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 dominated 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), IGFR1 (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 Arabadopsis 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...
identified a remarkable number of catalytically active HDV-like ribozymes in eukaryotic genomes across diverse phyla, including insects (*Anopheles gambiae*), invertebrates (*Pristionchus pacificus*, *Caenorhabditis japonica*, and *Stronglyocentrotus purpuratus*), and basal vertebrates (*Branchiostoma floridae* and *Petromyzon marinus*). Amazingly, most species contained more than one ribozyme or the ribozyme existed in multiple copies of the genome. Most notable are the nematodes. The *P. pacificus* genome has 32 apparent HDV-like ribozymes, whereas *C. japonica* harbors 122 apparent ribozymes. Most of these ribozymes tend to be associated with retrotransposons, but some of the ribozymes, including those from *A. gambiae*, are found in expressed sequence tags in or near predicted genes. They are expressed at different time points during development and may have a biological role in gene regulation. It is somewhat surprising that similar ribozymes were not identified in related model organisms, such as *Caenorhabditis elegans* and *Drosophila melanogaster*. It is likely that ribozymes will eventually be discovered in these and other organisms as the structural descriptors are further refined.

**Future directions**

It is clear that ribozymes are prevalent in eukaryotic genomes, but several important questions remain. In most cases, the biological function of eukaryotic small self-cleaving ribozymes has not been explored. It is now imperative to figure out why they are there and what they are doing. As new ribozymes are identified in genetically tractable model organisms, it will be possible to assess their functional role and hopefully identify general principles of eukaryotic ribozyme function. Are other structural variants of other self-cleaving ribozymes, such as the VS or hairpin ribozyme, hiding in eukaryotic genomes as well? What about new ribozyme classes, not just those related to previously characterized ribozymes? The in vitro selection strategy and computational tools are now in place to quickly answer these questions. The existence of functional hammerhead ribozymes with large discontinuous segments suggests the intriguing possibility that functional ribozymes may be built from multiple RNA sequences. Indeed, nearly all small self-cleaving ribozymes have been engineered to cleave RNA substrates in trans [1,36,37]. There are many non-coding regulatory RNAs found in all kingdoms of life. Perhaps some of them anneal to targets to form RNA enzymes, leading to self-cleavage. If so, they could possibly be the evolutionary precursor of modern small RNA silencing pathways. One thing is clear: RNA-guided RNA cleavage, whether catalyzed by protein or by RNA, is fundamental to life as we know it.

**Abbreviations**

Clec2, C-type lectin domain family 2; CPEB3, cytoplasmic polyadenylation element binding protein 3; HDV, hepatitis delta virus; UTR, untranslated region; VS, Varkud satellite.

**Competing interests**

The author declares that he has no competing interests.

**Acknowledgments**

I thank John Pagano, Brian Farley, Ruth Zearfoss, Nick Rhind, and Mary Munson for helpful comments. Work in my lab is funded by National Institutes of Health grant R01GM081422.

**References**


