

# Hidden ribozymes in eukaryotic genome sequence

Sean P Ryder

Address: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, LRB-906, Worcester, MA 01605, USA

Email: sean.ryder@umassmed.edu

*1000 Biology Reports* 2010, 2:50 (doi:10.3410/B2-50)

The electronic version of this article is the complete one and can be found at: <http://f1000.com/reports/biology/content/2/50>

## Abstract

The small self-cleaving ribozymes fold into complex tertiary structures to promote autocatalytic cleavage or ligation at a precise position within their sequence. Until recently, relatively few examples had been identified. Two papers now reveal that self-cleaving ribozymes are prevalent in eukaryotic genomes and, in some cases, might play a role in regulating gene expression.

## Introduction and context

### RNA catalysts

Twenty five years ago the surprising discovery that certain RNA sequences catalyze chemical reactions ignited the hunt for RNA enzymes (ribozymes) and spurred intense investigation into the chemical and structural basis of RNA catalysis [1-3]. Since then, only a few naturally occurring ribozymes have been discovered. These include the small self-cleaving Varkud satellite (VS), hepatitis delta virus (HDV), hairpin, and hammerhead ribozymes found in selfish genetic elements such as satellite plasmids, viroids, and viruses [4-10]. Ribozyme activity resolves rolling circle replication intermediates through autocatalytic cleavage of RNA multimers followed by autocatalytic ligation of individual monomeric units into closed single-stranded circles.

Incredible progress has been made in understanding how RNA enzymes fold into complex structures and enhance the rate of chemical reactions [1]. All naturally occurring ribozymes catalyze phosphodiester transfer or hydrolysis with exquisite substrate specificity. A series of high-resolution crystal structures reveal the diversity of molecular architectures that enable specific substrate recognition and catalysis [1,11-14]. Corresponding biochemical experiments describe the relative contributions of individual nucleotides and chemical groups that contribute to folding and catalysis [15-21]. Helical junctions, non-Watson-Crick base pairs, pseudoknots,

and metal ion-binding sites provide the foundation for RNA active sites capable of metal-assisted or general acid-base catalysis.

The observation that most extant ribozymes are found within selfish genetic elements or can trace their lineage through self-splicing mobile elements led to the hypothesis that ribozymes are modern vestigial remnants of an evolutionary era predominated by RNA 'life' forms [22]. This hypothesis led researchers to apply *in vitro* evolution technology (systematic evolution of ligands through exponential enrichment, or SELEX) to identify new artificial RNA catalysts, including ribozymes capable of RNA polymerization, peptide bond formation, tRNA charging, and many other biological activities required for primitive 'life' [23-26]. In contrast, the search for naturally occurring ribozymes moved forward at a relatively slow pace, suggesting that ribozymes may be rare entities – curiosities of unusual biological systems that have been largely discarded during evolution.

### Regulatory ribozymes

The discovery in 2004 of a metabolite-sensitive ribozyme with a demonstrable role in gene regulation hinted at an expanded biological role for RNA catalysts [27]. The *glmS* (glutamine-fructose-6-phosphate amidotransferase) ribozyme/riboswitch, first identified in *Bacillus subtilis* and related Gram-positive bacteria, is found in the 5' leader of transcripts that encode an enzyme necessary for

the biosynthesis of glucosamine-6-phosphate, a cell wall precursor. The ribozyme binds to glucosamine-6-phosphate in order to activate self-cleavage at a defined position. Cleavage destabilizes the message, initiating a negative feedback loop that effectively reduces the biosynthesis of new glucosamine-6-phosphate [28]. This finding reinvigorated the search for new RNA enzymes.

Two years later, Szostak and colleagues [29] devised a clever selection approach to identify self-cleaving ribozymes in the human genome. The authors generated a library of single-stranded, closed circular genomic DNA fragments approximately 150 nucleotides in length. This library was used as a template for rolling circle *in vitro* transcription to generate long RNA multimers. After incubating the multimers at physiological salt concentrations, the authors purified sequences that were capable of self-cleavage but migrated at dimer length thus retaining one copy of the intact ribozyme. The recovered RNAs were used to generate the next round of circular DNA template. After 12 rounds of selection and amplification, they cloned and sequenced three autocatalytic self-cleaving ribozymes present in the *OR4K15* (olfactory receptor family 4 subfamily K member 15), *IGF1R* (insulin-like growth factor receptor 1 gene), and *CPEB3* (cytoplasmic polyadenylation element binding protein 3) genes and a fourth in a LINE 1 (long interspersed repetitive element 1) retrotransposon. This result clearly demonstrates that ribozymes are not as rare as initially believed.

Further characterization of the *CPEB3* ribozyme demonstrated that it folds into a structure similar to the HDV ribozyme and catalyzes the same chemical reaction at the same relative position [29]. There is extensive secondary structure conservation between the two, but only six nucleotides are identical in the primary sequence. The *CPEB3* ribozyme is conserved in mammals but is not found in other vertebrates. It is present in the second intron of the *CPEB3* gene. Cleavage is hypothesized to regulate gene expression through destruction of pre-mRNA or possibly through formation of a truncated form of *CPEB3* that lacks the N-terminal domain.

The extensive structural homology between ribozymes within a functional class and the absence of strong primary sequence identity led to a new bioinformatic strategy to identify ribozymes in genome sequences. The development of pattern-matching tools powerful enough to combine conserved sequence identity with secondary structural features into a single descriptor opened the door to genome-wide searches for new variations of the known ribozymes [30,31]. In a

pioneering study, Przybilski and colleagues [32] used this approach to search the EMBL (European Molecular Biology Laboratory) database for self-cleaving hammerhead ribozymes. They identified two previously undiscovered ribozymes in chromosome IV of the *Arabidopsis thaliana* genome. The two ribozymes share conserved flanking sequences, are transcribed in a variety of plant tissues, and are capable of self-cleavage *in vitro*. The authors hypothesized that the ribozymes are not the result of viroid integration but instead evolved independently to perform a biological function, which to date has not been characterized. A clear outcome from this study is that pattern-based computational searches can successfully identify functional ribozyme variants that were previously hidden within a sequenced eukaryotic genome.

### Major recent advances

Two recent papers expand upon this approach to identify new ribozymes in a vast array of eukaryotic genomes [33,34]. Martick and colleagues [33] searched mammalian mRNA sequence databases for discontinuous hammerhead ribozymes that maintained required secondary structure features but allowed for significant flexibility in the loop regions. They reasoned that large insertions in loop 1 or loop 3 of the ribozyme should not adversely affect ribozyme activity but would make it difficult to search for ribozymes using standard sequence-based patterns (Figure 1a). Their search revealed three hammerhead-like ribozymes in the 3'-untranslated regions (UTRs) of murine *Clec2d* (C-type lectin domain family 2 member D), *Clec2e*, and *Clec2d11* transcripts. Subsequent conservation-based searches revealed hammerhead ribozyme-like sequences in the 3'-UTR of *Clec2* homologs in other mammals, including rat, horse, and platypus, suggesting a possible conserved function.

*Clec2d* is an osteoclast inhibitory lectin required for normal bone physiology [35]. *Clec2d* knockout mice display reduced bone volume and osteopenia but are otherwise normal. Insertion of the *Clec2d* ribozyme into the 3'-UTR of a luciferase reporter reduces expression by 80%, whereas a catalytically dead variant has no effect [33]. This suggests that the ribozyme is a negative regulatory element, presumably acting to destabilize the *Clec2d* message. If so, then ribozyme function may influence bone homeostasis, but this remains to be tested *in vivo*.

In a similar approach, Luptak and colleagues [34] used structural descriptors formulated by comparison of the HDV and *CPEB3* ribozymes to search for novel HDV-like ribozymes in sequenced genomes (Figure 1b). They



3. Zaug AJ, Kent JR, Cech TR: **A labile phosphodiester bond at the ligation junction in a circular intervening sequence RNA.** *Science* 1984, **224**:574-8.
4. Forster AC, Symons RH: **Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites.** *Cell* 1987, **49**:211-20.
5. Hutchins CJ, Rathjen PD, Forster AC, Symons RH: **Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid.** *Nucleic Acids Res* 1986, **14**:3627-40.
6. Kuo MY, Sharmeen L, Dinter-Gottlieb G, Taylor J: **Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus.** *J Virol* 1988, **62**:4439-44.
7. Prody GA, Bakos JT, Buzayan JM, Schneider IR, Bruening G: **Autolytic processing of dimeric plant virus satellite RNA.** *Science* 1986, **231**:1577-80.
8. Saville BJ, Collins RA: **A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria.** *Cell* 1990, **61**:685-96.
9. Pabon-Pena LM, Zhang Y, Epstein LM: **Newt satellite 2 transcripts self-cleave by using an extended hammerhead structure.** *Mol Cell Biol* 1991, **11**:6109-15.
10. Zhang Y, Epstein LM: **Cloning and characterization of extended hammerheads from a diverse set of caudate amphibians.** *Gene* 1996, **172**:183-90.
11. Martick M, Scott WG: **Tertiary contacts distant from the active site prime a ribozyme for catalysis.** *Cell* 2006, **126**:309-20.  

F1000 Factor 6.5 *Must Read*  
 Evaluated by Fritz Eckstein 01 Aug 2006, Sabine Müller 14 Aug 2006, Eric Westhof 18 Aug 2006
12. Rupert PB, Ferre-D'Amare AR: **Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis.** *Nature* 2001, **410**:780-6.
13. Ferre-D'Amare AR, Zhou K, Doudna JA: **Crystal structure of a hepatitis delta virus ribozyme.** *Nature* 1998, **395**:567-74.
14. Scott WG, Finch JT, Klug A: **The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage.** *Cell* 1995, **81**:991-1002.
15. Das SR, Fong R, Piccirilli JA: **Nucleotide analogues to investigate RNA structure and function.** *Curr Opin Chem Biol* 2005, **9**:585-93.
16. Oyelere AK, Kardon JR, Strobel SA: **pK(a) perturbation in genomic Hepatitis Delta Virus ribozyme catalysis evidenced by nucleotide analogue interference mapping.** *Biochemistry* 2002, **41**:3667-75.
17. Jones FD, Strobel SA: **Ionization of a critical adenosine residue in the neurospora Varkud Satellite ribozyme active site.** *Biochemistry* 2003, **42**:4265-76.  

F1000 Factor 3.0 *Recommended*  
 Evaluated by David MJ Lilley 26 Mar 2003
18. Kuzmin YI, Da Costa CP, Cottrell JW, Fedor MJ: **Role of an active site adenine in hairpin ribozyme catalysis.** *J Mol Biol* 2005, **349**:989-1010.
19. Kuzmin YI, Da Costa CP, Fedor MJ: **Role of an active site guanine in hairpin ribozyme catalysis probed by exogenous nucleobase rescue.** *J Mol Biol* 2004, **340**:233-51.  

F1000 Factor 6.0 *Must Read*  
 Evaluated by Donald Burke 14 Sep 2004
20. Schmidt S, Beigelman L, Karpeisky A, Usman N, Sorensen US, Gait MJ: **Base and sugar requirements for RNA cleavage of essential nucleoside residues in internal loop B of the hairpin ribozyme: implications for secondary structure.** *Nucleic Acids Res* 1996, **24**:573-81.
21. Grasby JA, Mersmann K, Singh M, Gait MJ: **Purine functional groups in essential residues of the hairpin ribozyme required for catalytic cleavage of RNA.** *Biochemistry* 1995, **34**:4068-76.
22. Gilbert W: **Origin of life: the RNA world.** *Nature* 1986, **319**:618.
23. Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP: **RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension.** *Science* 2001, **292**:1319-25.  

F1000 Factor 8.0 *Exceptional*  
 Evaluated by Ronald Breaker 21 Sep 2001, Andrew Ellington 17 Jul 2002
24. Lee N, Bessho Y, Wei K, Szostak JW, Suga H: **Ribozyme-catalyzed tRNA aminoacylation.** *Nat Struct Biol* 2000, **7**:28-33.
25. Zhang B, Cech TR: **Peptide bond formation by in vitro selected ribozymes.** *Nature* 1997, **390**:96-100.
26. Tuerk C, Gold L: **Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** *Science* 1990, **249**:505-10.
27. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR: **Control of gene expression by a natural metabolite-responsive ribozyme.** *Nature* 2004, **428**:281-6.  

F1000 Factor 12.8 *Exceptional*  
 Evaluated by David MJ Lilley 18 Mar 2004, Richard Maraia 22 Mar 2004, Peter Unrau 24 Mar 2004, Samuel Butcher 24 Mar 2004, Eric Westhof 24 Mar 2004, Scott K Silverman 26 Mar 2004, Fritz Eckstein 29 Mar 2004, Burckhard Seelig 01 Apr 2004, Andrew D Sharrocks 02 Apr 2004, Janet Morrow 04 May 2004, Andres Jaschke 10 May 2004, John Burke 13 May 2004, Reinhard Luehrmann 17 May 2004
28. Collins JA, Irnov I, Baker S, Winkler WC: **Mechanism of mRNA destabilization by the glmS ribozyme.** *Genes Dev* 2007, **21**:3356-68.
29. Salehi-Ashtiani K, Luptak A, Litovchick A, Szostak JW: **A genome-wide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene.** *Science* 2006, **313**:1788-92.  

F1000 Factor 9.6 *Exceptional*  
 Evaluated by Michael Famulok 28 Sep 2006, Samuel Gunderson 13 Oct 2006
30. Gautheret D, Major F, Cedergren R: **Pattern searching/alignment with RNA primary and secondary structures: an effective descriptor for tRNA.** *Comput Appl Biosci* 1990, **6**:325-31.
31. Dsouza M, Larsen N, Overbeek R: **Searching for patterns in genomic data.** *Trends Genet* 1997, **13**:497-8.
32. Przybilski R, Graf S, Lescoute A, Nellen W, Westhof E, Steger G, Hammann C: **Functional hammerhead ribozymes naturally encoded in the genome of Arabidopsis thaliana.** *Plant Cell* 2005, **17**:1877-85.  

F1000 Factor 3.2 *Recommended*  
 Evaluated by Peter Unrau 21 Jun 2005, Daniel Gallie 17 Aug 2005
33. Martick M, Horan LH, Noller HF, Scott WG: **A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA.** *Nature* 2008, **454**:899-902.  

F1000 Factor 10.8 *Exceptional*  
 Evaluated by Robert Batey 16 Jul 2008, Thorsten Dieckmann 21 Jul 2008, Sabine Müller 23 Jul 2008, Fritz Eckstein 23 Jul 2008, Andres Jaschke 04 Aug 2008, Barry Stoddard 14 Aug 2008, Ian Willis 19 Aug 2008, Eric Westhof 22 Aug 2008
34. Webb CH, Riccitelli NJ, Ruminski DJ, Luptak A: **Widespread occurrence of self-cleaving ribozymes.** *Science* 2009, **326**:953.
35. Kartsogiannis V, Sims NA, Quinn JM, Ly C, Cipetic M, Poulton IJ, Walker EC, Saleh H, McGregor NE, Wallace ME, Smyth MJ, Martin TJ, Zhou H, Ng KW, Gillespie MT: **Osteoclast inhibitory lectin, an immune cell product that is required for normal bone physiology in vivo.** *J Biol Chem* 2008, **283**:30850-60.
36. Uhlenbeck OC: **A small catalytic oligoribonucleotide.** *Nature* 1987, **328**:596-600.
37. Perrotta AT, Been MD: **Cleavage of oligoribonucleotides by a ribozyme derived from the hepatitis delta virus RNA sequence.** *Biochemistry* 1992, **31**:16-21.