

REPORTS

## A Genomewide Search for Ribozymes Reveals an HDV-Like Sequence in the Human *CPEB3* Gene

Kourosh Salehi-Ashtiani,<sup>1,2\*</sup> Andrej Lupták,<sup>1\*</sup> Alexander Litovchick,<sup>1</sup> Jack W. Szostak<sup>1†</sup>

Ribozymes are thought to have played a pivotal role in the early evolution of life, but relatively few have been identified in modern organisms. We performed an in vitro selection aimed at isolating self-cleaving RNAs from the human genome. The selection yielded several ribozymes, one of which is a conserved mammalian sequence that resides in an intron of the *CPEB3* gene, which belongs to a family of genes regulating messenger RNA polyadenylation. The *CPEB3* ribozyme is structurally and biochemically related to the human hepatitis delta virus (HDV) ribozymes. The occurrence of this ribozyme exclusively in mammals suggests that it may have evolved as recently as 200 million years ago. We postulate that HDV arose from the human transcriptome.

<sup>1</sup> Howard Hughes Medical Institute, Department of Molecular Biology, and Center for Computational and Integrative Biology (CCIB), 7215 Simches Research Center, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA.

<sup>2</sup> Center for Cancer Systems Biology (CCSB) and Department of Cancer Biology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA.

\* These authors contributed equally to this work. 

†To whom correspondence should be addressed. E-mail: [szostak@molbio.mgh.harvard.edu](mailto:szostak@molbio.mgh.harvard.edu)

RNAs play diverse roles in biology, including a variety of regulatory and catalytic functions that together provide support for the "RNA world" hypothesis (1). Catalytic RNAs, or ribozymes, include naturally occurring phosphoryl transferases and the ribosomal peptidyl transferase (2), as well as in vitro selected RNAs (3). Recent discoveries of regulatory cellular ribozymes, such as the bacterial cofactor-dependent ribozyme GImS (4) and the eukaryotic cotranscriptional cleavage (CoTC) ribozyme (5), have raised interesting questions about the prevalence and evolution of catalytic RNAs in higher eukaryotes.

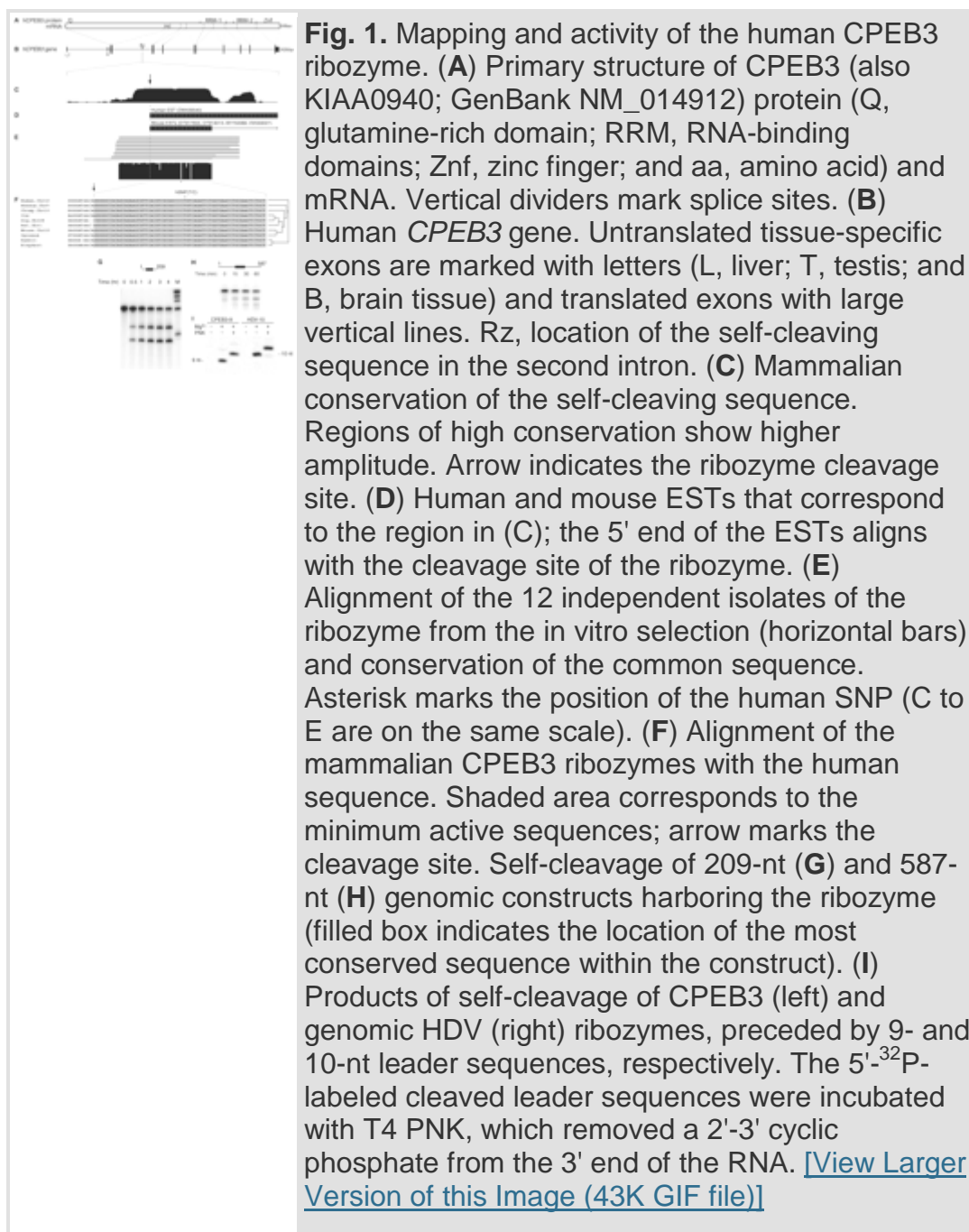
Most of the reactions carried out by naturally occurring ribozymes involve the chemistry of the phosphodiester RNA backbone. Among phosphotransferases, self-cleaving ribozymes form perhaps the most diverse subgroup, which includes the hammerhead (6), hairpin (7), and HDV ribozymes (8); the *Neurospora* Varkud satellite (VS) motif (9); the bacterial GImS ribozyme (4); the eukaryotic CoTC motif (5); and the group I intron-like

ribozyme GIR 1 ([10](#)). Only two of these are associated with mammals: the CoTC motif, the only known self-cleaving human ribozyme, with orthologs in other primates ([5](#)), and the HDV ribozyme, the only known self-cleaving ribozyme encoded by a human pathogen, with no known evolutionary orthologs. The sparse occurrence of self-cleaving motifs in humans, and mammals in general, could indicate that this class of catalytic RNA has been lost over the course of evolution or, alternatively, that these RNAs have escaped detection. All self-cleaving ribozymes found to date were discovered by careful analysis of transcripts of single genes or RNA genomes of pathogens. Notably, GImS was first identified as a riboswitch through computational analysis of conserved bacterial RNA structures and was later observed to self-cleave in the presence of a cofactor ([4](#)). CoTC was identified through detailed studies of cotranscriptional termination of  $\beta$ -globin mRNA synthesis ([5](#)).

To identify ribozymes encoded in the human genome, we devised an in vitro selection scheme for the isolation of self-cleaving sequences without imposing fixed cleavage target sites. In this scheme, a genomic library is constructed in which library elements are uniform in size [ $\sim$ 150 nucleotides (nt)] and are flanked by polymerase chain reaction (PCR) primer sequences [fig. S1 ([11](#))]. The library was converted into a single-stranded form, circularized by splint-ligation, and converted back to double-stranded DNA to form a "relaxed" double-stranded circle that can be in vitro transcribed ([11](#)). Rolling-circle transcription produced tandemly repeated RNAs that preserved the covalent linkage of the entire transcribed sequence, even after self-cleavage ([12](#)). If a sequence encoded a self-cleaving motif, it could self-cleave under appropriate conditions to produce unit-length copies, as well as kinetically trapped misfolded intermediates such as dimers, trimers, and so on. The size difference between the cleaved multimers and uncleaved sequences provided a positive selection criterion for the enrichment of active molecules. Dimers, which contain one intact copy of the sequence, were isolated by polyacrylamide gel electrophoresis (PAGE) purification, reverse-transcribed, and amplified to reinitiate the cycle.

We started the selection with about 580 ng of RNA (or  $10^{15}$  nucleotides) in concatemeric form. We allowed the RNA to self-cleave for 1 hour in the presence of 5 mM  $MgCl_2$  at near-physiological monovalent salt concentration, pH, and temperature (140 mM KCl, 10 mM NaCl, 50 mM Tris-HCl, pH 7.4, 37°C). After separation by denaturing PAGE, RNA fragments corresponding to dimers ( $\sim$ 400 nt) were isolated, reverse-transcribed, and amplified to reinitiate the cycle as outlined above. After six rounds of selection, self-cleavage could be detected in the library. Activity increased up to round 12, when the library became dominated by a small number of active sequences that exhibited robust self-cleavage in the presence of physiological concentrations of  $Mg^{2+}$  (0.5 to 1 mM  $MgCl_2$ ). The concatemeric RNA of this and preceding active rounds cleaved to form fragments that differed in length by about 200 nt, the unit size of the library. When the kinetically trapped trimers and dimers were isolated and renatured, they self-cleaved to yield smaller products (fig. S1). Thus, the selected library displayed properties in accordance with the experimental design and consistent with the presence of a single cleavage site per monomeric unit.

Cloning and sequencing of the round 12 library (fig. S2) revealed four self-cleaving ribozymes associated with the following human genes: olfactory receptor *OR4K15*, insulin-like growth factor 1 receptor (*IGF1R*), a LINE 1 retroposon, and the cytoplasmic polyadenylation element-binding protein 3 (*CPEB3*) (Fig. 1A and fig. S3). We studied the *CPEB3* ribozyme because it occurs within the transcript of a single-copy gene and because it is highly conserved (Fig. 1, B and C).

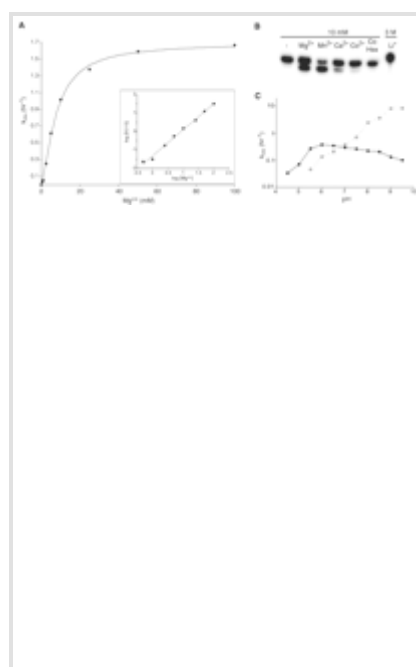


The isolated CPEB3 clones (including clones from round 10 and 11) originated from 12 independent progenitor sequences (i.e., independent DNase I fragments) of the library (fig. S4). The region of sequence overlap contains the minimal sequence necessary for self-cleavage activity and also delineates the local conservation boundaries among different species in this region of the intron (Fig. 1, C and E). The sequence alignment also revealed several mutations that are not present in the genome. Rather than having randomly accumulated during the repeated PCR amplification cycles, some of these mutations appear to have been highly selected (Fig. 1E and fig. S5). To rule out the possibility that the cleavage activity depends on the mutations that had been accumulated in the course of the selection, we tested the activity of RNA transcribed from a genomic DNA segment amplified directly from human genomic DNA. We tested the putative ribozyme sequence for self-cleavage in the absence of the selection PCR primer sequences and demonstrated that the 209-nt genomic sequence alone was sufficient to carry out self-cleavage (Fig. 1G). Additional 5' and 3' flanking sequences do not interfere with the activity of the ribozyme, measured as a first-order rate constant ( $k_{\text{obs}}$ ) of the RNA self-cleavage reaction ( $k_{\text{obs}} = 0.69 \pm 0.03 \text{ hour}^{-1}$ ,  $t_{1/2} = 1.00 \pm 0.04$  hours). Furthermore, we designed a set of primers to amplify a 587–base pair (bp) region that included the CPEB3 ribozyme and ~250 bp of flanking sequences on each side. Transcripts made from this genomic segment showed self-cleavage (Fig. 1H). Conversely, shortening the ribozyme to the region of highest conservation allowed us to narrow the sequence to 81 nt, to map precisely the site of cleavage, and to analyze the products of self-cleavage. The 81-nt construct cleaved to yield fragments of 9 nt upstream and 72 nt downstream of the cleavage site. Incubation of the products of the transesterification reaction with polynucleotide kinase (PNK), in the absence of adenosine triphosphate, shifted the mobility of the upstream 9-nt fragment; this change suggested that the upstream fragment contains a 2'-3' cyclic phosphate and that the downstream 72-nt fragment has a 5' terminal hydroxyl (Fig. 1I). We confirmed the identity of the products of self-cleavage by noting that the downstream sequence could be phosphorylated at its 5' end with PNK and that the upstream sequence was not a substrate for T4 RNA ligase. Altogether these experiments demonstrate that the 72-nt core of the sequence is sufficient to carry out self-cleavage, that the ribozyme can function in the presence of its native flanking sequences, and that the chemistry of self-cleavage most likely proceeds via a nucleophilic attack of a 2' hydroxyl on the adjacent phosphate, which yields a 2'-3' cyclic phosphate and a 5' terminal hydroxyl.

BLAST and BLAT analyses (13, 14) demonstrate that the CPEB3 ribozyme sequence, like the *CPEB3* gene itself, is found as a single copy in the genome and is highly conserved in all mammalian species examined, including opossum (a marsupial) (Fig. 1F). The ribozyme resides in a large intron (~46 kbp in human, ~37 kbp in rat, and ~35 kbp in mouse) about 10 to 14 kbp upstream of the second coding exon in all mammals except opossum, where the distance to the exon is about 25 kbp. Because the CPEB3 ribozyme is present in eutheria (placental mammals) and metatheria (marsupials), it is likely to have appeared at least 130 million years ago, before the two groups diverged. We found no ortholog of the ribozyme in orthologs of the *CPEB3*

gene in nonmammalian vertebrates, which suggests a mammalian origin for this sequence and indicates that it is likely to have appeared less than ~200 million years ago. The high conservation of the ribozyme sequence among mammals suggests that the sequence is functionally important.

The CPEB3 ribozyme requires divalent metal ions for its cleavage-transesterification. The ribozyme has a relatively low magnesium requirement, half-saturating at ~7.8 mM with a Hill coefficient of 1.4, which suggests that some degree of cooperativity is involved in the metal-dependent catalysis (Fig. 2A). The ribozyme efficiently self-cleaves in the presence of  $Mn^{2+}$ ,  $Mg^{2+}$ , and  $Ca^{2+}$ ; less efficient cleavage is observed in the presence of  $Co^{2+}$ ; and no products are detectable in cobalt (III) hexamine, an exchange-inert structural analog of magnesium (II) hexahydrate (Fig. 2B). When subjected to high-resolution PAGE, the products of the reaction comigrate, which suggests a common function for the divalent metal ions in catalysis. Given our data, the simplest model for the metal ion requirement in CPEB3 self-cleavage is that a hydrated divalent metal ion is required to accelerate the transesterification reaction. In some self-cleaving RNAs, such as the hammerhead, hairpin, and VS ribozymes, but not HDV, high concentrations of monovalent metal ions can promote catalysis at rates similar to magnesium-promoted scission (15). When we tested the CPEB3 ribozyme in the presence of 3 M  $Li^+$ , we did not detect self-cleavage (Fig. 2B), which suggests that CPEB3 ribozyme uses a different catalytic mechanism.



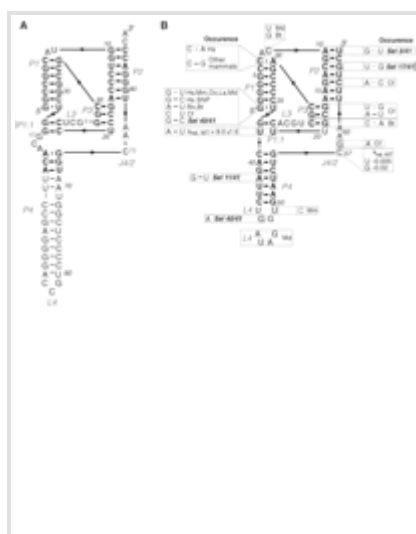
**Fig. 2.** Biochemical characterization of the CPEB3 ribozyme. (A)  $Mg^{2+}$  dependence of the self-cleavage kinetics. The cleavage rate reaches half-maximum at 7.8 mM  $Mg^{2+}$ . (Insert) Hill analysis of the dependence shows a slope of ~1.4, indicating slight cooperativity in  $Mg^{2+}$  binding ( $f$ , fraction of activity at saturating  $Mg^{2+}$ ). (B) Metal ion requirement for ribozyme activity. Identity and concentration of the metals are indicated above the gel lanes [CoHex, cobalt (III) hexamine]. CPEB3 ribozyme was incubated at 37°C for 10 min. (C) pH-rate profile (22°C, 10 mM  $Mg^{2+}$ ) of the CPEB3 ribozyme (filled circles) and a pool of in vitro selected self-cleaving ribozymes (open circles) (16). [\[View Larger Version of this Image \(13K GIF file\)\]](#)

A pH-rate profile of the CPEB3 ribozyme indicates that in 5 mM  $MgCl_2$  the activity of the ribozyme is almost constant between pH 5.5 and 8.5, dropping only at higher or lower pHs (Fig. 2C). This profile is consistent with the presence of two functional groups with distinct  $pK_a$  values (where  $K_a$  is the



acid dissociation constant) involved in rate-limiting proton transfers. For comparison, we examined the pH profile of a pool of *in vitro* selected self-cleaving ribozymes that we had previously isolated from random sequences (16). About 83% of this pool consists of new ribozymes, and 17% are diverse hammerhead-like sequences. In contrast to the CPEB3 ribozyme, the activity of this pool increased monotonically with pH, which suggests that acidic apparent  $pK_a$  values are not commonly found among self-cleaving ribozymes. Furthermore, when we carried out a cleavage reaction in 80%  $D_2O$ , the cleavage rate was less than that in 100%  $H_2O$  by a factor of  $\sim 1.7$ . The relatively flat kinetic pH profile and a solvent kinetic isotope effect support a catalytic model in which the rate-limiting step involves at least one proton transfer.

Because the biochemical properties of the CPEB3 ribozyme resemble those of the HDV ribozymes, we next investigated whether these ribozymes have structural similarities. The HDV RNA genome encodes two self-cleaving motifs, which reside in different locations and on opposite strands of the HDV RNA genome and are used to cut nascent rolling-circle replicated RNAs into genome-size copies (8, 17). The ribozymes fold into similar secondary structures characterized by a nested double pseudoknot (18, 19), but, with the exception of several conserved nucleotides, they diverge in primary sequence. The sequence of the CPEB3 ribozyme is not similar to either of the HDV ribozymes; however, the CPEB3 sequence can be folded into an HDV-like secondary structure without violating known HDV structural parameters (Fig. 3). Indeed, the experimentally determined cleavage site of the CPEB3 ribozyme (Fig. 1F) coincides precisely with the cleavage site predicted by this model. Like the HDV ribozyme, the CPEB3 ribozyme can tolerate nucleotide changes immediately upstream of its cleavage site. The homologous sequences in other mammals are all consistent with the HDV-like secondary structure (Fig. 3B). The mutations that appear to have been selectively accumulated during the course of the selection [from 41 sequenced CPEB3 clones (fig. S5)] would result in either conservative changes such as Watson-Crick base pairs to wobble pairs, or changes in a distal loop (Fig. 3B).



**Fig. 3.** Secondary structure of the genomic HDV ribozyme (A) and the human CPEB3 ribozyme (B). Names of ribozyme segments are boxed (P, paired; L, loop; and J, joining region). Numbers represent positions in the chain relative to the cleavage site (marked by 5'). (B) Variants of the CPEB3 ribozyme found in mammals (Hs, human; Mm, mouse; Rn, rat; Oc, rabbit; Cf, dog; La, elephant; Bt, cow; and Md, opossum) and among 41 sequenced clones from the human genomic selection described in the text (Sel). *In vitro* self-cleavage rates of U38A and C57 mutants are shown relative to wild-type ribozyme.

[\[View Larger Version of this Image \(29K GIF file\)\]](#)

To test whether the secondary structure of the CPEB3 ribozyme is HDV-like, we carried out covariation and mutational analyses at CPEB3 positions that are critical for activity in HDV (fig. S7). Disruption of the last base pair of the P3 helix, G19-C27, dramatically lowered the cleavage rate, whereas ribozymes in which the base pair was reversed to C19-G27 retained the wild-type rate. Similarly, in the P1 region, disruption of the third base-pair (G3-C34) to form a C3-C34 mispair lowered the observed cleavage rate by a factor of 200, whereas a G3-G34 pair reduced it by a factor of 4. Reversing the base pair to C3-G34 increased the cleavage rate of the ribozyme by a factor of 2.7 (fig. S7). The 3' end of the P1 helix forms a wobble pair in the human sequence (C7-A30) and a Watson-Crick pair (C7-G30) in all other mammals (Fig. 3B). Sequences that differ from the human sequence only at this position (e.g., rabbit and elephant, Fig. 1F) appear to self-cleave several times faster, which supports a structural role of the P1 helix.

The 5' end of the P1 helix is anchored in the ribozyme active site, which retains the 3' product of the cleavage reaction (18). In the CPEB3 ribozyme, the first nucleotide downstream of the cleavage site forms a wobble pair (G1-U36), as do the corresponding bases in HDV ribozymes. However, the human CPEB3 ribozyme has a single-nucleotide polymorphism (SNP) (U36C), and there are several mammalian orthologs that also form a Watson-Crick base pair at this position and retain cleavage activity (Figs. 1F and 3B and fig. S7).

The secondary structure proposed for CPEB3 ribozyme defines nonhelical elements predicted to be important for RNA self-cleavage. In HDV, a nucleobase, C75 in the genomic and C76 in the antigenomic ribozyme, forms part of the active site and has been proposed to be directly involved in proton shuttling during transesterification (18, 20–23). The CPEB3 ribozyme has a nucleobase C57 at the corresponding position. As in the HDV ribozymes, mutation of this residue to uridine or guanosine decreases the cleavage rate by at least a factor of 200 (Fig. 3B).

To define the critical regions of the CPEB3 ribozyme, we performed phosphorothioate interference mapping, in which sulfur substitution for nonbridging oxygens in the RNA backbone can disrupt key interactions within the ribozyme (24). The interference mapping of the CPEB3 ribozyme revealed two regions (nucleotides 20 to 23 and 58 to 61 downstream of the cleavage site) where such sulfur substitutions inactivated the ribozyme. Mapping of these positions onto the proposed HDV-like model (fig. S7) placed the phosphorothioate interference sites at positions corresponding to those previously mapped in the genomic HDV ribozyme (25) and to solvent-inaccessible backbone positions of the ribozyme core (26).

The HDV ribozyme folds into an intricate secondary structure that allows two stacked helical structures, P1-P1.1-P4 and P2-P3, respectively, to pack next to each other (18, 22). The stacked helices are held together on one end by

the cross-over strands J1/2 and J3/1. The other end of the P2-P3 helical structure is held in place through the 2 bp of the P1.1 stem, which defines the second nested pseudoknot. The P1.1 minihelix forms one side of the active site of the ribozyme; hence, disruption of this interaction decreases the cleavage rate in both genomic and antigenomic HDV ribozymes ([19](#)). It is surprising that the CPEB3 ribozyme only forms 1 bp of the P1.1 interaction (C22-G37). The second pair forms a weak U21-U38 interaction that can be strengthened with a U38A mutation, thus forming a 2-bp P1.1 and increasing the cleavage rate by almost 10-fold ([Fig. 3B](#)).

There are five expressed sequence tags (ESTs) (one from human liver and four from three different murine tissues) in public databases that overlap with the conserved region of the self-cleaving sequence, all of which share the same 5' terminal sequence that maps to the CPEB3 ribozyme cleavage site ([Fig. 1D](#)). The existence of these ESTs suggests that the CPEB3 ribozyme is active in vivo and that the downstream product of self-cleavage is sufficiently stable that it can be subsequently detected by EST mapping. To investigate in vivo expression and self-cleavage of the ribozyme, we used reverse transcriptase (RT)–PCR to analyze human and murine total RNA extracts prepared in the absence of divalent metal ions from several tissues ([fig. S9](#)) ([11](#)). Mapping by rapid amplification of 5' cDNA ends (5' RACE) of those extracts revealed that the site of self-cleavage is the same as in the ESTs and at the position mapped in vitro, which together indicate that the ribozyme is expressed and self-cleaves in vivo ([fig. S8](#)).

The CPEB3 ribozyme is an example of a class of self-cleaving ribozyme previously seen only in the HDV. The HDV ribozymes are structurally ([18](#), [22](#)) and informationally ([27](#), [28](#)) complex. Thus, in contrast to the hammerhead fold ([16](#)), the HDV fold is unlikely to have arisen independently multiple times. In further support of this suggestion, we found no sequences that could fold into an HDV-like secondary structure among the self-cleaving sequences isolated from a random pool ([16](#)) or among conserved mammalian sequences. On the basis of these observations, we hypothesize that HDV ribozymes and the CPEB3 ribozyme are evolutionarily related and that HDV, a single-stranded RNA satellite virus for hepatitis B (HBV) infection, arose from the human transcriptome. The reverse scenario—i.e., that the CPEB3 ribozyme came from HDV—is less likely. In order for HDV to transfer the ribozyme to mammals, the transfer would have to have occurred early in mammalian evolution (~200 million years ago), and ancestral versions of HDV and HBV would have to have existed as well. However, although HBV has been isolated from other animals, even nonmammals, and may have ancient origins, HDV has been isolated only from human tissues. The absence of HDV isolates from other animals suggests that HDV is of recent origin and that it arose from the human transcriptome, acquiring both the delta antigen protein ([29](#)) and the self-cleaving ribozyme from its host. Furthermore, if the HDV fold is exclusive to mammalian genomes and pathogens, then it is unlikely to be a descendant of the RNA world, which implies that structurally complex ribozymes have evolved in modern protein-dominated organisms. The rapidly evolving intron sequences may have been a fertile ground for the emergence of novel regulatory RNAs.



The CPEB3 ribozyme seems to be optimized to cleave more slowly than the HDV ribozymes; this inference is based on the conservation of a weak P1.1 stem among all mammalian sequences. We speculate that this slow cleavage activity allows normal splicing of the CPEB3 pre-mRNA to occur under basal conditions. The ribozyme is located about 11.5 kbp upstream of the third CPEB3 exon, so that a transcribing RNA polymerase will take about 10 min to reach the next intron-exon junction after synthesizing the CPEB3 ribozyme (30). Thus, most of the time, the spliceosome will join the exons before intron cleavage. If, however, the rate of ribozyme self-cleavage is up-regulated, the number of truncated CPEB3 pre-mRNAs could increase significantly. The CPEB3 ribozyme sequences in all species examined retain the weak P1.1 interaction, which suggests that slow cleavage is a positively selected feature of this ribozyme. Because there are many possible ways to decrease the cleavage rate, conserving this specific way of lowering the rate suggests that a trans-acting factor interacting at or near this site could readily stimulate the activity of the ribozyme. Rapid degradation of the cleaved RNA fragments could switch off the expression of the CPEB3 protein. Alternatively, cleavage of the CPEB3 pre-mRNA, perhaps followed by subsequent processing events, might produce smaller mRNAs that could be translated into truncated versions of the protein.

CPEB3 is a member of a protein family that regulates local polyadenylation of mRNAs in the cytoplasm of, among other tissues, neurons and oocytes (31). The CPEB3 full mRNA has been detected in human brain, skeletal muscle, and heart and to a lesser degree in liver, kidney, testis, and ovary tissues (32, 33). In mouse hippocampus, the gene is up-regulated transiently after induction of seizure, and it has been implicated in long-term potentiation (33). CPEB3 differs from other members of the CPEB family by its glutamine-rich N-terminal domain, which is encoded upstream of the CPEB3 ribozyme (Fig. 1, A and B). The function of this domain is unknown, but the glutamine-rich N-terminal domain of the *Aplysia* CPEB protein has been shown to aggregate into a prionlike structure that in turn makes the protein more active in promoting polyadenylation and that may play a role in the molecular basis for long-term facilitation (34, 35).

The four ribozymes isolated from our selection most likely represent only a small fraction of all cellular ribozymes. Our selection operated under a number of constraints including defined limits for size, cleavage rate, and selectable biochemical properties. For instance, ribozymes larger than ~150 nt could not have been isolated from our library. By changing these parameters and by including cellular factors such as small molecules, it may be possible to isolate additional genomic ribozymes, some of which may be modulated by cofactors, as is the case with the recently discovered bacterial cofactor-dependent ribozyme and the eukaryotic CoTC ribozyme (4, 5).

## References and Notes

- 1. R. F. Gesteland, T. R. Cech, J. F. Atkins, Eds., *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 3, 2006).

- 2. M. J. Fedor, J. R. Williamson, *Nat. Rev. Mol. Cell Biol.* **6**, 399 (2005). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 3. D. S. Wilson, J. W. Szostak, *Annu. Rev. Biochem.* **68**, 611 (1999). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 4. W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins, R. R. Breaker, *Nature* **428**, 281 (2004). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 5. A. Teixeira *et al.*, *Nature* **432**, 526 (2004). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 6. A. C. Forster, R. H. Symons, *Cell* **50**, 9 (1987). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 7. A. Hampel, R. Tritz, *Biochemistry* **28**, 4929 (1989). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 8. M. Y. Kuo, L. Sharmeen, G. Dinter-Gottlieb, J. Taylor, *J. Virol.* **62**, 4439 (1988). [\[Abstract/Free Full Text\]](#)
- 9. B. J. Saville, R. A. Collins, *Cell* **61**, 685 (1990). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 10. H. Nielsen, E. Westhof, S. Johansen, *Science* **309**, 1584 (2005). [\[Abstract/Free Full Text\]](#)
- 11. Materials and methods are available as supporting material on Science Online.
- 12. S. L. Daubendiek, E. T. Kool, *Nat. Biotechnol.* **15**, 273 (1997). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 13. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 14. A. S. Hinrichs *et al.*, *Nucleic Acids Res.* **34**, D590 (2006). [\[Abstract/Free Full Text\]](#)
- 15. J. B. Murray, A. A. Seyhan, N. G. Walter, J. M. Burke, W. G. Scott, *Chem. Biol.* **5**, 587 (1998). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 16. K. Salehi-Ashtiani, J. W. Szostak, *Nature* **414**, 82 (2001). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 17. L. Sharmeen, M. Y. Kuo, G. Dinter-Gottlieb, J. Taylor, *J. Virol.* **62**, 2674 (1988). [\[Abstract/Free Full Text\]](#)
- 18. A. R. Ferré-D'Amaré, K. Zhou, J. A. Doudna, *Nature* **395**, 567 (1998). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 19. T. S. Wadkins, A. T. Perrotta, A. R. Ferre-D'Amare, J. A. Doudna, M. D. Been, *RNA* **5**, 720 (1999). [\[Abstract\]](#)
- 20. S. Nakano, D. M. Chadalavada, P. C. Bevilacqua, *Science* **287**, 1493 (2000). [\[Abstract/Free Full Text\]](#)
- 21. S. R. Das, J. A. Piccirilli, *Nat. Chem. Biol.* **1**, 45 (2005). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 22. A. Ke, K. Zhou, F. Ding, J. H. Cate, J. A. Doudna, *Nature* **429**, 201 (2004). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)
- 23. A. T. Perrotta, I. Shih, M. D. Been, *Science* **286**, 123 (1999). [\[Abstract/Free Full Text\]](#)
- 24. E. L. Christian, M. Yarus, *J. Mol. Biol.* **228**, 743 (1992). [\[CrossRef\]](#) [\[ISI\]](#) [\[Medline\]](#)

- 25. N. S. Prabhu, G. Dinter-Gottlieb, P. A. Gottlieb, *Nucleic Acids Res.* **25**, 5119 (1997). [[Abstract/Free Full Text](#)]
- 26. S. P. Rosenstein, M. D. Been, *Biochemistry* **35**, 11403 (1996). [[CrossRef](#)] [[ISI](#)] [[Medline](#)]
- 27. A. Nehdi, J. P. Perreault, *Nucleic Acids Res.* **34**, 584 (2006). [[Abstract/Free Full Text](#)]
- 28. M. Legiewicz, A. Wichlacz, B. Brzezicha, J. Ciesiolka, *Nucleic Acids Res.* **34**, 1270 (2006). [[Abstract/Free Full Text](#)]
- 29. R. Brazas, D. Ganem, *Science* **274**, 90 (1996). [[Abstract/Free Full Text](#)]
- 30. D. S. Ucker, K. R. Yamamoto, *J. Biol. Chem.* **259**, 7416 (1984). [[Abstract/Free Full Text](#)]
- 31. Y. S. Huang, J. D. Richter, *Curr. Opin. Cell Biol.* **16**, 308 (2004). [[CrossRef](#)] [[ISI](#)] [[Medline](#)]
- 32. T. Nagase *et al.*, *DNA Res.* **6**, 63 (1999). [[Abstract](#)]
- 33. M. Theis, K. Si, E. R. Kandel, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9602 (2003). [[Abstract/Free Full Text](#)]
- 34. K. Si *et al.*, *Cell* **115**, 893 (2003). [[CrossRef](#)] [[ISI](#)] [[Medline](#)]
- 35. K. Si, S. Lindquist, E. R. Kandel, *Cell* **115**, 879 (2003). [[CrossRef](#)] [[ISI](#)] [[Medline](#)]
- 36. We thank L. Yen, G. Short, E. Meisami, E. Weiler, D. Treco, R. Bruckner, S. Schalm, J. Dennis, L. Davidoff, J. Ichida, J. Carothers, R. Larralde, and members of the Szostak laboratory for useful discussions and/or critical comments on the manuscript; P. Svec, D. Levin, and L. Almeida for administrative and laboratory support; and the Massachusetts General Hospital DNA core facility. Supported by NIH grants GM53936 and HL66678 to J.W.S. and the Howard Hughes Medical Institute.

---

## Supporting Online Material