



RNA

A PUBLICATION OF THE RNA SOCIETY

The ubiquitous hammerhead ribozyme

Christian Hammann, Andrej Luptak, Jonathan Perreault, et al.

RNA 2012 18: 871-885 originally published online March 27, 2012

Access the most recent version at doi:[10.1261/rna.031401.111](https://doi.org/10.1261/rna.031401.111)

References

This article cites 78 articles, 40 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/18/5/871.full.html#ref-list-1>

Email alerting service

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

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

The ubiquitous hammerhead ribozyme

CHRISTIAN HAMMANN,^{1,5,6} ANDREJ LUPTAK,^{2,6} JONATHAN PERREAULT,^{3,6} and MARCOS DE LA PEÑA^{4,6}

¹Heisenberg Research Group Ribogenetics, Technical University of Darmstadt, 64287 Darmstadt, Germany

²Department of Pharmaceutical Sciences, University of California–Irvine, Irvine, California 92697, USA

³Centre INRS – Institut Armand-Frappier, Laval, Québec, H7V 1B7, Canada

⁴Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC), 46022 Valencia, Spain

ABSTRACT

The hammerhead ribozyme is a small catalytic RNA motif capable of endonucleolytic (self-) cleavage. It is composed of a catalytic core of conserved nucleotides flanked by three helices, two of which form essential tertiary interactions for fast self-scission under physiological conditions. Originally discovered in subviral plant pathogens, its presence in several eukaryotic genomes has been reported since. More recently, this catalytic RNA motif has been shown to reside in a large number of genomes. We review the different approaches in discovering these new hammerhead ribozyme sequences and discuss possible biological functions of the genomic motifs.

Keywords: catalytic RNA; database searches; homology; retrotransposons; structure

INTRODUCTION

The hammerhead ribozyme (HHR) (Prody et al. 1986) belongs to the family of small endonucleolytic ribozymes that have sizes in the range of from 50 to 150 nucleotides (nt) (Lilley 2005). As with the other members of this family, the Varkud satellite (VS), the hairpin, and the hepatitis delta ribozymes (Buzayan et al. 1986; Kuo et al. 1988; Saville and Collins 1990), the HHR catalyzes the scission of its own phosphodiester backbone by means of a transesterification reaction that proceeds under inversion of the configuration and, albeit intramolecularly, according to an S_N2 mechanism (Fig. 1A). No cofactor is required for these reactions, although divalent metal ions are known to play varying roles in the processes of RNA folding and self-cleavage of each small endonucleolytic ribozyme (Wilson and Lilley 2009). In contrast, the *glmS* catalytic riboswitch catalyzes the same reaction, but it requires glucosamine-6-phosphate as an essential cofactor (Winkler et al. 2004; Klein and Ferré-D'Amaré 2006).

Formally, the minimal HHR motif is made up of a catalytic core of conserved nucleotides flanked by three helices (Fig. 1B) that are absolutely required for catalysis. Since its introduction (Forster and Symons 1987b; Uhlenbeck 1987), this minimal structure has been extensively studied in vitro (summarized in Hammann and Lilley 2002), but in 2003, additional experiments showed that interacting peripheral loops 1 and 2 of full-length HHRs are also required to reach efficient self-cleavage under physiological conditions (De la Peña et al. 2003; Khvorova et al. 2003). These interactions are required for efficient folding and fast cleavage kinetics (Canny et al. 2004; Penedo et al. 2004), and were elegantly confirmed in a first crystal structure of a natural HHR (Martick and Scott 2006). The discovery that the HHR requires tertiary interactions distal from the core to allow efficient cleavage reconciled biochemical data with the atomic-resolution structure (Przybilski and Hammann 2006; Nelson and Uhlenbeck 2008a) and explained how this ribozyme could function in physiological magnesium concentrations (~1 mM).

Although HHRs with potential to cleave other RNAs in *trans* have been identified (Luzi et al. 1997; Martick et al. 2008), the vast majority of known natural HHRs are *cis*-, i.e., self-cleaving entities that contain the motif within a single RNA strand. This gives rise to the three circularly permuted forms (Fig. 1C) that are named types I, II, or III, according to the open-ended helix that connects the motif with the flanking sequences.

The HHR was discovered originally as type III in subviral plant pathogens, where it is involved in the processing of

⁵Present address: Ribogenetics@Biochemistry Laboratory, Jacobs University Bremen, 28759 Bremen, Germany.

⁶Corresponding authors.

E-mail hammann@bio.tu-darmstadt.de.

E-mail aluptak@uci.edu.

E-mail Jonathan.Perreault@iaf.inrs.ca.

E-mail rivero@ibmcp.upv.es.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.031401.111>.

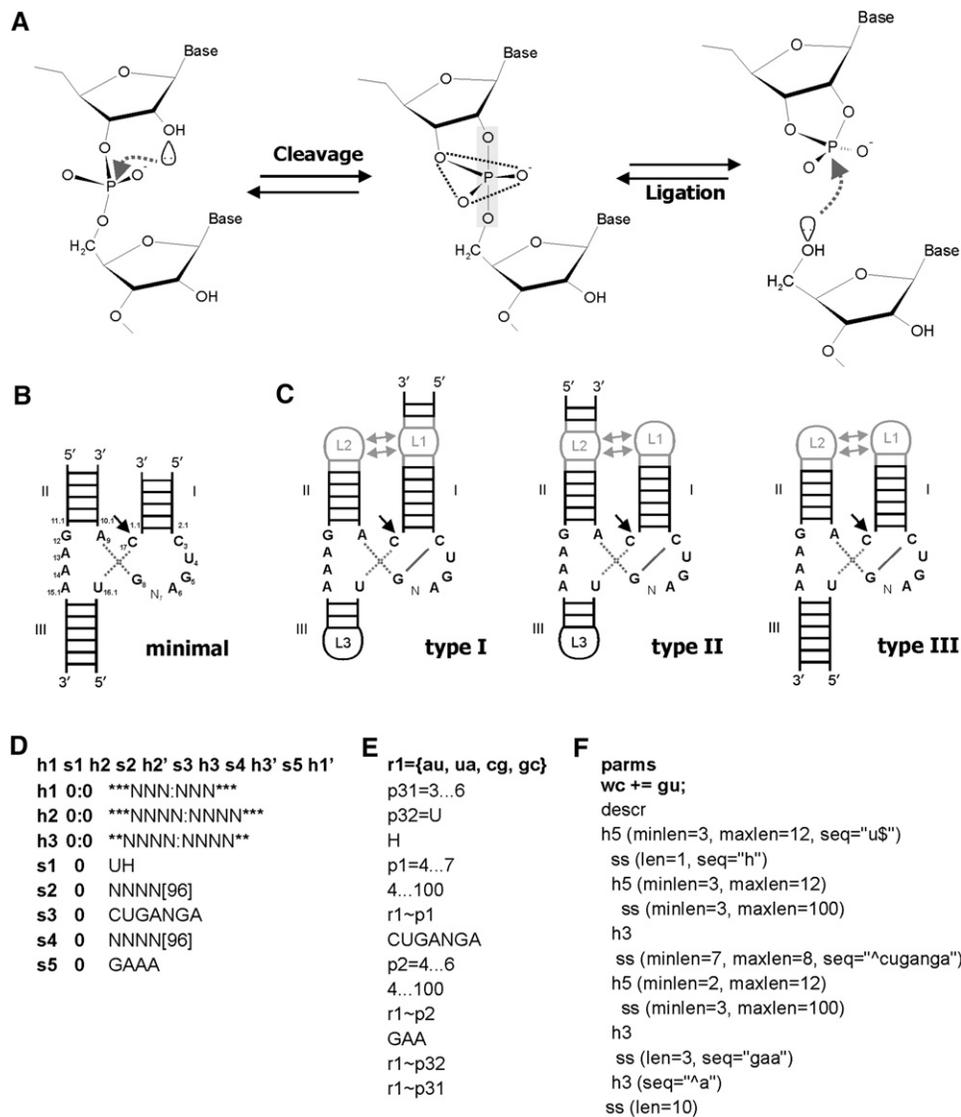


FIGURE 1. The hammerhead ribozyme. (A) The reaction proceeds via an S_N2 mechanism, in which the hydroxy moiety at C2' attacks the neighboring 3'-5' phosphodiester bond. The pentagonal bipyramidal transition state is adopted both in the forward cleavage and in the reverse ligation reaction. The cleavage products are a 2',3'-cyclic phosphate at the 5' product and 5'-hydroxyl at the 3' cleavage product, which serve as substrates for the ligation reaction. (B) The catalytic core of HHR consists of conserved nucleotides (bold) flanked by helices I-III. The conventional numbering system is indicated (Hertel et al. 1992). Cleavage takes place between nucleotides 17 and 1.1, as indicated by the arrow. Dotted lines indicate backbone continuity. (C) Natural forms occur in three types, named after the open stem. Interactions between nucleotides of loops L1 and L2 are observed in all three topologies (double-headed arrows). Additionally, a base pair is formed between nucleotides C₃ and G₈, or variants thereof (see text). Only the three natural topologies are active under physiological magnesium ion conditions, while the minimal format is not. Pattern descriptors for RNABOB (D), PatScan (E), or RNAmotif (F), describing for each a type III motif with the following features: stem III of 3-6 bp, U, H, stem I of 4-7 bp, loop L1 of 4-100 nt, CUGANGA, stem II of 4-6 bp, loop L2 of 4-100 nt GAAA (for details, see D'Souza et al. 1997; Macke et al. 2001; Eddy 2005).

multimeric replication intermediates (Prody et al. 1986). Two different types of subviral plant pathogens harbor HHRs, viroids, and plant virus satellite RNAs, both of which are circular, single-stranded RNA molecules of ~250-400 nt. Viroids are infectious by themselves (Flores et al. 2004; Tabler and Tsagris 2004), while plant virus satellite RNAs depend in their propagation on their cognate virus and modulate the virus' symptoms (Buzayan et al. 1986;

Kaper et al. 1988; Rubino et al. 1990). These two types of subviral plant pathogens are thought to replicate in a rolling circle mechanism (Branch and Robertson 1984), which can proceed in either a symmetric or asymmetric manner (Flores et al. 2004; Tabler and Tsagris 2004).

The first genomic HHRs, of type I, were reported more than two decades ago in *newt*, where they are contained in tandem repeats (~300 bp) in satellite DNA (Epstein and

Gall 1987). Two different HHR-containing transcripts are expressed in the amphibian. They feature different transcription start sites, but have the same length (Epstein and Coats 1991). Similar genomic HHR were also shown for a number of other amphibians (Zhang and Epstein 1996), where they were shown to reside in RNPs, and isolated RNPs were capable to exert cleavage in *trans* on a short synthetic model RNA substrate (Luzi et al. 1997; Denti et al. 2000). Despite these advances, the biological function of the amphibian HHR remains elusive—as the title of an earlier manuscript stated (Cremisi et al. 1992). Similar to these amphibian motifs, genomic HHRs were also reported in satellite DNA of schistosomes and cave crickets (Ferbeyre et al. 1998; Rojas et al. 2000). Together with early HHR motif searches through sequence databases, these discoveries hinted at a broad distribution of these ribozymes in nature, despite a low theoretical chance of occurrence (Ferbeyre et al. 2000).

An unusual case of a genomically incorporated HHR was found in carnation, where the ribozyme exists in the form of a retroviroid element in DNA tandem repeats, similar to the newt (Daros and Flores 1995). Unlike in newts, however, the carnation self-cleaving HHR motifs are found in both DNA strands. Since this is analogous to the subviral plant pathogens, where HHR motifs can be observed in the plus and minus RNA strand, this result points toward a genomic integration of a viroid in carnation.

Given the seemingly sporadic appearance of the HHRs in various genomes, we have set out in the past several years to investigate the prevalence of this catalytic RNA motif. In a pilot study, the presence of two individual HHRs was reported in the genome of *Arabidopsis thaliana* (Przybilski et al. 2005) in addition to a few other genomes (Gräf et al. 2005). The *A. thaliana* motifs are distinct from the aforementioned genomic HHRs, as they do not reside in satellite DNA, but in two individual loci on chromosome IV. Subsequently, a split version of the HHR was uncovered in mammalian genomes and this motif has the potential to cleave in *trans* (Martick et al. 2008). Searches in large data sets carried out since have revealed a surprisingly high number of genomic locations that feature HHRs, and also led to the discovery of the first instances of the type II topology (de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011). In this review, we will summarize our different approaches and results, and discuss possible functions of the various genomic HHRs.

DIFFERENT APPROACHES FOR THE IDENTIFICATION OF GENOMIC HHRs

Small RNA domains are known to show high informational content, a feature highlighted by this family of endonucleolytic ribozymes. The small size of these motifs, together with their low-sequence conservation makes any bioinformatic detection a challenging task. This results from the

nature of the HHR that requires not only the catalytic core, but also helical elements in stems I–III, as well as interacting nucleotides within loops I and II (De la Peña et al. 2003; Khvorova et al. 2003). Thus, a full HHR features varying levels of conservation: high conservation of nucleotides in the catalytic core (see below); sequence covariation, but not sequence identity, as dictated by Watson-Crick base-pairing within helices I–III; and positional and sequence restriction of loop I and II nucleotides involved in tertiary interactions. The latter feature a weak degree of sequence conservation (Martick and Scott 2006; Chi et al. 2008; Dufour et al. 2009), which might be explained by the observation that non-Watson-Crick interactions (Leontis et al. 2002) can also contribute to the stabilization of the HHR (Przybilski et al. 2005). Because of the isostericity of many non-Watson-Crick base pairs (Leontis et al. 2002), the sequence conservation of interaction patterns might be blurred to some degree. To address these features of the HHR motif in the process of identifying novel examples, we have applied two different approaches: sequence homology-based and structure-based searches, as detailed next.

FINDING THE TREASURE BY LOOKING FOR SMALL PIECES: SEQUENCE HOMOLOGY-BASED SEARCHES

This approach is mainly based on the observation that not only the conserved core, but also tertiary interactions between loops 1 and 2 (Fig. 1B,C) are somewhat conserved in most natural HHRs (Martick and Scott 2006; Chi et al. 2008; Dufour et al. 2009) and could be used as an anchoring point for the identification of these ribozymes. All of these constraints were recently used to follow a new HHR search strategy (de la Peña and Garcia-Robles 2010a,b), which first looked for short sequence seeds of 20–30 nt derived from natural HHR sequences. A seed corresponded to a single helix (usually helix II sequence) flanked by conserved catalytic boxes (Fig. 1C) that were used for BLAST homology searches of sequence databases. Identified hits were initially filtered in order to preserve (1) the stem II, and (2) the two conserved boxes of canonical HHRs. Subsequently, 5' and 3' ends of the positive hits were extended and examined for the presence of the HHR remnants: (3) helix I, (4) helix III, (5) self-cleavage site, and (6) potential tertiary interactions between loops 2 and 1. New BLAST searches were carried out recursively, such that seeds derived from HHRs identified in previous rounds of searches were used as new queries in subsequent rounds. Interestingly, hits that already satisfied the two first criteria (CUGANGA-helix II-GAAA) frequently fulfilled the other four criteria not included in the search seeds. The conjuncture of six unrelated traits together with the HHR occurrence in similar genomic locations (i.e., short tandem repeats) or their high degree of phylogenetic conservation (i.e., amniota HHRs, see below) led to the

conclusion that the detected motifs correspond to bona fide HHRs.

STRUCTURE-BASED SEARCH APPROACHES

As a complementary approach to the sequence homology-based searches, several structure-based computational screens were applied to find novel HHRs (Ferbeyre et al. 1998; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011). In these, different software packages such as RNAMotif (Macke et al. 2001), PatScan (D'Souza et al. 1997), or RNABOB (Eddy 2005; Riccitelli and Luptak 2010) were used to describe in detail both the consensus sequence and the stems of HHRs (Fig. 1C). Resulting descriptors (Fig. 1D–F) were then applied for screening sequences in public databases for the minimal motif. The largely automated searches resulted in numerous primary hits, which, however, did not yet consider the interactions between loops 1 and 2 (Ferbeyre et al. 1998; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011) due to the aforementioned sequence flexibility of these interactions. Subsequent steps were necessary to filter out active ribozymes. In contrast, in the sequence homology-based searches, the iterative steps are an integral part of the search procedure.

In a search for type III motifs, free energies of folding were computed using MFold (Zuker 2003) to discriminate between motifs that are more-or-less likely to adopt the HHR fold (Seehafer et al. 2011). In this process, which was implemented in a program pipeline written in Perl, more than 60,000 primary hits were subjected to two folding procedures, one fully unrestricted and one restricted to adopt the hammerhead fold in order to apply thermodynamic selection criteria, as also applied by others (Reeder and Giegerich 2009). The difference between the resulting folding energies was then used to identify true positive HHR motifs among the primary hits. In vitro cleavage kinetics (Kalweit et al. 2012) served to verify the chosen energy parameters on the example of newly identified motifs from *Xenopus tropicalis* (Seehafer et al. 2011).

Finally, the RNA structure and sequence searching software Infernal (Nawrocki et al. 2009) also revealed many novel HHRs (Perreault et al. 2011). This algorithm uses an alignment in Stockholm format (sequence alignment with secondary structure annotation). It has the advantage of considering non-consensus sequence identities in the covariance model, thus taking into account all base pairs and conserved positions, whether they are part of the catalytic

core or not. In addition, it will also score hits that do not conform perfectly to the consensus core if most of the HHR requirements are fulfilled, making it possible to find natural variants of HHRs (Perreault et al. 2011).

NOVEL HHRs IN ALL KINGDOMS OF LIFE

Our various approaches outlined above resulted in an unprecedented number of HHR motifs that were found in all three kingdoms of life, including viruses or subviral entities (de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011). There is a significant overlap between our results, as the majority of conventional motifs was identified by more than one of our approaches. If differences are observed, they are either due to the used sequence sources, or, alternatively, the idiosyncracies of the different approaches and pattern descriptors. We are currently in the process of setting up a web-based data resource for the new HHR motifs that we uncovered. In the following, selected novel motifs will be summarized according to their genomic location.

Prokaryotes

Initial searches among metagenomic and bacterial sequences gave around 100 HHRs of both types I and III from bacterial species (Table 1), next to a single occurrence in the genome of the archaeon *Cenarchaeum symbiosum* A (Perreault et al. 2011). The bacterial motifs mapped to intergenic regions and were associated with bacteriophage-related sequences (de la Peña and Garcia-Robles 2010b;

TABLE 1. Occurrence of hammerhead ribozymes in prokaryotic genomes

Group	Number of species	Experimentally verified
Acidobacteria	1 sp.	<i>Solibacter usitatus</i>
Actinobacteria	4 sp.	
Cyanobacteria	3 sp.	
Chloroflexi	1 sp.	
Firmicutes	9 Bacillus sp. 9 Clostridium sp.	<i>Desulfotomaculum reducens</i> , <i>Clostridium scindens</i>
Alphaproteobacteria	18 others 6 sp.	<i>Agrobacterium tumefaciens</i> , <i>Azorhizobium caulinodans</i>
Betaproteobacteria	12 sp.	
Deltaproteobacteria	9 Desulfovibrio sp.	
Gammaproteobacteria	8 Legionella sp. 6 Shewanella sp. 11 Vibrio sp. 9 others	
Archaea	1 sp.	<i>Cenarchaeum symbiosum</i>

Note: based on studies by de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011.

Perreault et al. 2011), as exemplified in Figure 2A. More recent results have unveiled an extensive occurrence of HHR motifs among metagenomic data from very diverse origins (marine, freshwater, soil, epibionts, microbiomes, etc.), indicative of a widespread presence of the HHR in viruses, prokaryotes, archaea, and, more specifically, their bacteriophages (Jimenez et al. 2011; Perreault et al. 2011; I Garcia-Robles, R Jimenez, A Luptak, and M de la Peña, unpubl.).

Eukaryotes

Our searches of HHRs were especially fruitful among eukaryotic genomes (Table 2). The genomic localization of these motifs varies strongly, both in number and genetic

context. For example, the genome of the elephant *Loxodonta africana* features an orphan type III HHR motif, whereas the genome of the frog *Xenopus tropicalis* has several hundreds of type I and a few type III motifs, and the genome of the blood fluke *Schistosoma mansoni* contains thousands of type I and several type III motifs (de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011). The genetic contexts of these motifs are frequently tandem repeats, intergenic regions or introns, but they appear only rarely within coding sequences (see below).

In oomycetes (unicellular protists, *Chromalveolata* kingdom), for example, a new family of atypical type I HHRs was found (de la Peña and Garcia-Robles 2010b). The motifs showed unusually long and short stems II and III,

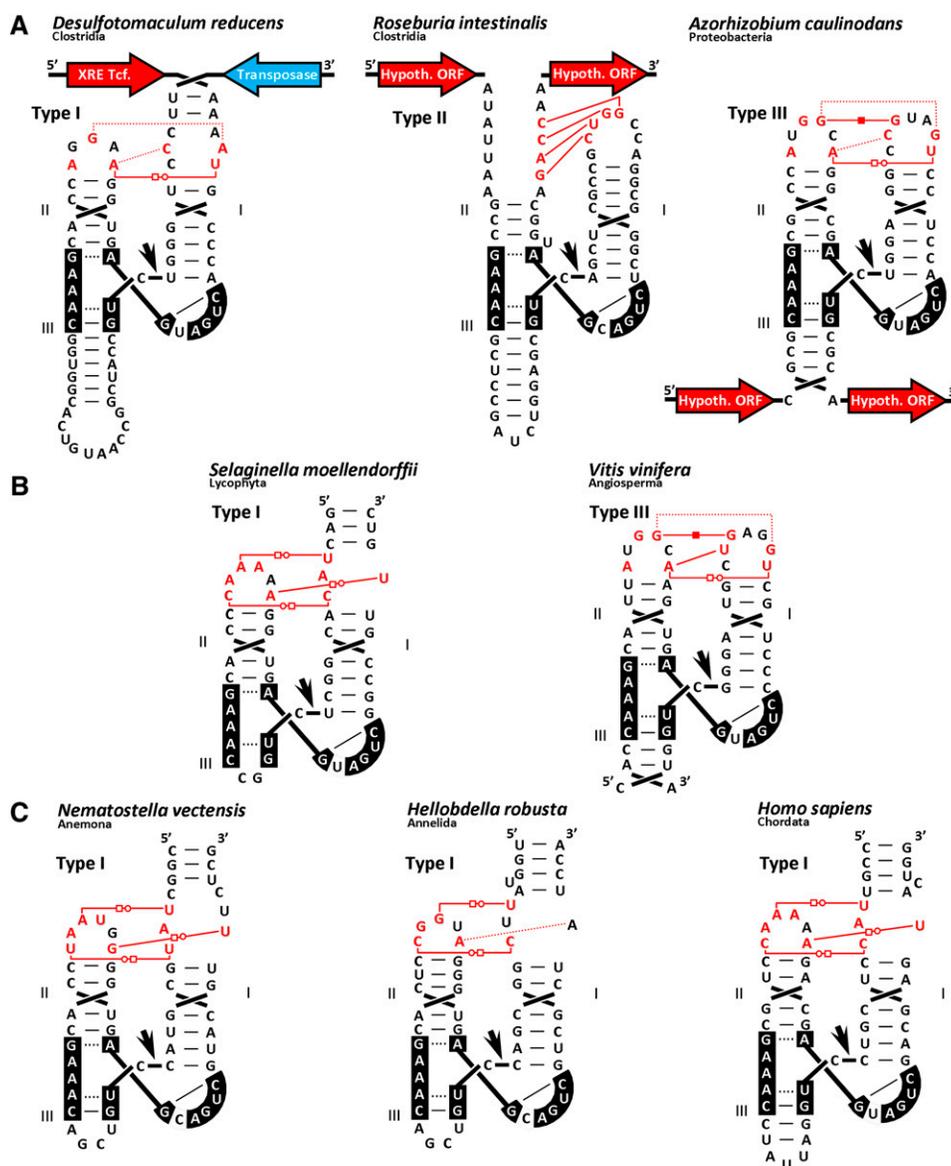


FIGURE 2. Some examples of genomic HHRs detected in bacteria (A), plants (B), and metazoans (C). Main tertiary contacts based on previous experimental models (Martick and Scott 2006; Chi et al. 2008; Dufour et al. 2009), and the conserved nucleotides involved in the interactions are tentatively depicted in red.

TABLE 2. Occurrence of hammerhead ribozymes in eukaryal genomes

Amoebozoa		<i>Entamoeba histolytica</i> HM-1:IMSS <i>Physarum polycephalum</i>
Chromalveolata		2 <i>Hyaloperonospora</i> sp. 2 <i>Phytophthora</i> sp.
Euglenozoa		<i>Trypanosoma cruzi</i> strain CL Brener
Fungi		<i>Ajellomyces capsulatus</i> Nam1 2 <i>Aspergillus</i> sp. <i>Laccaria bicolor</i> S238N-H82 <i>Neosartorya fischeri</i> NRRL 181 <i>Pichia stipitis</i> <i>Postia placenta</i> Mad-698-R <i>Trypanosoma cruzi</i> strain CL Brener <i>Yarrowia lipolytica</i> CLIB122 ^a
Metazoa	Rotifera	<i>Adineta vaga</i> <i>Philodina roseola</i>
	Arthropoda	<i>Acanthoscurria gomesiana</i> 2 <i>Aedes</i> sp. ^a <i>Anopheles stephensi</i> <i>Apis mellifera</i> <i>Argas monolakensis</i> <i>Culex quinquefasciatus</i> <i>Diaphorina citri</i> 11 Dolichopoda sp. ^a 2 <i>Drosophila</i> sp. ^a <i>Ixodes scapularis</i> <i>Nasonia vitripennis</i> <i>Penaeus monodon</i> <i>Reticulitermes flavipes</i> <i>Tribolium castaneum</i>
	Mollusca	<i>Chlamys farreri</i> <i>Doryteuthis pealeii</i> <i>Euprymna scolopes</i> <i>Idiosepius paradoxus</i> <i>Lottia gigantea</i> <i>Pinctada martensii</i>
	Chordata	<i>Branchiostoma floridae</i> <i>Ciona intestinalis</i> <i>Xenopus tropicalis</i> ^a
	Vertebrata (fish)	<i>Danio rerio</i> ^a <i>Labeotropheus fuelleborni</i> <i>Tetraodon nigroviridis</i>
	Vertebrata (sauropsids)	<i>Anolis carolinensis</i> ^a <i>Gallus gallus</i> <i>Meleagris gallopavo</i> <i>Taeniopygia guttata</i>
	Vertebrata (mammals)	<i>Bos taurus</i> <i>Canis lupus familiaris</i> <i>Cavia porcellus</i> <i>Choloepus hoffmanni</i> <i>Dasyurus novemcinctus</i> <i>Echinops telfairi</i> <i>Equus caballus</i> <i>Homo sapiens</i> ^a <i>Loxodonta africana</i> <i>Macaca mulatta</i> <i>Macropus eugenii</i> <i>Microcebus murinus</i> <i>Monodelphis domestica</i> ^a <i>Mus musculus</i> ^a <i>Myotis lucifugus</i> <i>Ochotona princeps</i>

(continued)

respectively, and occurred within 700–1000-bp tandem repeats (Table 3). Similar HHRs have also been found in related plant pathogens (*Hyaloperonospora arabidopsidis* and *Pythium ultimum*) and other protists (slime mold *Physarum polycephalum*), suggesting a wider occurrence of the HHR motif among these eukaryotic microorganisms. More typical type I HHRs were also detected in the genomes of the unicellular *Chlamydomonas reinhardtii* and the related algae *Volvox carteri* (M de la Peña, unpubl.). Although many of these motifs were found associated with larger (2–3 kb) tandem repeats, their particular occurrence within intronic regions suggest the possibility of conserved roles other than retrotransposition (see below).

Genomic sequences of plants (from a spike moss to about 40 angiosperms, both monocots and dicots) showed examples of both types I and III HHRs (Fig. 2B). The motifs usually mapped within short tandem repeats of 300–900 bp (Table 3) at intergenic or repetitive DNA regions (i.e., telomeric repeats in the spike moss *Sellaginella moellenhoffii*). Homology-based searches among EST databases reinforced the in vivo activity of some HHRs with examples of putatively self-cleaved sequences (de la Peña and Garcia-Robles 2010b).

Among metazoans, the vast majority of detected cases corresponded to type I HHRs, while type III motifs appeared to be the exception. Ribozyme motifs were found widespread in the genomes of most metazoan phyla, from simple cnidarians (sea anemones and corals) and poriferans (sea sponge), to more complex organisms like platyhelminthes (trematodes and flat worms), rotifers, annelids (leech and marine polychaetes), arthropods (beetles, mosquitos, or arachnids among many others), and molluscs (scallops, oysters, and squids). Again, some of the detected HHRs mapped within tandem repeats (from 170 to 510 bp) (see Table 3) located in highly repetitive regions (i.e., telomeric sequences in rotifers), although their incidence as single motifs along the genomes was more frequent (Fig. 2C).

TABLE 2. Continued

		<i>Oryctolagus cuniculus</i>
		<i>Otolemur garnettii</i>
		<i>Pan troglodytes</i>
		<i>Rattus norvegicus</i>
		<i>Sciurognathi Muroidea</i>
		<i>Sorex araneus</i>
		<i>Spermophilus tridecemlineatus</i>
		<i>Sus scrofa</i> ^a
		<i>Tarsius syrichta</i>
		<i>Tursiops truncatus</i>
		<i>Vicugna pacos</i>
	Cnidaria	3 <i>Acropora</i> sp.
		<i>Anemonia viridis</i>
		<i>Carukia barnesi</i>
		<i>Hydra magnipapillata</i> ^a
		<i>Hydractinia symbiolongicarpus</i>
		<i>Montastraea faveolata</i>
		<i>Nematostella vectensis</i>
		<i>Porites astreoides</i>
		<i>Strongylocentrotus purpuratus</i>
	Porifera	<i>Sycon raphanus</i>
	Nematoda	2 <i>Caenorhabditis</i> sp.
	Platyhelminthes	<i>Clonorchis sinensis</i>
		<i>Opisthorchis viverrini</i>
		4 <i>Schistosoma</i> sp. ^a
		<i>Schistosomatium douthitti</i>
		<i>Schmidtea mediterranea</i>
Viridiplantae	Chlorophyta	<i>Chlamydomonas reinhardtii</i>
		<i>Volvox carteri</i>
	Plantae	2 <i>Arabidopsis</i> sp. ^a
		<i>Artemisia annua</i>
		3 <i>Citrus</i> sp.
		<i>Daucus carota</i>
		<i>Dianthus caryophyllus</i>
		<i>Fragaria x ananassa</i>
		<i>Guizotia abyssinica</i>
		2 <i>Helianthus</i> sp.
		<i>Jatropha curcas</i> ^a
		2 <i>Lactuca</i> sp.
		<i>Manihot esculenta</i>
		<i>Mimulus guttatus</i>
		<i>Morus australis</i>
		<i>Nicotiana tabacum</i>
		2 <i>Oryza sativa</i> sp.
		<i>Physcomitrella patens</i>
		<i>Sellaginella moellenforfii</i>
		<i>Striga asiatica</i>
		<i>Vitis vinifera</i> ^a
		<i>Zinnia elegans</i>

Note: based on studies by de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011.

^aSome of the corresponding HHRs were experimentally verified.

Genomic sequences from chordates also revealed similar type I HHRs. In the genomes of the *Xenopus tropicalis* frog or marine lampreys (i.e., *Petromyzon marinus*) some examples of HHRs were detected within tandem repeats, similar to the satellite DNA in amphibians (Epstein and Gall 1987). Most HHRs, however, were found widespread as single type I or type III motifs in either intergenic or

intronic (either strand) regions. In contrast, only a few, but highly conserved HHRs were found in amniotes (reptiles, birds, and mammals), all mapping to the sense strands of large introns of three different genes (de la Peña and Garcia-Robles 2010a). The HHRs shared clear sequence and structural similarity with those described in other metazoans, notably trematodes or *X. tropicalis*, with the exception of a larger stem III in the amniota motifs (from 2 to 14 more base pairs) (Fig. 2C). This longer stem III resulted in a robust in vitro self-cleavage activity for one of the human HHRs, with a $k_{\text{obs}} = 2.4 \text{ min}^{-1}$ versus a level of activity that is barely greater than background for the *X. tropicalis* HHR under low Mg^{2+} conditions ($\sim 10^{-4}$ – 10^{-5} min^{-1}) (Garrett et al. 1996). In vivo self-cleavage of one of the amniota HHRs was supported by mammalian ESTs, which intriguingly showed RNA fusions of a cleaved intron with U5 or U6 snRNAs.

Finally, viral entities displayed a rather limited occurrence of the HHR motif. Nevertheless, the already known motifs from virus satellite RNAs and certain Avsunviroidae can be considered as a positive control and were found in all of our approaches (de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011). Next to these, the HHR was found in a small set of dsDNA viruses (Table 4) and in many prophages.

DISCOVERY OF TYPE II HHRs

Soon after the discovery and elucidation of the secondary structure of HHRs (Forster and Symons 1987b), a viral sequence from the tobacco ringspot virus was engineered into a *trans*-cleaving ribozyme by opening the stem-loop I (Haseloff and Gerlach 1988). This experiment formally demonstrated that,

in principle, *cis*-cleaving ribozymes could assume any of the three topologies resulting from embedding the catalytic core within the surrounding transcripts through any of its three helices. All early HHRs discovered were of types I and III, and large-scale sequence-based searches were designed to look for these elements (de la Peña and Garcia-Robles 2010a,b; Seehafer et al. 2011). Motif searches

TABLE 3. Representative examples of new HHRs in eukaryotes within tandem repetitive elements

	Eukaryotic species	Tandem repeats (size [nt])
Chromalveolata	<i>Phytophthora sojae</i>	1088
	<i>Hyaloperonospora parasitica</i>	786
	<i>Phytophthora infestans</i>	730–739
	<i>Hyaloperonospora arabidopsidis</i>	699
Plantae	<i>Selaginella moellendorffii</i> (Spike moss)	530–983
	<i>Vitis vinifera</i> (Grape)	854–865 ^a
	<i>Nicotiana tabacum</i> (Tobacco)	499
	<i>Oryza sativa</i> (Rice)	466
	<i>Lactuca perennis</i> (Blue lettuce)	441 ^a
Metazoa		
Cnidaria	<i>Nematostella vectensis</i> (Sea anemone)	330–356
	<i>Acropora millepora/A. digitifera</i> (Corals)	230
Platyhelmintha	<i>Schmidtea mediterranea</i> (Planarian worm)	379–422
	<i>Clonorchis sinensis</i> (Liver fluke)	304–310
Rotifera	<i>Adineta vaga</i> (Rotifer)	216
	<i>Philodina roseola</i> (Rotifer)	170–184
Annelida	<i>Helobdella robusta</i> (Leech)	385–448
Mollusca	<i>Pinctada martensii</i> (Pearl oyster)	350
Arthropoda	<i>Argas monolakensis</i>	229–231
Chordata	<i>Ambystoma mexicanum</i> (Salamander)	328
	<i>Pleurodeles waltl</i> (Newt)	325
	<i>Xenopus tropicalis</i> (Frog)	251–291 ^a
	Amniotes (Reptiles, birds, and mammals)	Only single motifs ^a

Note: based on studies by de la Peña and Garcia-Robles 2010a,b; I Garcia-Robles and M de la Peña, unpubl.

^aSome of the corresponding HHRs were experimentally verified.

designed to look for all three topologies revealed that the type II motifs are very common, particularly in microbial genomes, where they often comprise the largest fraction of HHRs (Jimenez et al. 2011; Perreault et al. 2011). Given these results, it is somewhat surprising that type II HHRs were not found earlier. Analysis of self-scission of some of these ribozymes in vitro showed that at physiological-like conditions (for HHRs found in the human microbiome), the type II are also the fastest HHRs, with half-lives of just several seconds in 1 mM Mg²⁺ (Jimenez et al. 2011). Similarly, a type II ribozyme discovered in a marine metagenome showed fast self-scission at 0.5 mM Mg²⁺ and 23°C (Perreault et al. 2011). These ribozymes were found in diverse genetic contexts, but, similar to types I and III, seem often associated with phage-related genes (de la Peña and Garcia-Robles 2010b; Perreault et al. 2011).

STEM I–II TERTIARY INTERACTION AND UNUSUAL VARIANTS

The interaction between stems I and II was observed among many different HHRs, as was the case of the two terminal loops in type III viroid and plant HHRs (De la Peña et al. 2003; Khvorova et al. 2003; Przybilski et al. 2005; Shepotinovskaya and Uhlenbeck 2008), and between inter-

nal and terminal loops in type I *S. mansoni* HHRs (Fig. 3A; Canny et al. 2004; Martick and Scott 2006). Recently, a pseudoknot has been shown to also fulfill this structural role in all three types of HHR (Jimenez et al. 2011; Perreault et al. 2011). In fact, close to 40% of HHRs not associated with repeat elements are predicted to possess base-pairing that bring stems I and II in proximity (Perreault et al. 2011). In addition, this interaction displays diversity in the number of different permutations in which it occurs: kissing loops between terminal loops, or terminal loop and internal loop; pseudoknots between a terminal loop and the 3' or 5' region outside of the minimal HHR (Fig. 3B; Perreault et al. 2011). The various possible structures used to bring stems I and II closer emphasize the importance of this interaction for the ribozyme's function. Examples observed in nature and with confirmed in vitro cleavage, are schematized in Figure 2.

Because it is an ideal model to understand the RNA structure–function relationship, the HHR has been a subject of exhaustive mutagenesis analyses (Ruffner et al. 1990; Nelson and Uhlenbeck 2008b). Furthermore, numerous chemogenetic studies used analogs of the conserved bases to acquire details about molecular interactions within the collection of all chemical groups in the core and their involvement in RNA folding and catalysis (Nelson and Uhlenbeck 2008a). Therefore, the consensus structure was not only supported by sequence alignments, but also experimentally scrutinized. As a result, it was generally accepted that for biologically significant activity, the ribozyme could not diverge from a minimum of 11 bases (CUGANGA, GAAA, and NUH) (Fig. 1B). However, given that the interaction of stems I and II dramatically increases activity, it is easier to imagine that slightly impaired catalytic cores diverging from the consensus would still self-cleave efficiently enough to achieve their function. This observation could explain natural occurrences of some HHR variants, a subset of which had already been observed earlier in viroid isolates (Przybilski and Hammann 2007a). At least 10 different natural variants have been found and, although some of them have reduced cleavage rates, all of them self-cleave in vitro and are found in the genetic context of consensus HHRs (Table 5).

Among the natural variants, some had been foretold by previous mutational analyses, like in the case of the C3G8

TABLE 4. HHR Motifs in viral and subviral entities

Type	Class	Species
dsDNA Virus	Caudovirales	5 Burkholderia phages ^a
		9 <i>Pseudomonas</i> phages ^a
		Xanthomonas phage Xp15
Viroid	Haloviruses	Halorubrum phage HF2 ^a
	Avsunviroidae	Avocado sunblotch viroid ^a
		Chrysanthemum chlorotic mottle viroid ^a
		Eggplant latent viroid ^a
		Mulberry small circular viroid like RNA 1 ^a
Virus satellite RNAs		Peach latent mosaic viroid ^a
		Arabidopsis mosaic virus satellite RNA ^a
		Cereal yellow dwarf virus RPV satellite RNA ^a
		Chicory yellow mottle virus satellite RNA ^a
		Lucerne transient streak virus satellite RNA ^a
		Solanum nodiflorum mottle virus satellite RNA
		Subterranean clover mottle virus satellite RNA
		Tobacco ringspot virus satellite RNA ^a
Unclassified elements		Velvet tobacco mottle virus satellite RNA
		Cherry small circular viroid-like RNA ^a
		Cherry small circular viroid-like RNA 2 ^a
		Carnation stunt associated viroid ^a

Note: based on studies by de la Peña and Garcia-Robles 2010a,b; Jimenez et al. 2011; Perreault et al. 2011; Seehafer et al. 2011.

^aSome of the corresponding HHRs were experimentally verified.

base pair (Fig. 1C), which is essential for catalytic activity (Martick and Scott 2006; Przybilski and Hammann 2007b; Nelson and Uhlenbeck 2008b). Covariation for that base pair has been observed, U3A8 to be more precise (Perreault et al. 2011). Similarly, a U10.1–U11.1 mismatch at the base of stem II, instead of the typical G10.1–C11.1 conserved base pair, illustrates the malleability of stem II. This result is not surprising, since that stem was known to be dispensable in vitro (McCall et al. 1992; Tuschl and Eckstein 1993; Long and Uhlenbeck 1994) and explains the observation of many HHRs with a stem II composed of only 1 or 2 bp (Perreault et al. 2011). Many ribozymes with short stem III have been observed as well, but in this case it was found that the ribozymes could fold into a dimeric configuration, permitted by a palindromic loop III, effectively yielding a stem III of 6–10 bp (Forster and Symons 1987a).

Other variants with a predicted absence of core-stabilizing interactions were expected to be inactive, based on previous biochemical and structural data. Nevertheless, they exist and self-cleave, albeit less efficiently than the well-established consensus. It is interesting to note that the compilation of variants observed in nature matches very well with the mutations that are the least disruptive in minimal and extended HHR studies (Table 5). Indeed, for all positions of the catalytic core, the identity of bases for which k_{obs} are shown in Table 5 are the most active mutants, after wild type. This correlation is also supported by the fact that no natural variants have been observed yet for the positions G5 and A14, both of which are essential

for the activity in the minimal HHR (Ruffner et al. 1990). Additional natural variants revealed by an ongoing study fit with this correlation as well (F Bolduc and JP Perreault, pers. comm.).

The only case where a natural variant does not correspond to the most active mutation assayed earlier (Ruffner et al. 1990) is the first base pair of stem II; the ribozyme is more active with a G10.1–U11.1 wobble pair than with a U10.1–U11.1 mismatch, but the latter variant is still more active than most sequences with conventional base pairs (A–U, U–A, U–G, and C–G). In spite of several very thorough mutational studies, including in vitro selection, there are several genomic variants that involve insertions that have never been evaluated before (see Table 5). Insertions studied earlier corresponded to a single base between positions A9 and G10.1 of many natural variants, which exhibited k_{obs} comparable to typical HHRs (De la Peña and Flores 2001;

Shepotinovskaya and Uhlenbeck 2008, 2010). The most notable novel insertions were two bases found upstream of C3 in two different ribozyme sequences. Both of these instances originate from solar salterns (Perreault et al. 2011), environments with saturating concentrations of salts. Although neither of them cleave efficiently at standard concentrations of Mg^{2+} , their cleavage rate in 100 mM Mg^{2+} (physiological concentration in this case) is comparable to other ribozymes (Perreault et al. 2011). Potential implications for Mg^{2+} -dependent regulation are evident, as the corresponding organisms require high-salt concentration for growth, dilution of their environment thus represents a dangerous situation requiring rapid adaptation (Torsvik and Dundas 1980).

Prospective roles of variant HHRs go beyond this example. Indeed, regulation implies an on/off switching mechanism, which is improbable if the ribozyme cleaves constitutively, like in the rolling circle replication of viroids. Hence, less active variant ribozymes are more likely to be turned-on in specific conditions. This is the case of the glucosamine-6-phosphate-activated *glmS* ribozyme (Winkler et al. 2004) and has been hypothesized for a discontinuous HHR found in some mammals (Martick et al. 2008; Scott et al. 2009). The latter harbors a very large loop I (150–1739 nt), which could contain protein-binding sites in addition to greatly reducing cleavage efficiency, analogous to experimentally selected allosteric HHR with an aptamer in loop II (Zivarts et al. 2005). Analysis of HDV-like ribozymes in mosquitoes showed that both expression levels and the cleavage extent vary through the developmental stages of

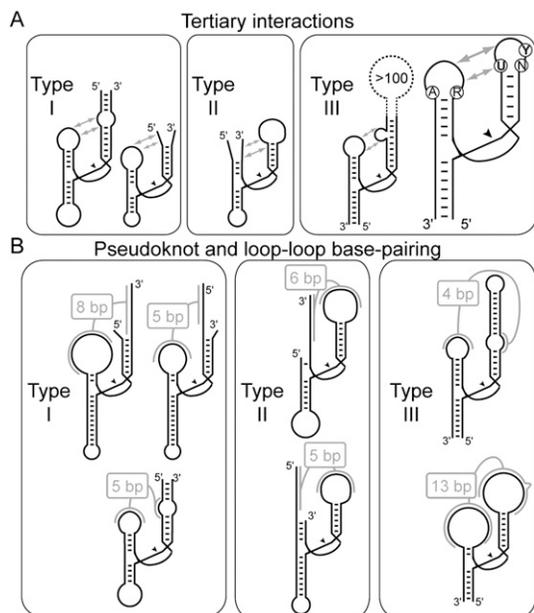


FIGURE 3. Structural versatility of stem–stem II interaction. Many different possible permutations that allow these two stems to interact are observed. (A) Tertiary interactions between stem I and II. Gray arrows illustrate the tertiary interaction. HHR examples pictured from left to right are found in the nematode *Schistosoma mansoni* (this example has been crystallized and shows the importance of the tertiary interaction), the viroid ASBVd, the bacterium *Bukholderia oklahomensis*, the mouse, and the viroid CChMVd (circled nucleotides are typically involved in this interaction for HHR sequences found in viroids). (B) Base-pairing interactions between stems I and II. The large array of different putative pseudoknots or kissing-loop interactions observed in HHR sequences is pictured. HHR examples pictured from left to right on the first row are found in the fungus *Yarrowia lipolytica*, the mouse and human microbiome, the bacterium *Agrobacterium tumefaciens*, and the fly *Drosophila pseudoobscura*. HHRs of the second row are found in the crickets *Dolichopoda*, marine metagenome, and in the bee *Apis mellifera*.

the insect, suggesting that the ribozyme self-scission is indeed regulated (Webb et al. 2009). A potential mechanism of ribozyme regulation may involve short antisense RNAs, such as those found to map to a putative HDV-like ribozyme in silkworm retrotransposons (Kawaoka et al. 2008; Ruminski et al. 2011).

DISCUSSION

The various approaches used to find the HHRs reviewed herein are complementary, and each has specific advantages and disadvantages. A sequence homology-based approach has the advantage of indirectly following evolution's course, without any preconceived functional biases and can thus help find HHRs even if, for example, the tertiary interaction between stems I and II is not fully understood. Moreover, sequence-homology searches are computationally very fast and allow for scanning of the whole GeneBank for a HHR motif in a few seconds. A drawback of this ap-

proach is that it relies on known sequences, making it difficult to find very divergent HHRs. On the other hand, structure-based approaches rely on the structural and biochemical data of the RNA that we look for, and this knowledge is extensive in the case of the HHR. The disadvantage is that the sequences to be found must perfectly match the string that is searched for. Therefore, designing a good descriptor is critical. Furthermore, an all-encompassing descriptor generates numerous false positives. To filter these false positives, it is possible to estimate the folding free energies of the candidate sequences and exclude those that do not conform to hammerhead fold. This filtering might exclude real HHRs that compete with a more stable structure; on the other hand, it allows one to predict stably folding HHRs motifs embedded in longer RNA sequences. Finally, the covariance model in Infernal incorporates both sequence and structure constraints. While it relies on a secondary structure alignment, it can find sequences with only a tenuous sequence homology, relying mostly on structure-based criteria. It also allows some degree of divergence from the consensus; however, it is computationally intensive and can require weeks to complete a search on large sequence databases. While all of these approaches yield many HHRs and overlapping results have been obtained by our groups, a side-by-side comparison has not been performed on the same sequence data set to directly compare the methods and evaluate the identified ribozyme candidates.

A particular feature of natural HHRs, the tertiary interactions between helix I and helix II, have been experimentally confirmed for only a few cases. Nevertheless, most of the naturally occurring HHRs appear to have such a conserved interaction (Dufour et al. 2009; Jimenez et al. 2011; Perreault et al. 2011). For canonical types I and III HHRs, these interactions require not only a few conserved nucleotides, but also specific lengths of stem I (~6 bp) and stem II (~4 bp) as found in many of the detected HHRs (Fig. 2). For those new HHRs showing atypical helix lengths and loops, future biochemical and structural studies will help us understand their global architecture. Even so, our current knowledge of other natural three-helical RNAs indicates that many different forms of tertiary interactions are possible among these motifs (de la Peña et al. 2009; Jimenez et al. 2011; Perreault et al. 2011).

From a functional point of view, most of the detected eukaryotic HHRs appear to be associated with short tandem repeats, which we suggest are retrotransposons of the SINE family. SINEs are repetitive elements amplified in the genomes via RNA intermediates that use the propagation machinery of autonomous retrotransposons. For the case of HHR-containing SINEs, the ribozymes may function in self-cleavage of oligomeric transcripts to yield SINE monomers in a fashion similar to the originally proposed rolling-circle replication of RNA plant pathogens, like viroids and viral satellite RNAs. Alternatively, the overhangs generated

TABLE 5. Divergence from consensus observed in natural variants

Consensus	Natural variants	Reference	Artificial variants	Reference
Stem I (N1N2)				
C3G8	ins UG ^a or GC U3, ^a U3A8	(Perreault et al. 2011) (Perreault et al. 2011)	U3 < 0.02, U3A8 ~0.4, G3C8 ~0.08–0.8	(Ruffner et al. 1990; Przybilski and Hammann 2007b)
U4	C4	(Ambros and Flores 1998)	C < 0.007	(Ruffner et al. 1990)
G5	none observed		no cleavage	(Ruffner et al. 1990)
A6	C6 ins A	(Perreault et al. 2011) (Perreault et al. 2011)	C < 0.005	(Ruffner et al. 1990)
N7				
A9	U9 frequent ins N	(Przybilski and Hammann 2007a) (Rocheleau and Pelchat 2006; Forster and Symons 1987a; De la Peña and Flores 2001)	U < 0.06 N ~1	(Ruffner et al. 1990) (De la Peña and Flores 2001; Warashina et al. 2004)
G10.1C11.1	U10–U11 mispair	(Perreault et al. 2011)	U10U11 = 0.35, (G10U11 = 0.43)	(Ruffner et al. 1990; Long and Uhlenbeck 1994)
Stem II	single bp stem	(Perreault et al. 2011)	single bp stem ~0.1	(Tuschl and Eckstein 1993)
G12	none observed		G 0.002–0.02	(Ruffner et al. 1990)
A13	G13 ins U	(Przybilski and Hammann 2007a) (Perreault et al. 2011)	G < 0.03	(Ruffner et al. 1990)
A14	none observed		no cleavage	(Ruffner et al. 1990)
A15.1U16.1	U15A16	(Perreault et al. 2011)	U15A16 < 0.002, 0.02–0.3	(Ruffner et al. 1990; Kore et al. 2000)
Stem III	single bp stem	(Forster and Symons 1987a)	single bp stem	(Forster and Symons 1987a)

(ins) Insertion, (N) any base, (H) A, C, or U.

^aIndicate multiple covarying positions for a given ribozyme.

Artificial mutants, list of relative cleavage rates of mutants compared with wt, data from Ruffner et al. (1990) correspond to a minimal HHR model.

by the self-cleavage reaction might be a requirement for the genomic integration of such sequences (Fig. 4A). This could explain the predominance of type I HHRs in SINEs, because this topology is more favorable to the proposed mechanism in Figure 4A. On the other hand, HHR activity in repetitive genomic sequences might negatively influence the production of piRNAs (Fig. 4B). HHRs are not the only catalytic RNA motifs associated with retrotransposons: Recent analysis of autonomous retrotransposons (LINEs) revealed that many are terminated by HDV-like self-cleaving ribozymes, suggesting that in some species HHRs and HDV-like ribozymes are dedicated to processing of SINE and LINE elements, respectively (Ruminski et al. 2011). This dedicated processing, however, does not seem to take place in mammals, where the two ribozyme families do not appear to be associated with retrotransposons (Salehi-Ashtiani et al. 2006; Martick et al. 2008; de la Peña and Garcia-Robles 2010a; Perreault et al. 2011; Seehafer et al. 2011).

In amniotes, the elevated conservation of intronic HHRs suggests that they have a specific biological function in mRNA biogenesis yet to be characterized. As a parsimonious hypothesis, these HHRs, which show a high sequence and structural similarity to those from retrotransposons from lower metazoans (Martick et al. 2008; de la Peña and Garcia-Robles 2010a), may play an

exapted/domesticated role in the control of alternative splicing in a similar way to the recently described artificial intronic HHRs (de la Peña and Garcia-Robles 2010a; Pastor et al. 2011).

The ubiquity of HHRs suggests that these ribozymes are either ancient and were maintained in functional form throughout evolution, or that they can arise independently several times. The 11 conserved nucleotides of the catalytic core, together with three flanking helices require about 40 bits of information (two bits for all invariant positions and two for each base pair, assuming a minimum of three base pair per helix), without any requirements for tertiary contacts. Thus, the frequency of this minimal ribozyme structure in a random sequence would be about 10^{-12} and lower if longer helices and tertiary contacts are required, suggesting that the ribozymes are not likely to have evolved independently many times. On the other hand, a large scale meta-analysis of the newly discovered sequences described here indicated distinct coevolutionary patterns for HHRs of different topology and genomic location (F Hoffgaard, C Seehafer, K Hamacher, and C Hammann, unpubl.). This observation points toward an independent appearance of the HHRs a few times in evolution. This scenario is also supported by the analysis of self-cleaving ribozymes isolated using an in vitro selection from a random library (Salehi-Ashtiani and Szostak 2001), which showed that active HHRs

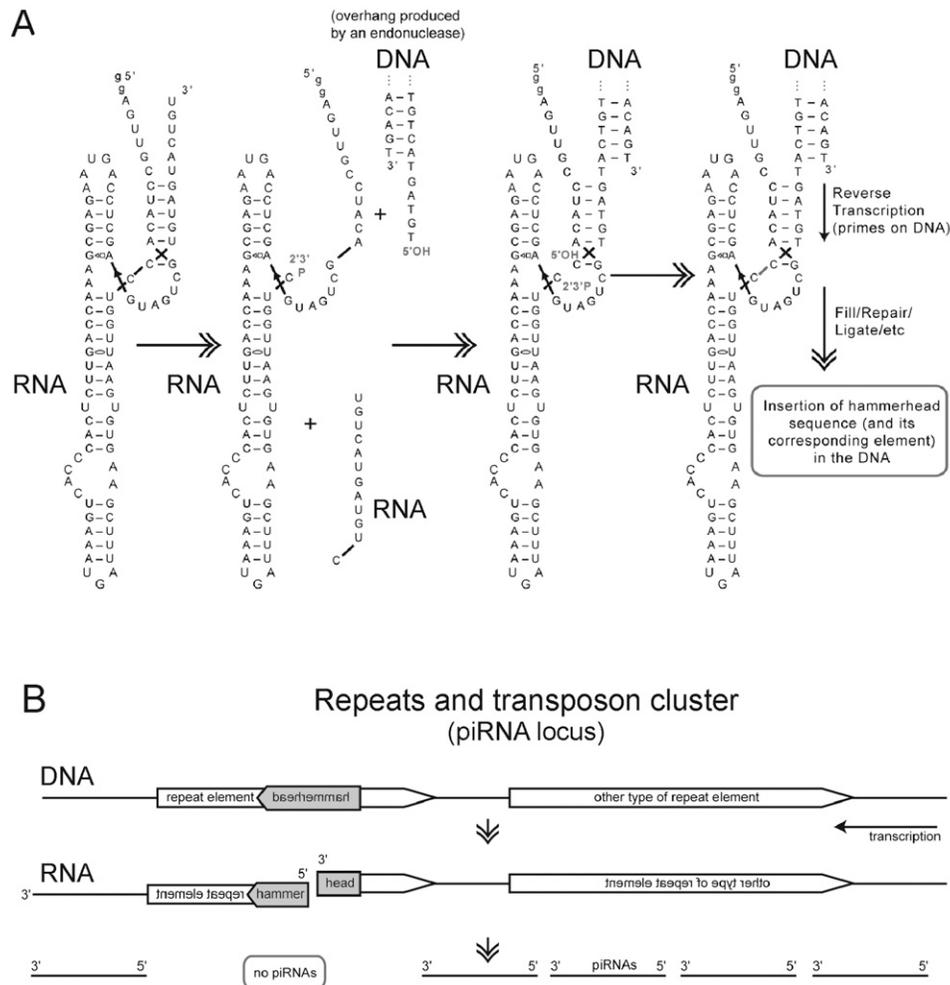


FIGURE 4. Possible function of the HHR in genomes. (A) HHR-mediated integration of nucleic acid sequences in genomes. (B) Self-cleavage of HHR sequences within transcripts of repeats or transposon clusters might interfere with the production of piRNAs or other small RNAs.

can arise independently several times, providing an example of convergent molecular evolution. However, the diversity of the starting DNA pool was likely on the order of 10^{15} , and so the HHR motif is statistically likely to be found there. This result stands in contrast to the structurally and informationally more complex HDV-like ribozymes, which are also widespread in nature (Webb et al. 2009; Ruminiski et al. 2011; Webb and Luptak 2011), but have not been identified using in vitro selection experiments.

The biological roles of the newly discovered HHRs are currently under investigation, and several scenarios can be imagined. The common feature is the bisection of a transcribed RNA by the catalytic RNA motif. This reaction in itself might serve, for example, to uncouple two or more physically coupled genetic units. Alternatively, the reaction products might be substrates for subsequent reactions, be it RNA-catalyzed ligation, *trans*-cleavage, or a protein-catalyzed reaction. These alternative scenarios are exemplified

by mammalian ESTs, which intriguingly showed RNA fusions of HHR-cleaved introns with U5 or U6 snRNAs (de la Peña and Garcia-Robles 2010a).

Figure 5 displays selected scenarios for the biological function of different genetic contexts of HHR motifs. They might simply serve in the processing of polycistronic tRNA transcripts, as exemplified by a sequence from the marine metagenome (Fig. 5A). All transcribed HHR motifs might result in the generation of regulatory small RNAs (Fig. 5B) and may also help to prevent expression from a bacterial promoter. This might be instrumental for integration of the phage DNA sequence in the bacterial genome, as seems conceivable for the *Pseudomonas aeruginosa* phage PaP3 (Fig. 5B). A particularly intriguing arrangement is that of multiple HHRs, as observed in *Azorhizobium caulinodans*. Here, the self-cleavage and re-ligation activities of the HHR motifs might lead to the generation of novel open reading frames, which might represent a bacterial form of alterna-

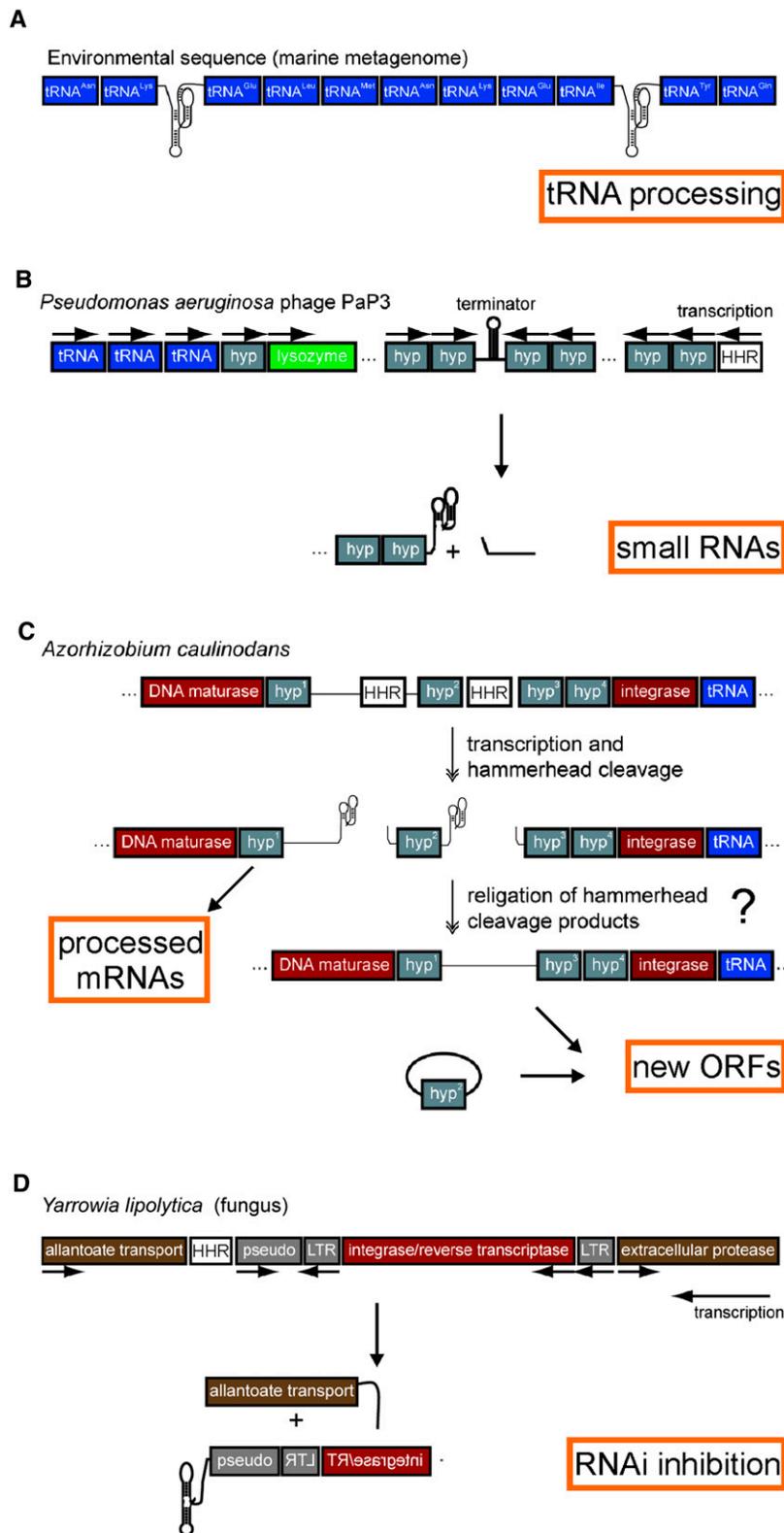


FIGURE 5. Possible functions of HHR motifs in various transcriptional contexts. (A) Processing of polycistronic tRNA transcripts. (B) Generation of small RNAs. (C) Generation of alternative open reading frames by HHR self-cleavage and re-ligation. (D) Inhibition of RNAi by prevention of the generation of complementary RNA strands from fungal and retrotransposon promoters.

tive splicing (Fig. 5C). Further, RNA self-cleavage by an HHR might prevent the generation of two RNA strands of opposite polarity, which might otherwise be generated from an mRNA and a retrotransposon promoter, as exemplified for an HHR motif from the fungus *Yarrowia lipolytica* (Fig. 5D). This, in turn, would prevent the generation of a double strand, and consequently lead to the prevention or inhibition of RNAi. Association with mobile elements, such as phages or SINEs, could explain the patchy distribution of HHRs, which are present in all kingdoms, but have been found in a fraction of sequenced species (Tables 1, 2, 4).

Either of these pathways appears attractive, and many more are conceivable. Future experiments will reveal the real functions that the most widely distributed small endonucleolytic ribozyme exerts. With this resurgence in HHR research activity, we can expect exciting discoveries that will unveil at least some of the biological roles of HHRs in the years to come.

ACKNOWLEDGMENTS

This work was supported by grant HA3459/3 and a Heisenberg fellowship HA3459/5 by the Deutsche Forschungsgemeinschaft to C.H., by the NIH (R01GM094929) and the Pew Charitable Trusts to A.L., and by the Ministerio de Ciencia e Innovacion of Spain (BFU2011-23398) to M. d-I P.

REFERENCES

- Ambros S, Flores R. 1998. In vitro and in vivo self-cleavage of a viroid RNA with a mutation in the hammerhead catalytic pocket. *Nucleic Acids Res* **26**: 1877–1883.
- Branch AD, Robertson HD. 1984. A replication cycle for viroids and other small infectious RNAs. *Science* **223**: 450–455.
- Buzayan JM, Gerlach WL, Bruening G. 1986. Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature* **323**: 349–353.
- Canny MD, Jucker FM, Kellogg E, Khvorova A, Jayasena SD, Pardi A. 2004. Fast cleavage kinetics of a natural hammerhead ribozyme. *J Am Chem Soc* **126**: 10848–10849.
- Chi YI, Martick M, Lares M, Kim R, Scott WG, Kim SH. 2008. Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol* **6**: e234. doi: 10.1371/journal.pbio.0060234.

- Cremisi F, Scarabino D, Carluccio MA, Salvadori P, Barsacchi G. 1992. A new ribozyme: a catalytic activity in search of a function. *Proc Natl Acad Sci* **89**: 1651–1655.
- D'Souza M, Larsen N, Overbeek R. 1997. Searching for patterns in genomic data. *Trends Genet* **13**: 497–498.
- Daros JA, Flores R. 1995. Identification of a retroviroid-like element from plants. *Proc Natl Acad Sci* **92**: 6856–6860.
- De la Peña M, Flores R. 2001. An extra nucleotide in the consensus catalytic core of a viroid hammerhead ribozyme: implications for the design of more efficient ribozymes. *J Biol Chem* **276**: 34586–34593.
- De la Peña M, Gago S, Flores R. 2003. Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity. *EMBO J* **22**: 5561–5570.
- de la Peña M, Garcia-Robles I. 2010a. Intronic hammerhead ribozymes are ultraconserved in the human genome. *EMBO Rep* **11**: 711–716.
- de la Peña M, Garcia-Robles I. 2010b. Ubiquitous presence of the hammerhead ribozyme motif along the tree of life. *RNA* **16**: 1943–1950.
- de la Peña M, Dufour D, Gallego J. 2009. Three-way RNA junctions with remote tertiary contacts: A recurrent and highly versatile fold. *RNA* **15**: 1949–1964.
- Denti MA, Martinez de Alba AE, Sagesser R, Tsagris M, Tabler M. 2000. A novel RNA-binding protein from *Triturus carnifex* identified by RNA-ligand screening with the new hammerhead ribozyme. *Nucleic Acids Res* **28**: 1045–1052.
- Dufour D, de la Peña M, Gago S, Flores R, Gallego J. 2009. Structure-function analysis of the ribozymes of chrysanthemum chlorotic mottle viroid: a loop-loop interaction motif conserved in most natural hammerheads. *Nucleic Acids Res* **37**: 368–381.
- Eddy S. 2005. RNABOB 2.1. Available: <ftp://selab.janelia.org/pub/software/rnabob/>.
- Epstein LM, Coats SR. 1991. Tissue-specific permutations of self-cleaving newt satellite-2 transcripts. *Gene* **107**: 213–218.
- Epstein LM, Gall JG. 1987. Self-cleaving transcripts of satellite DNA from the newt. *Cell* **48**: 535–543.
- Ferbeyre G, Smith JM, Cedergren R. 1998. Schistosome satellite DNA encodes active hammerhead ribozymes. *Mol Cell Biol* **18**: 3880–3888.
- Ferbeyre G, Bourdeau V, Pageau M, Miramontes P, Cedergren R. 2000. Distribution of hammerhead and hammerhead-like RNA motifs through the GenBank. *Genome Res* **10**: 1011–1019.
- Flores R, Delgado S, Gas ME, Carbonell A, Molina D, Gago S, De la Peña M. 2004. Viroids: the minimal non-coding RNAs with autonomous replication. *FEBS Lett* **567**: 42–48.
- Forster AC, Symons RH. 1987a. Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* **49**: 211–220.
- Forster AC, Symons RH. 1987b. Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell* **50**: 9–16.
- Garrett TA, Pabon-Peña LM, Gokaldas N, Epstein LM. 1996. Novel requirements in peripheral structures of the extended satellite 2 hammerhead. *RNA* **2**: 699–706.
- Gräf S, Przybilski R, Steger G, Hammann C. 2005. A database search for hammerhead ribozyme motifs. *Biochem Soc Trans* **33**: 477–478.
- Hammann C, Lilley DM. 2002. Folding and activity of the hammerhead ribozyme. *ChemBioChem* **3**: 690–700.
- Haseloff J, Gerlach WL. 1988. Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature* **334**: 585–591.
- Hertel KJ, Pardi A, Uhlenbeck OC, Koizumi M, Ohtsuka E, Uesugi S, Cedergren R, Eckstein F, Gerlach WL, Hodgson R, et al. 1992. Numbering system for the hammerhead. *Nucleic Acids Res* **20**: 3252. doi: 10.1093/nar/20.12.3252.
- Jimenez RM, Delwart E, Luptak A. 2011. Structure-based search reveals hammerhead ribozymes in the human microbiome. *J Biol Chem* **286**: 7737–7743.
- Kalweit A, Przybilski R, Seehafer C, de la Peña M, Hammann C. 2012. Characterization of hammerhead ribozyme reactions. *Methods Mol Biol* **848**: 5–20.
- Kaper JM, Tousignant ME, Steger G. 1988. Nucleotide sequence predicts circularity and self-cleavage of 300-ribonucleotide satellite of arabis mosaic virus. *Biochem Biophys Res Commun* **154**: 318–325.
- Kawaoka S, Hayashi N, Katsuma S, Kishino H, Kohara Y, Mita K, Shimada T. 2008. Bombyx small RNAs: genomic defense system against transposons in the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* **38**: 1058–1065.
- Khvorova A, Lescoute A, Westhof E, Jayasena SD. 2003. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nat Struct Biol* **10**: 708–712.
- Klein DJ, Ferré-D'Amaré AR. 2006. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* **313**: 1752–1756.
- Kore AR, Carola C, Eckstein F. 2000. Attempts to obtain more efficient GAC-cleaving hammerhead ribozymes by in vitro selection. *Bioorg Med Chem* **8**: 1767–1771.
- Kuo MY, Sharmeen L, Dinter-Gottlieb G, Taylor J. 1988. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *J Virol* **62**: 4439–4444.
- Leontis NB, Stombaugh J, Westhof E. 2002. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res* **30**: 3497–3531.
- Lilley DM. 2005. Structure, folding and mechanisms of ribozymes. *Curr Opin Struct Biol* **15**: 313–323.
- Long DM, Uhlenbeck OC. 1994. Kinetic characterization of intramolecular and intermolecular hammerhead RNAs with stem II deletions. *Proc Natl Acad Sci* **91**: 6977–6981.
- Luzi E, Eckstein F, Barsacchi G. 1997. The newt ribozyme is part of a riboprotein complex. *Proc Natl Acad Sci* **94**: 9711–9716.
- Macke TJ, Ecker DJ, Gutell RR, Gautheret D, Case DA, Sampath R. 2001. RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Res* **29**: 4724–4735.
- Martick M, Scott WG. 2006. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* **126**: 309–320.
- Martick M, Horan LH, Noller HF, Scott WG. 2008. A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA. *Nature* **454**: 899–902.
- McCall MJ, Hendry P, Jennings PA. 1992. Minimal sequence requirements for ribozyme activity. *Proc Natl Acad Sci* **89**: 5710–5714.
- Nawrocki EP, Kolbe DL, Eddy SR. 2009. Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25**: 1335–1337.
- Nelson JA, Uhlenbeck OC. 2008a. Hammerhead redux: Does the new structure fit the old biochemical data? *RNA* **14**: 605–615.
- Nelson JA, Uhlenbeck OC. 2008b. Minimal and extended hammerheads utilize a similar dynamic reaction mechanism for catalysis. *RNA* **14**: 43–54.
- Pastor T, Dal Mas A, Talotti G, Bussani E, Pagani F. 2011. Intron cleavage affects processing of alternatively spliced transcripts. *RNA* **17**: 1604–1613.
- Penedo JC, Wilson TJ, Jayasena SD, Khvorova A, Lilley DM. 2004. Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements. *RNA* **10**: 880–888.
- Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, Breaker RR. 2011. Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput Biol* **7**: e1002031. doi: 10.1371/journal.pcbi.1002031.
- Prody GA, Bakos JT, Buzayan JM, Schneider IR, Bruening G. 1986. Autolytic processing of dimeric plant virus satellite RNA. *Science* **231**: 1577–1580.
- Przybilski R, Hammann C. 2006. The hammerhead ribozyme structure brought in line. *ChemBioChem* **7**: 1641–1644.
- Przybilski R, Hammann C. 2007a. Idiosyncratic cleavage and ligation activity of individual hammerhead ribozymes and core sequence variants thereof. *Biol Chem* **388**: 737–741.
- Przybilski R, Hammann C. 2007b. The tolerance to exchanges of the Watson-Crick base pair in the hammerhead ribozyme core is determined by surrounding elements. *RNA* **13**: 1625–1630.
- Przybilski R, Gräf S, Lescoute A, Nellen W, Westhof E, Steger G, Hammann C. 2005. Functional hammerhead ribozymes naturally encoded in the genome of *Arabidopsis thaliana*. *Plant Cell* **17**: 1877–1885.

- Reeder J, Giegerich R. 2009. RNA secondary structure analysis using the RNAsnp package. In *Current Protocols in Bioinformatics*, Chapter 12, Unit 12, p. 18. John Wiley & Sons, Inc., New York.
- Riccitelli NJ, Luptak A. 2010. Computational discovery of folded RNA domains in genomes and in vitro selected libraries. *Methods* **52**: 133–140.
- Rocheleau L, Pelchat M. 2006. The Subviral RNA Database: a toolbox for viroids, the hepatitis *delta virus* and satellite RNAs research. *BMC Microbiol* **6**: 24. doi: 10.1186/1471-2180-6-24.
- Rojas AA, Vazquez-Tello A, Ferbeyre G, Venanzetti F, Bachmann L, Paquin B, Sbordoni V, Cedergren R. 2000. Hammerhead-mediated processing of satellite pDo500 family transcripts from *Dolichopoda* cave crickets. *Nucleic Acids Res* **28**: 4037–4043.
- Rubino L, Tousignant ME, Steger G, Kaper JM. 1990. Nucleotide sequence and structural analysis of two satellite RNAs associated with chicory yellow mottle virus. *J Gen Virol* **71**: 1897–1903.
- Ruffner DE, Stormo GD, Uhlenbeck OC. 1990. Sequence requirements of the hammerhead RNA self-cleavage reaction. *Biochemistry* **29**: 10695–10702.
- Ruminski DJ, Webb CH, Riccitelli NJ, Luptak A. 2011. Processing and translation initiation of non-long terminal repeat retrotransposons by hepatitis delta virus (HDV)-like self-cleaving ribozymes. *J Biol Chem* **286**: 41286–41295.
- Salehi-Ashtiani K, Szostak JW. 2001. In vitro evolution suggests multiple origins for the hammerhead ribozyme. *Nature* **414**: 82–84.
- Salehi-Ashtiani K, Luptak A, Litovchick A, Szostak JW. 2006. A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science* **313**: 1788–1792.
- Saville BJ, Collins RA. 1990. A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell* **61**: 685–696.
- Scott WG, Martick M, Chi YI. 2009. Structure and function of regulatory RNA elements: ribozymes that regulate gene expression. *Biochim Biophys Acta* **1789**: 634–641.
- Seehafer C, Kalweit A, Steger G, Gräf S, Hammann C. 2011. From alpaca to zebrafish: Hammerhead ribozymes wherever you look. *RNA* **17**: 21–26.
- Shepotinovskaya IV, Uhlenbeck OC. 2008. Catalytic diversity of extended hammerhead ribozymes. *Biochemistry* **47**: 7034–7042.
- Shepotinovskaya I, Uhlenbeck OC. 2010. Enhanced product stability in the hammerhead ribozyme. *Biochemistry* **49**: 4494–4500.
- Tabler M, Tsagris M. 2004. Viroids: petite RNA pathogens with distinguished talents. *Trends Plant Sci* **9**: 339–348.
- Torsvik T, Dundas ID. 1980. Persisting phage infection in *Halobacterium salinarum* str. 1. *J Gen Virol* **47**: 29–36.
- Tuschl T, Eckstein F. 1993. Hammerhead ribozymes: importance of stem-loop-II for activity. *Proc Natl Acad Sci* **90**: 6991–6994.
- Uhlenbeck OC. 1987. A small catalytic oligoribonucleotide. *Nature* **328**: 596–600.
- Warashina M, Kuwabara T, Nakamatsu Y, Takagi Y, Kato Y, Taira K. 2004. Analysis of the conserved P9-G10.1 metal-binding motif in hammerhead ribozymes with an extra nucleotide inserted between A9 and G10.1 residues. *J Am Chem Soc* **126**: 12291–12297.
- Webb CH, Luptak A. 2011. HDV-like self-cleaving ribozymes. *RNA Biol* doi: 10.4161/ma.8.5.16226.
- Webb CH, Riccitelli NJ, Ruminski DJ, Luptak A. 2009. Widespread occurrence of self-cleaving ribozymes. *Science* **326**: 953. doi: 10.1126/science.1178084.
- Wilson TJ, Lilley DM. 2009. Biochemistry. The evolution of ribozyme chemistry. *Science* **323**: 1436–1438.
- Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**: 281–286.
- Zhang Y, Epstein LM. 1996. Cloning and characterization of extended hammerheads from a diverse set of caudate amphibians. *Gene* **172**: 183–190.
- Zivarts M, Liu Y, Breaker RR. 2005. Engineered allosteric ribozymes that respond to specific divalent metal ions. *Nucleic Acids Res* **33**: 622–631.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.