluescriptII (SK+) (Stratagene) as XbaI/HindIII fragments. Plasmids were linearized with an appropriate restriction enzyme prior to in vitro transcription (18). Transcripts from VetPCR4 and Pst3 (both 558 nt in length) were synthesized by T7 RNA polymerase in the presence of [γ-32P]dCTP and α-restriction enzyme prior to HinIII were subcloned into pBluescriptII (SK+) (Stratagene) as XbaI/HindIII fragments. Plasmids were linearized with an appropriate restriction enzyme prior to in vitro transcription (18). Transcripts from VetPCR4 and Pst3 (both 558 nt in length) were synthesized by T7 RNA polymerase in the presence of [γ-32P]dCTP and α-restriction enzyme prior to HinIII were subcloned into pBluescriptII (SK+) (Stratagene) as XbaI/HindIII fragments. Plasmids were linearized with an appropriate restriction enzyme prior to in vitro transcription (18).

Cleavage assays and kinetics

To determine the effect of incubation temperature on the rate of self-cleavage activity (kobs), gel-purified radiolabeled RNA molecules (0.5 µg/ml) were denatured at 95°C for 2 min and then cooled slowly at room temperature. They were equilibrated at various temperatures in 10 mM Tris–HCl (pH 8.0) for 2 min and self-cleavage reactions were initiated by adding 5 mM MgCl2. Aliquots (5 ng RNA) were taken at 30–60 s intervals and the reaction was stopped immediately by adding 2 vol of loading buffer (90% formamide, 30 mM EDTA, 1 M urea, 0.1% each of xylene cyanol and bromophenol blue) and resolved on denaturing 6% polyacrylamide gels containing 8 M urea. The radiolabeled RNAs were eluted from gel slices overnight at 37°C in buffer containing 0.4 M ammonium acetate and 0.1% SDS, then extracted with phenol:chloroform and recovered by ethanol precipitation.

Cricket RNA purification and northern hybridization

Total RNA was extracted from adult, male or female Dolichopoda cave crickets (collected in Mount Argentario, Tuscany, Italy) by the acid–guanidinium thiocyanate–phenol–chloroform method (22). Typically, 1–20 µg of total RNA along with 1–100 ng of in vitro transcribed RNA from the repetitive satellite Pst3 clone were separated on a denaturing 12.5% agarose–formaldehyde gel and transferred to nylon membranes (Hybond-N, Amersham). The full-length monomeric Pst3 clone was radiolabeled with [γ-32P]dCTP by random priming and T7 DNA polymerase (Pharmacia) and used as a probe. The oligonucleotide 5′-TACAATAAAGGAAGGACA-AACACGATC-3′, complementary to the 3′ end of Pst3 RNA but 89% identical to the 3′ end of other satellite members, was phosphorylated with [γ-32P]ATP and polynucleotide kinase (Pharmacia) and used as a probe. Solutions, hybridization temperature and membrane washes were done as described (23). After autoradiography, the integrated intensity of individual bands was quantified by densitometry with the computer program Scion ImagePC (Scion Corp.).

RESULTS

A hammerhead ribozyme structure in repetitive satellite DNA

Using the computer program RNAMOT (19), we searched for hammerhead motifs in the GenBank sequence database. We identified 13 sequences from the tandemly repeated satellite pDo500 DNA family of Dolichopoda cave crickets (20), which predictably fold into hammerhead structures with the triplet CUA upstream of the cleavage site. The hammerhead domain is highly conserved in these 13 sequences whereas eight additional satellite clones contain nucleotide substitutions expected to inactivate the catalytic structure (Fig. 1A). We concentrated our efforts on four clones representing three Dolichopoda species, Pst3 (D. baccettii), Forb-2 (D. laetitia), VetPCR3 and VetPCR4 (both from D. schiavazzii). They were chosen because they encompass several nucleotide substitutions occurring within the hammerhead domain. None of the isolates numbered 3–11 from D. schiavazzii with identical hammerhead domain sequences were chosen for the analysis because the most significant substitution, the U–U mismatch in the stem–loop III is also represented in the Forb-2 isolate.

The folding of the complete monomeric sequences was assessed using the computer program MFOLD (version 3.0) (24). Figure 1B shows the anticipated Pst3 hammerhead structure and the nucleotide variations among these repetitive sequences. The Pst3 hammerhead domain, presenting a short leader of 4038 nt from the vector. Synthesis reactions were incubated at 37°C for 1 h and stopped by adding an equal volume of 20 mM EDTA, extracted with phenol:chloroform and precipitated with ethanol. Full-length in vitro transcribed RNAs were purified on denaturing 6% polyacrylamide gels containing 8 M urea. The radiolabeled RNAs were eluted from gel slices overnight at 37°C in buffer containing 0.4 M ammonium acetate and 0.1% SDS, then extracted with phenol:chloroform and recovered by ethanol precipitation.

Cleavage assays and kinetics

To determine the effect of incubation temperature on the rate of self-cleavage activity (kobs), gel-purified radiolabeled RNA molecules (0.5 µg/ml) were denatured at 95°C for 2 min and then cooled slowly at room temperature. They were equilibrated at various temperatures in 10 mM Tris–HCl (pH 8.0) for 2 min and self-cleavage reactions were initiated by adding 5 mM MgCl2. Aliquots (5 ng RNA) were taken at 30–60 s intervals and the reaction was stopped immediately by adding 2 vol of loading buffer (90% formamide, 30 mM EDTA, 1 M urea, 0.1% each of xylene cyanol and bromophenol blue) and resolved on denaturing 6% polyacrylamide–TBE gels containing 8 M urea. The intact RNA and cleavage products were cut out of the gel and counted by liquid scintillation. To calculate kobs values for each incubation temperature, the logarithm of disappearing substrate (%) was plotted versus the incubation time (min) and the slope of the straight line was calculated with the computer program SigmaPlot 4.01 (21) (SPSS Inc.).
obtain the structures shown in Figure 1D and E, respectively. It is unlikely that active hammerhead structures would form from these two sequences, based on their highest free energy ($\Delta G = -8.3$ and $-9.0$ kcal/mol for VetPCR3 and VetPCR4, respectively) with respect to other non-hammerhead, more stable alternative conformations ($\Delta G = -19.0$ and $-12.3$ kcal/mol). In addition, the mismatch U1.1–G2.1 in the stem I of VetPCR3, immediately adjacent to the cleavage triplet CUA, is likely to inhibit self-cleavage (Fig. 1C). The stems II and III of the hammerhead-like structure from VetPCR4 contain only 2 bp, making the structure thermodynamically less stable. This, together with the presence of two inhibitory nucleotide substitutions (A6 $\rightarrow$ G6 and A9 $\rightarrow$ G9) in the catalytic core suggest that this hammerhead structure is inactive (Fig. 1E).

Active hammerhead self-cleavage of repetitive satellite Pst3 and For6-2 RNA in vitro

We sought first to determine whether in vitro transcribed monomer RNAs from four members of the satellite pDo500 family fold into active self-cleaving structures. As expected and according to computer predictions, both Pst3 and For6-2 RNAs displayed self-cleavage activity. The effect of Mg$^{2+}$ concentration on the rate of cleavage was determined. Both Pst3 and For6-2 RNAs showed highest activity with 5 mM MgCl$_2$, independent of the incubation temperature. Higher concentrations led to a decrease in the self-cleavage activity (Fig. 2A and B). For VetPCR3 and VetPCR4 RNAs, self-cleavage activity was undetectable at the conditions tested (results not shown).

The identity of the cleavage site, CUA, was confirmed by primer extension-reverse transcription of the self-cleaved in vitro transcribed Pst3 RNA (Fig. 3A). Mutational analysis gave further support to the autocatalytic hammerhead structure of the Pst3 clone. When we introduced the substitution G5 $\rightarrow$ C5 within the hammerhead core domain, the self-cleavage activity was completely eliminated (Fig. 3B).

The rate of cleavage ($k_{\text{obs}}$) of the Pst3 hammerhead was much higher than that of For6-2. The temperature of incubation had an important effect on $k_{\text{obs}}$, being highest in the range of 50–55°C ($0.26$ min$^{-1}$) (Fig. 4). Above this temperature, the sharp decrease in self-cleavage activity suggests that the hammerhead structure is destabilized or denatured, perhaps accentuated by the G–U mismatch present in the stem–loop III. The much lower rate of cleavage observed for the monomeric For6-2 is also in agreement with a predicted thermodynamically less stable catalytic structure, based on the short (2 bp) stem–loop III. The maximum $k_{\text{obs}}$ was observed at $37^\circ$C ($0.03$ min$^{-1}$), then decreased in parallel to the increase of temperature (Fig. 4). A weak self-cleavage activity could be detected at $50^\circ$C but not at higher temperatures (Figs 2 and 4).

Figure 1. Satellite pDo500 family members from the cave cricket Dolichopoda coding for self-cleaving transcripts. (A) Alignment of the hammerhead domain from 13 different satellite members (numbered from 1 to 13) that were predicted to fold into active hammerhead ribozymes with the computer program RNAMOT. Eight additional satellite isolates (numbered from 14 to 21) that contain nucleotide substitutions (shown in bold and underlined) probably fold into inactive conformations. Satellite clones numbered 3–21 were isolated from $D.\text{schiauvazii}$ whereas Pst3 and For6-2 clones were isolated from $D.\text{buccetti}$ and $D.\text{laetitia}$, respectively (20). (B) Hammerhead structural motif predicted with the computer program MFOLD V3.0 from the Pst3 isolate. Helix III contains the non-canonical Watson–Crick G–U base pair. The cleavage site is denoted by CS. The mutation G5 $\rightarrow$ C5 was artificially introduced for functional studies. Additional substitutions (indicated by arrows directed out of the structure) and insertions (indicated by arrows directed toward the structure) found in three other isolates are shown. (C) For6-2 isolate. The short helix III, with only 2 bp, is predicted to be destabilized at high incubation temperatures. (D) VetPCR3 isolate. The mismatch U1.1–G2.1, which is the principal difference with respect to For6-2 clone, is predicted to inhibit self-cleavage. (E) VetPCR4 isolate. Two substitutions within the conserved catalytic core, A6 $\rightarrow$ G6 and A9 $\rightarrow$ G9, predictably abolish its self-cleavage activity.
The satellite pDo500 family is transcribed in vivo

In order to determine whether these repeats are transcribed in vivo, we performed a northern analysis. We used the full-length monomer Pst3 DNA to probe total RNA isolated from either male or female adult D.baccettii cave crickets. We detected a single uniform transcript of ∼500 nt, similar in size to that of the in vitro transcribed monomer Pst3 (Fig. 5A). Putative multimer Pst3 transcripts of heterogeneous size are also visible when the autoradiography is over exposed. The detected monomeric Pst3 transcript was equally abundant in the total RNA samples from adults of both gender. Quantitation analysis indicated that the monomer Pst3 RNA is not an abundant transcript, representing ∼0.07% of the total RNA. Additional northern blots with an oligonucleotide probe complementary to the 3′ end of the Pst3 transcript confirmed these results (Fig. 5B).

DISCUSSION

In this work we report the discovery and the in vitro characterization of active self-cleaving hammerhead structures within the satellite pDo500 DNA family from Dolichopoda cave crickets. Our work corroborates the feasibility of the ribonomics approach, defined by the search, distribution analysis and functional characterization of various RNA motifs in sequence databases (25) and recently illustrated by the discovery of a hammerhead within the repetitive DNA of schistosomes (18). Major questions raised concerned the relevance of the 13 positive hits obtained by in silico prediction of hammerhead structures within the repetitive sequences from Dolichopoda. As predicted, we showed that the complex RNA sequences from two representative satellite family members fold into an active self-cleaving hammerhead structure in vitro (Fig. 2). Further, since the observed hits are conserved among genetically related species (20), we could infer that these structures have resisted random drift divergence, possibly because they fulfill an important physiological role or function.
The issue of primordial importance was to determine whether these structural motifs are actively transcribed. Northern analysis demonstrated that the catalytic hammerhead-encoding repetitive sequences are indeed transcribed \textit{in vivo}, despite the fact that no obvious promoter elements could be identified (Fig. 5). Considering that this repetitive satellite family represents $\sim 5\%$ of the total genomic DNA and that various members (unit-copy sizes ranging from 420 to 510 bp in length) are arranged in multiple tandem repeats with variable unit-copy numbers (20), we expected to detect discrete multimeric transcripts. Surprisingly, only a major discrete band corresponding to monomeric-sized transcripts was detected. This is compatible with the hypothesis of a hammerhead-mediated self-processing that produces precisely defined 5$'$ and 3$'$ termini. Discrete transcripts corresponding to multimers being processed were not detected perhaps due to their very low abundance or because they are indistinguishable from the heterogeneous population of transcripts with variable number of tandemly repeated units (20). This situation corroborates results observed with the SINE retroposon-like SM$\alpha$ from schistosomes (18) but contrasts with the hammerhead-encoding repetitive sat2 from the salamanders for which both monomer and discrete multimer transcripts are detected \textit{in vivo} (17).

In addition to the single hammerhead \textit{cis}-cleavage mechanism, two tandem copies within the long primary transcript can hybridize to each other because of extensive base complementation and form a double hammerhead structure so that each strand is trans-cleaved (Fig. 6). This hypothetical mechanism may explain the efficient processing of the long primary transcripts \textit{in vivo}, thus producing the uniform monomeric-sized RNA detected by northern blotting.

The physiological relevance of the hammerhead-mediated processing of transcripts from repetitive sequences in organisms as varied as cricket, schistosomes and salamanders is intriguing. These examples prompt us to consider that the association between catalytically active hammerhead motifs and transcribed repetitive DNA sequences could not be purely coincidental. Rather, this suggests that the hammerhead structures exert a common function or alternatively, that they may have arisen by a similar mechanism. Our interpretation is that the hammerhead structure operates as an accessory only to process the long primary transcripts into monomers. However, the possibility that monomeric RNAs could \textit{trans}-cleave host transcripts \textit{in vivo} cannot be rejected.
It is speculated that repetitive DNA sequences evolved from ancient independent genetic elements that propagated within the genome of many organisms (26). At that earlier period, it could be suggested that the catalytic hammerhead played an important role in replication and propagation of these elements, which later became obsolete once they were stably inserted into the host genome and had reached an equilibrium with the host cell. How then does one explain the presence of active hammerhead motifs in the repetitive pDo500 sequences from genetically isolated Dolichopoda cave cricket populations? Indeed, functionless nuclear DNA can be readily eliminated by selection (27). The remarkable conservation of pDo500 sequences in several Dolichopoda species, even among genetically isolated D.schiavazzii populations (20), along with the fact that they are actively transcribed in vivo, strongly suggests that these repetitive sequences are under selection pressure. Our analysis showed that the Dolichopoda satellite pDo500 family, like the SINE SMα family from schistosomes (18) and the sat2 family from salamanders (17,28), consists of members with either active or mutated hammerhead structures (Fig. 1A). A possible explanation may be spontaneous reversion of mutations. This leaves open the possibility that the ribozyme-mediated processing of multimeric transcripts may be somehow beneficial to the host cells. Except for the three repetitive DNA families cited above, so far all characterized animal satellite DNAs and SINEs that are transcribed in vivo lack catalytically active structures (29–33). However, the absence of sequence similarity in the repetitive elements has led to speculation that their corresponding transcripts may adopt common secondary structures, such as stem–loops with double-stranded RNA stretches, or higher order structural motifs, which could determine RNA–protein interactions (34,35).

It is notable that, so far and without exception, all natural RNAs containing sequences folding into an active hammerhead structure do exist as monomeric sized RNA copies. A number of plant pathogenic satellite RNAs such as sTobRV, propagate by a rolling-circle amplification mechanism, in which the circular monomer templates result from the hammerhead-mediated multimeric RNA processing (or by other catalytic structures) into linear monomers and subsequent circularization by a host-specific RNA ligase (12,36,37). The possibility that the hammerhead-mediated processing of repetitive pDo500 sequences constitute the remains of an ancient propagation mechanism deserves further analysis. Although it has not been determined whether the self-cleaved RNA monomers from cricket satellite pDo500 and schistosome SINE SMα are circularized in vivo, the report that the satellite RPCS DNA from the rodent Ctenomys porteous could have propagated by a rolling-circle mechanism is significant (38). Moreover, a model of replication mechanism similar to that of plant pathogenic satellite RNAs is an attractive alternative that could explain both the propagation of repetitive sequences and the function of their associated catalytic structures.

ACKNOWLEDGEMENTS

We thank Alice Rae and Jose Esteban for critical reading of the manuscript, and the members of the laboratory, especially C. Gomez-Aguilera, G. Tremblay, P. Blancafort and A. Ioudovitch for helpful discussions. This work was supported by a grant from the Medical Research Council of Canada.

REFERENCES


Figure 6. Hypothetical mechanism that may explain the efficient processing of the long primary transcripts from tandemly repeated pDo500 Pst3 and For6-2 members in vivo. This model proposes that two copies with the tandemly repeated transcript could hybridize to each other by extensive base pair complementation, resulting in the formation of a double-hammerhead structure so that each opposite strand is trans-cleaved producing monomer RNA copies. Intra-repeat (cis-) cleavage also contributes to the rapid processing. This partial sequence is from the Pst3 isolate. Arrows pointing out from the structure indicate the nucleotide substitutions occurring in the For6-2 isolate. The cleavage sites (CS) are also shown with an arrow.