Published in final edited form as:

Nature. 2008 August 14; 454(7206): 899–902. doi:10.1038/nature07117.

A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA

Monika Martick^{1,*}, Lucas H. Horan^{1,*}, Harry F. Noller¹, and William G. Scott¹

¹ Center for Molecular Biology of RNA, University of California, Santa Cruz, California 95064, USA

Abstract

Structured RNAs embedded in the untranslated regions (UTRs) of messenger RNAs can regulate gene expression. In bacteria, control of a metabolite gene is mediated by the self-cleaving activity of a ribozyme embedded in its 5' UTR¹. This discovery has raised the question of whether gene-regulating ribozymes also exist in eukaryotic mRNAs. Here we show that highly active hammerhead ribozymes^{2,3} are present in the 3' UTRs of rodent C-type lectin type II (*Clec2*) genes^{4–7}. Using a hammerhead RNA motif search with relaxed delimitation of the non-conserved regions, we detected ribozyme sequences in which the invariant regions, in contrast to the previously identified continuous hammerheads^{8–10}, occur as two fragments separated by hundreds of nucleotides. Notably, a fragment pair can assemble to form an active hammerhead ribozyme structure between the translation termination and the poly-adenylation signals within the 3' UTR. We demonstrate that this hammerhead structure can self-cleave both *in vitro* and *in vivo*, and is able to reduce protein expression in mouse cells. These results indicate that an unrecognized mechanism of post-transcriptional gene regulation involving association of discontinuous ribozyme sequences within an mRNA may be modulating the expression of several CLEC2 proteins that function in bone remodelling and the immune response of several mammals.

The hammerhead ribozyme is a small, self-cleaving motif composed of a three-helical junction with a core of invariant nucleotides required for activity. To identify hammerhead ribozymes in mammalian mRNAs, we searched mRNA sequence databases using a pattern descriptor that allowed for insertions of up to 5,000 nucleotides at the ends of stem 1 or stem 3 (Supplementary Fig. 1)^{11,12}. Three hammerhead ribozymes were identified in the 3' UTRs of known rodent mRNAs. Two are found embedded in the transcripts of mouse *Clec2d* (osteoclast inhibitory lectin, also known as *Ocil*, *Clr-b* and *Clec2d8*)¹³ and its paralogue *Clec2e* (also known as *Clra* and *Clec2d7*)¹⁴, genes that belong to a group of phylogenetically related sequences within the natural killer receptor gene complex of chromosome 6. The third ribozyme is found in rat *CLEC2D11* (ref. 7)—a homologue of mouse *Clec2d*—which resides in the syntenic natural killer receptor gene complex region on chromosome 4. We extended our search to the genomic sequences of other organisms using the UCSC genome browser's comparative genomics tool¹⁵. Alignments using the natural killer receptor gene complex regions of mouse and rat led to the identification of nine candidate hammerhead ribozymes: four in the 3' regions of predicted rat, horse and platypus *Clec2*-like genes, and five in the unannotated regions of five other mammalian genomes (Supplementary Fig. 2).

Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.M. (mmartick@yahoo.com).

*These authors contributed equally to this work.

Author Contributions L.H.H. and M.M. did the sequence searches, designed the study, performed the experiments, analysed the data and wrote the manuscript.

Unlike most known self-cleaving RNA motifs that are contiguous^{8–10,16–19}, the hammerhead ribozymes identified here (referred to as CLEC2 ribozymes) are split into two fragments separated by a long ribozyme-unrelated insertion in the stem-1-capping loop. The insertion (which is 250 and 696 nucleotides in mouse *Clec2d* and *Clec2e*, respectively, and 145–1,739 nucleotides in other candidate ribozyme sequences) segregates the upstream substrate region, residing within 7–44 nucleotides of the stop codon, from the downstream enzyme fragment (Fig. 1 and Supplementary Fig. 2). When assembled from substrate and enzyme fragments, the secondary structure reflects a characteristic hammerhead ribozyme core of fifteen conserved nucleotides flanked by three helices. Comparison of all twelve CLEC2 ribozymes showed conservation of stem 2 as well as the presence of compensatory mutations maintaining stem 1 and the atypically long stem 3. Stem 1 and stem 2 contain secondary structure elements required to form a tertiary contact known to greatly enhance catalysis (Fig. 1 and Supplementary Fig. 2)^{3,20,21}. Six loop and bulge residues (Fig. 1)—necessary for the active structure stabilization and optimal catalysis in the *Schistosoma mansoni* hammerhead ribozyme²²—are also found in most CLEC2 ribozyme sequences. Taken together, the secondary structure features suggest that the 3' UTR-embedded hammerhead ribozyme motifs form catalytically active structures despite their discontinuity.

To test whether the hammerhead ribozymes embedded within the 3' UTRs possess catalytic activity, we examined the *in vitro* transcription products of mouse 3' UTRs from *Clec2d* (D-hcu, hammerhead containing UTR) and *Clec2e* (E-hcu). Cleavage of both 3' UTR RNAs occurs during the course of the transcription reaction (data not shown) and at the predicted hammerhead ribozyme cleavage sites (Fig. 2). The DM1-hcu construct—which is expected to create an inactive hammerhead ribozyme due to a triple mutation that prevents the 'substrate' from being cleaved—yields an intact 3' UTR transcript, thus confirming that the cleavage is due to the embedded hammerhead ribozyme activity. The mouse *Clec2e*-derived 3' UTR cleaves to a lesser extent than D-hcu, perhaps because of misfolding effects due to a ~450 nucleotide longer intervening sequence and/or weakened secondary structure in the stem 1 and stem 3 regions (Supplementary Fig. 3). These results demonstrate that the disconnected 'substrate' and 'enzyme' fragments form a fully active hammerhead ribozyme.

To determine whether the D-hcu sequence can reduce gene expression by forming an active ribozyme *in vivo*, it was incorporated into a dual-luciferase gene expression vector and assayed for activity in NIH 3T3 cells (Fig. 3). D-hcu reduces luciferase expression by 80% when placed downstream from the luciferase reporter (Fig. 3b). Nonspecific downregulation effects were ruled out because inversion of the same sequence (R-D-hcu) restores luciferase expression levels to that of the positive control. To verify that the reduction of gene expression was caused by the embedded ribozyme, mutations specifically targeting hammerhead ribozyme activity were introduced into D-hcu. Both a single transversion mutation at position G8 in the enzyme fragment (DM2-hcu), previously shown to compromise hammerhead ribozyme activity³, and a triple mutation in the substrate fragment (DM1-hcu) restore gene expression to control levels, indicating that the hammerhead ribozyme embedded in the mouse *Clec2d* 3' UTR is responsible for reducing reporter protein expression in mouse cells.

RNA from cells transfected with either wild-type (D-hcu) or mutant (DM1-hcu) dual luciferase reporter vectors was compared using reverse transcription–polymerase chain reaction (RT–PCR) to assess the effect of the discontinuous ribozyme-containing 3' UTR on mRNA integrity. A mixture of two primers, each recognizing one of two reporter genes transcribed from the same vector, was used for first-strand complementary DNA synthesis (Fig. 4a). Amplification of the region spanning the ribozyme cleavage site indicates that mouse *Clec2d* 3' UTR self-cleaves via the hammerhead ribozyme activity (Fig. 4b). The uncleaved product of D-hcu is at least tenfold reduced from that of DM1-hcu, in accordance with results of the protein expression assay (Fig. 4c). Previous studies have shown that ribozymes artificially engineered

into 3' UTRs can downregulate gene expression²³ by promoting rapid destruction of the transcript by cytoplasmic RNA degradation machinery²⁴. Consistent with these observations, the catalytic activity of mouse *Clec2d* 3' UTR correlates with the overall reduction of firefly luciferase transcript levels, as compared to those of the mutant UTR-containing mRNA, indicating that the cleaved *Clec2* transcript is degraded *in vivo*.

Structural conservation observed throughout the identified sequences, in conjunction with the tight association of the ribozyme with *Clec2*-like genes, suggests that the discontinuous ribozymes are orthologous to the mouse hammerhead motif characterized here. Within the rodent lineage, the rat *CLEC2D11* ribozyme sequence shares 86% homology with the mouse *Clec2d* ribozyme. Similarly, two predicted rat *Clec2d* paralogues are closely (61% and 67%) related to the mouse *Clec2e* ribozyme (Supplementary Fig. 2). These sequence relationships positively correlate with phylogenetic evidence for multiple duplication events of rodent *Clec2* genes⁷. Further genome searches using separate parts of the hammerhead ribozyme detected four additional mouse *Clec2* paralogues that contain substrate-like, but not enzyme, sequences in their 3' UTRs (data not shown). These partial motifs are also positioned proximally to the stop codon and share extensive primary sequence identity with the substrate regions of functional ribozymes of either mouse *Clec2d* or *Clec2e*. However, it is not clear if the incomplete ribozyme sequences are functional substrates for the *Clec2d*- and *Clec2e*-derived trans-acting ribozyme elements, or are non-functional evolutionary relics of the rodent *Clec2* gene expansion.

Identification of CLEC2 ribozymes in the genomes of mammalian species other than rodents (notably platypus) indicates that these ribozymes share an early mammalian ancestor. Active hammerhead motifs have been found in several divergent species, such as newts¹⁰, plants⁸ and viroids¹⁰. However, only the *Schistosoma* satellite DNA is known to encode hammerhead ribozyme sequences that have significant homology to the CLEC2 ribozymes^{3,18}. Although schistosomes parasitize rodents, their ribozymes are found in lineages more modern than the time of divergence of mouse and rat species^{18,25}, suggesting a horizontal gene transfer of the hammerhead sequence from host to the parasite genome as the probable scenario.

In eukaryotes, post-transcriptional regulation of gene expression through the 3' UTR involves an interaction between mRNA-stabilizing and -destabilizing protein-based processes²⁶. Together with evidence of cross-species secondary structure conservation, our results indicate that the CLEC2 hammerhead ribozymes have roles analogous to those of destabilizing protein factors. Paired with a yet undiscovered antagonistic process, the hammerhead motif may constitute a system of rapid regulation for CLEC2d, a transmembrane C-type lectin that has a dual role: the inhibition of natural killer (NK)-cell-mediated lysis through the interaction with the NK cell receptor NKR-P1, and the inhibition of osteoclast formation^{4,6}.

The core sequences of CLEC2 ribozymes resemble those of the smaller members of its ribozyme class, but the discontinuous arrangement of the substrate and enzyme regions exhibit properties previously associated only with catalytic introns. Although functioning through a fundamentally different mechanism and residing in transcripts distinct from processed mRNA, catalytic sequences of some self-splicing group I and II introns span long stretches of RNA¹⁹ on a scale analogous to that of the ribozymes described here. A question remains as to whether CLEC2 ribozymes can also process substrates encoded on detached transcripts similarly to RNase P, which possesses trans-cleaving properties unique among naturally occurring ribozymes¹⁹. Most self-cleaving RNAs, however, are considered to be compact sequences suitable for single-step functions such as multimer-to-monomer conversion of genomes in the course of viroid and satellite RNA replication, or for response to small molecule binding such as in metabolite-dependent gene regulation in bacteria. However, in the context of complex gene regulatory systems such as those of mammalian cells, the discontinuous

configuration of the CLEC2 ribozyme sequences may provide an opportunity for allosteric regulation.

METHODS SUMMARY

mRNA sequence data from the Ensembl database (release 48, <http://www.ensembl.org>) were searched using the RNABOB program^{11,12} with a descriptor (Supplementary Fig. 1) for the hammerhead ribozyme motif. Orthologues of the rodent CLEC2 ribozymes were identified using BLAT and the UCSC Genome Browser (<http://genome.ucsc.edu/>)^{15,27}. For *in vitro* RNA synthesis and analysis, the *Clec2* 3' UTR-containing plasmids were linearized and transcribed using T7 RNA polymerase at a pH of 8.0 and in the presence of 28 mM MgCl₂. Reverse transcriptase primer extensions and RNA sequencing were adapted for a [³²P]-end-labelled primer from previously published methods²⁸. For the expression assays, the NIH 3T3 mouse fibroblast cells were transiently transfected with *Clec2* 3' UTR-containing psiCHECK-2 reporter plasmids (Promega) and analysed using the Dual-Glo Luciferase Assay System (Promega). Firefly reporter expression was normalized using expression of the *Renilla* luciferase encoded on the same plasmid. RNA from the transfected cells was purified using Trizol (Sigma). For RT-PCR, cDNA was synthesized using primers specific to plasmid-derived transcripts using previous methods²⁸. PCR was performed with Accuprime Taq (Invitrogen). PCR products were resolved on an agarose gel and the product quantification was performed using ImageQuant 5.2 (Molecular Dynamics).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank M. Hall and K. Chakrabarti for assistance with cell culture, members of Haussler and Ares laboratories for sharing tissue culture space, M. Robertson for the transcriptase and A. Zahler for discussion. This work was supported by the National Institutes of Health grant R01043393.

References

1. Winkler WC, et al. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 2004;428:281–286. [PubMed: 15029187]
2. Khvorova A, Lescoute A, Westhof E, Jayasena SD. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature Struct Biol* 2003;10:708–712. [PubMed: 12881719]
3. Martick M, Scott WG. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* 2006;126:309–320. [PubMed: 16859740]
4. Zhou H, et al. Osteoclast inhibitory lectin, a family of new osteoclast inhibitors. *J Biol Chem* 2002;277:48808–48815. [PubMed: 12374791]
5. Plougastel B, Dubbelde C, Yokoyama WM. Cloning of *Clr*, a new family of lectin-like genes localized between mouse *Nkrp1a* and *Cd69*. *Immunogenetics* 2001;53:209–214. [PubMed: 11398965]
6. Carlyle JR, et al. Missing self-recognition of Ocil/Clr-b by inhibitory NKR-P1 natural killer cell receptors. *Proc Natl Acad Sci USA* 2004;101:3527–3532. [PubMed: 14990792]
7. Hao L, Klein J, Nei M. Heterogeneous but conserved natural killer receptor gene complexes in four major orders of mammals. *Proc Natl Acad Sci USA* 2006;103:3192–3197. [PubMed: 16492762]
8. Przybilski R, et al. Functional hammerhead ribozymes naturally encoded in the genome of *Arabidopsis thaliana*. *Plant Cell* 2005;17:1877–1885. [PubMed: 15937227]
9. Rojas AA, et al. Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets. *Nucleic Acids Res* 2000;28:4037–4043. [PubMed: 11024185]

10. Flores R, et al. Hammerhead ribozyme structure and function in plant RNA replication. *Methods Enzymol* 2001;341:540–552. [PubMed: 11582803]
11. Gautheret D, Major F, Cedergren R. Pattern searching/alignment with RNA primary and secondary structures: an effective descriptor for tRNA. *Comput Appl Biosci* 1990;6:325–331. [PubMed: 1701686]
12. Eddy, SR. RNABOB: a program to search for RNA secondary structure motifs in sequence databases. 2008. <http://selab.janelia.org/software.html>
13. Zhou H, et al. A novel osteoblast-derived C-type lectin that inhibits osteoclast formation. *J Biol Chem* 2001;276:14916–14923. [PubMed: 11278931]
14. Carlyle JR, et al. Molecular and genetic basis for strain-dependent NK1.1 alloreactivity of mouse NK cells. *J Immunol* 2006;176:7511–7524. [PubMed: 16751398]
15. Kent WJ, et al. The human genome browser at UCSC. *Genome Res* 2002;12:996–1006. [PubMed: 12045153]
16. Salehi-Ashtiani K, Luptak A, Litovchick A, Szostak JW. A genomewide search for ribozymes reveals an HDV-like sequence in the human *CPEB3* gene. *Science* 2006;313:1788–1792. [PubMed: 16990549]
17. Teixeira A, et al. Autocatalytic RNA cleavage in the human β -globin pre-mRNA promotes transcription termination. *Nature* 2004;432:526–530. [PubMed: 15565159]
18. Ferbeyre G, Smith JM, Cedergren R. Schistosome satellite DNA encodes active hammerhead ribozymes. *Mol Cell Biol* 1998;18:3880–3888. [PubMed: 9632772]
19. Serganov A, Patel DJ. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nature Rev Genet* 2007;8:776–790. [PubMed: 17846637]
20. Canny MD, et al. Fast cleavage kinetics of a natural hammerhead ribozyme. *J Am Chem Soc* 2004;126:10848–10849. [PubMed: 15339162]
21. Osborne EM, Schaak JE, Derose VJ. Characterization of a native hammerhead ribozyme derived from schistosomes. *RNA* 2005;11:187–196. [PubMed: 15659358]
22. Martick, M. PhD thesis. Univ. California; 2007. Structural Study of how the Hammerhead Ribozyme Solves the Problem of Catalysis. <http://www.proquest.com/>
23. Yen L, et al. Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* 2004;431:471–476. [PubMed: 15386015]
24. Meaux S, Van Hoof A. Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA* 2006;12:1323–1337. [PubMed: 16714281]
25. Lockyer AE, et al. The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. *Parasitology* 2003;126:203–224. [PubMed: 12666879]
26. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. *Nature Rev Mol Cell Biol* 2007;8:113–126. [PubMed: 17245413]
27. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res* 2002;12:656–664. [PubMed: 11932250]
28. Stern S, Moazed D, Noller HF. Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol* 1988;164:481–489. [PubMed: 2468070]

Figure 1. Sequence arrangement and secondary structure model of the rodent *Clec2d* hammerhead ribozymes

The mouse ribozyme sequence is shown in black and the rat ribozyme sequence length, single nucleotide, and base pair differences are denoted in red. The stop codon is shown in white. The substrate and enzyme sequences are shown on orange and blue backgrounds, respectively. The insertion sequence separating two ribozyme parts is abridged with a thick arrow. The predicted cleavage site (white arrowhead) is 3' of the active site cytosine (circled). The three-helical junction, composed of conserved (with the exception of 2.1 and 1.1) nucleotides (nt) that make catalytically important interactions, is shown in greater detail (together with canonical numbering scheme) in the inset. The small black arrowheads indicate conserved nucleotides of the catalytically important loop/bulge interactions.

Figure 2. The discontinuous hammerhead ribozyme-containing 3' UTRs (hcu) self-cleave *in vitro* Reverse transcriptase primer extensions of *in vitro* transcribed mouse *Clec2d* (D-hcu, lane 5) and *Clec2e* (E-hcu, lane 7) 3' UTRs show that both RNAs are cleaved at a single location (open arrowheads). D-hcu cleavage site 5' of U_{1.1} was identified by dideoxy sequencing of transcription products. Uncleaved E-hcu and triple mutant DM1-hcu are indicated with a black arrowhead. Asterisks indicate non-extended labelled primers.

Figure 3. The discontinuous mouse *Clec2* hammerhead ribozymes embedded in the 3' UTRs downregulate protein expression *in vivo*

a. Layout of the constructs (top) and the schematic of hammerhead ribozyme-containing 3' UTR (hcu) sequences used for expression. The colours of substrate (orange) and enzyme (blue) regions correspond to those of the structure in Fig. 1 and the insertion is highlighted by a thick line. The cleavage sites are indicated by white arrowheads and asterisks denote the positions of mutations. CHR (control hammerhead ribozyme) and CHRM (G8 to C8 mutant of CHR) are the controls; bp, base pairs. **b.** *In vivo* analysis of firefly luciferase protein expression from different 'hcu' constructs. Black denotes expression from the vector containing wild type 3' UTR; grey represents mutants and controls; and white indicates a vector without an insert as

a negative control. Relative firefly luciferase (*hLuc*⁺) expression was determined using a dual luciferase assay using the *Renilla* luciferase (*hRLuc*) expression for normalization. The results shown are means $\pm$ s.e.m. of triplicate experiments; the error bars for the vector value are averages of s.e.m. from three assays.

Figure 4. RT-PCR analysis of the *in vivo* expression products

a, Schematic of the relative positions of the RT-PCR primers. The hammerhead-specific cleavage site is indicated by an open arrowhead. The 368 base pair (bp) product acts as the internal control for transfection efficiency and the extent of transcription. **b**, RT-PCR of RNA isolated from NIH 3T3 cells that have been transfected with pD-hcu or pDM1-hcu. The asterisk indicates nonspecific PCR products. **c**, Quantification of the difference in PCR product intensity between D-hcu and DM1-hcu. The product intensity was used to determine the ratio of D-hcu/DM1-hcu at each listed cycle. Each value represents the mean $\pm$ s.d. of intensity ratios from three listed cycles.