

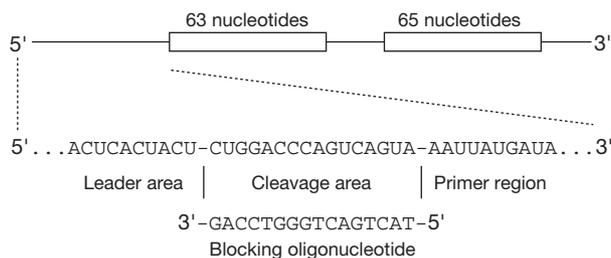
# In vitro evolution suggests multiple origins for the hammerhead ribozyme

Kourosh Salehi-Ashtiani & Jack W. Szostak

Howard Hughes Medical Institute, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

The hammerhead ribozyme was originally discovered in a group of RNAs associated with plant viruses, and has subsequently been identified in the genome of the newt (*Notophthalmus viridescens*)<sup>1,2</sup>, in schistosomes<sup>3</sup> and in cave crickets (*Dolichopoda* species)<sup>4</sup>. The sporadic occurrence of this self-cleaving RNA motif in highly divergent organisms could be a consequence of the very early evolution of the hammerhead ribozyme<sup>4,5</sup>, with all extant examples being descended from a single ancestral progenitor. Alternatively, the hammerhead ribozyme may have evolved independently many times. To better understand the observed distribution of hammerhead ribozymes, we used *in vitro* selection to search an unbiased sample of random sequences for comparably active self-cleaving motifs. Here we show that, under near-physiological conditions, the hammerhead ribozyme motif is the most common (and thus the simplest) RNA structure capable of self-cleavage at biologically observed rates. Our results suggest that the evolutionary process may have been channelled, in nature as in the laboratory, towards repeated selection of the simplest solution to a biochemical problem.

We designed a selection scheme for the isolation of self-cleaving RNAs with the goal of isolating ribozymes that are highly active at near-physiological pH and concentration of divalent ions. We started with a DNA pool coding for large random-sequence RNAs (Fig. 1) to allow for the isolation of complex self-cleaving molecules. Self-cleavage during transcription has limited the ability of previous *in vitro* selection experiments to evolve highly active self-cleaving ribozymes<sup>6</sup>. We inhibited unwanted premature self-cleavage by including in the transcription reaction a high concentration of a 16-base oligodeoxynucleotide that was complementary to the designed cleavage site. Control experiments showed that self-cleavage of a hammerhead ribozyme during transcription decreased from 90% to essentially undetectable levels in the presence of 50  $\mu$ M of the inhibitory oligonucleotide. We then purified, by denaturing gel electrophoresis, full-length RNA transcripts free of the inhibitory oligonucleotide. The self-cleavage reaction was subsequently initiated by the addition of magnesium ions, and terminated by the addition of EDTA and denaturant. Molecules that had self-cleaved within a 16-nucleotide region near the 5' end of the pool were then

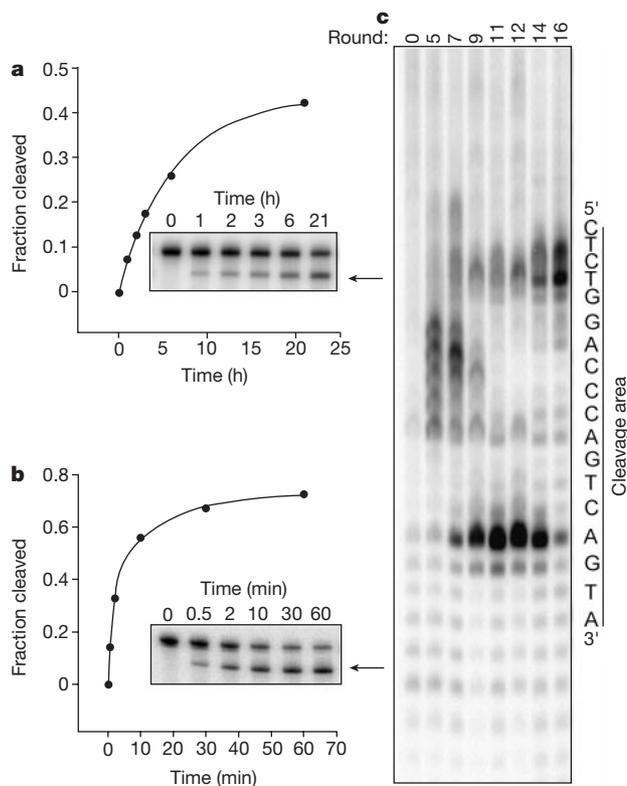


**Figure 1** Pool design. Pool RNAs consisted of a 60-nucleotide 5' constant region (containing the cleavage region), two random regions (63 and 65 nucleotides long) interrupted by an 18-nucleotide constant region, and an 18-nucleotide 3' constant region. The sequences of the central portion of the 5' constant region and of the blocking oligonucleotide are shown.

selected on the basis of their altered electrophoretic mobility. The selected RNAs were reverse transcribed (RT), the complementary DNA was amplified by polymerase chain reaction (PCR), and the DNA transcribed to re-initiate the selection cycle.

After five rounds of selection, a low level of self-cleavage activity was observed in the pool, corresponding to an average rate of  $\sim 0.003 \text{ min}^{-1}$  (Fig. 2a). In subsequent rounds, the reaction-incubation time was gradually reduced to enrich for more active molecules. By round 12, the pool activity had increased by almost 100-fold compared with round 5 (Fig. 2b). After round 12, pool molecules were mutagenized and subjected to additional rounds of selection. Four more rounds were carried out at lower pH and magnesium concentration (pH 7.2 and 0.5 mM  $\text{MgCl}_2$ ); a further modest increase in activity to  $\sim 0.8 \text{ min}^{-1}$  was observed by round 14 (Table 1).

We examined the evolution of the population in two ways. An initial analysis of the RNA cleavage sites showed that the ribozyme population changed significantly during the course of the selection (Fig. 2c). The round 5 RNAs cleaved largely in a cluster of sites in the 5' half of the 'cleavage area' (defined by the inhibitory oligonucleotide and the size of the gel-fractionated selected RNAs). However, as the selection progressed, the cleavage sites shifted mainly to a single site in the 3' region of the cleavage area in rounds 11 and 12, and then migrated again to the most 5' region of the cleavage area in rounds 14 and 16. To obtain a more detailed picture of the evolving RNA population, we cloned and sequenced 30–50 RNAs from rounds 5, 7, 9, 11, 12, 14 and 16. In round 5, when ribozyme



**Figure 2** Self-cleavage of pool RNAs. After completion of each round, RNAs were transcribed in the presence of  $\alpha$ -<sup>32</sup>P-UTP and blocking oligonucleotide, then purified and assayed for self-cleavage in 5 mM  $\text{MgCl}_2$  at pH 7.8 and 26 °C. **a, b**, Time courses of round 5 and 12 pools, respectively. Arrowheads mark 3' cleavage products; the smaller 5' fragments are not shown. **c**, Cleavage sites of round 5–16 RNAs. End-labelled primers complementary to the 5' constant region were used to reverse transcribe pool RNAs after self-cleavage. The lengths of the RT products were determined by comparison with markers from a sequencing reaction run in the same gel.

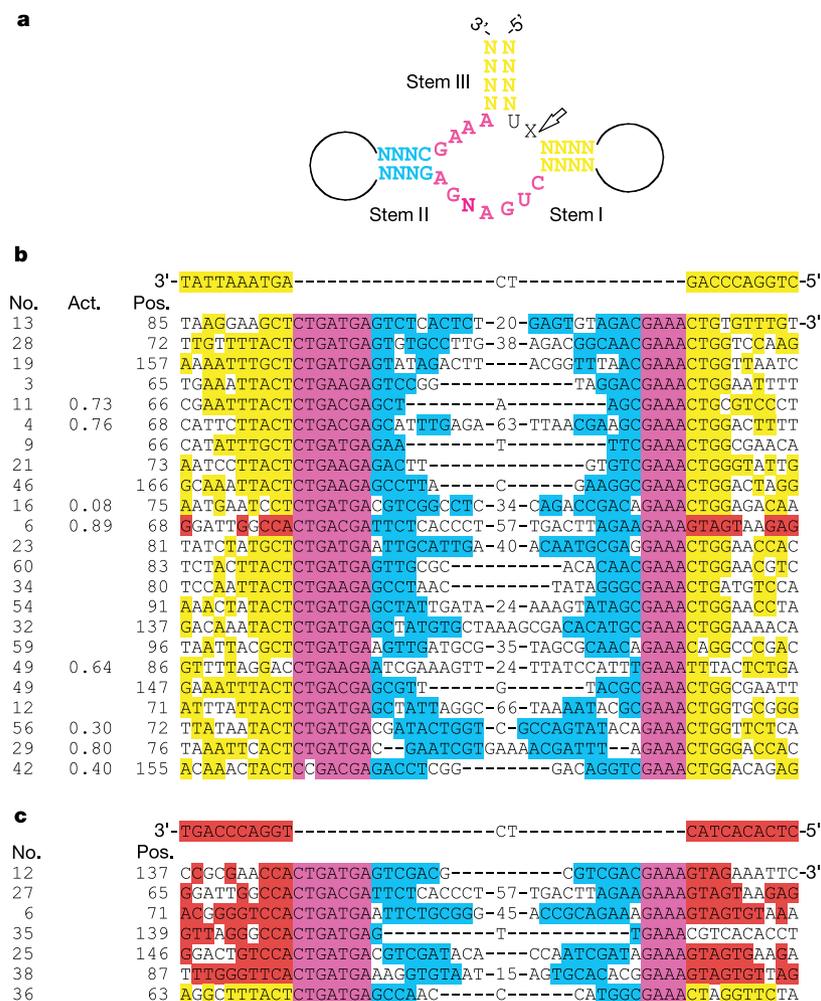
activity first clearly emerged, many unique sequences were present, consistent with the high diversity of low-activity ribozymes previously observed by Tang and Breaker<sup>7</sup>.

Clones that contained the following consensus sequence 5'-CTGANGA...GAAA-3', which is characteristic of the hammerhead ribozyme<sup>8,9</sup>, were rare in round 5, but increased in frequency in the round 7 and 9 sequences, and completely dominated the round 11 and 12 sequences (Table 1). Remarkably, three of the round 11 clones (not shown) and one round 12 clone contained two complete hammerhead motifs. Further evidence that clones containing the hammerhead consensus sequences do encode hammerhead ribozymes comes from examination of the aligned sequences (Fig. 3). The sequences 5' of the CTGANGA and 3' of the GAAA are complementary to the cleavage region of the pool, and if base paired would be consistent with the secondary structure of the hammerhead ribozyme (Fig. 3a) and the observed cleavage products (Fig. 2c). In most round 12 clones (Fig. 3b), the region between the CTGANGA and GAAA elements can form a stem-loop structure, which would correspond to stem-loop II of the hammerhead ribozyme. In some clones, this putative stem is capped by an extensive loop, whereas in others a compact stem-loop structure is present. We tested the activity of several clones from this round individually (Fig. 3); they ranged from 0.89 min<sup>-1</sup> (clone 6) to 0.083 min<sup>-1</sup> (clone 16), comparable

to the activities of natural hammerhead ribozymes.

On the basis of the sequence and structural features noted above, we consider the selected ribozymes to belong to the hammerhead class of self-cleaving ribozymes. However, a closer examination of the aligned sequences revealed an unexpected degree of structural variation in what are normally viewed as conserved elements of the hammerhead ribozyme. In addition, some of the other hammerhead ribozymes that we observed contain mismatched or surprisingly short base-paired stems. Our mutational analysis of these ribozymes strongly supports the idea that their core structure is that of a hammerhead ribozyme, and suggests that these deviations from the hammerhead consensus are compensated for by other regions of the selected sequence. The single non-hammerhead clone found in round 12 (clone 10, not shown) was highly active with a self-cleavage rate constant of 0.74 ± 0.014 min<sup>-1</sup>. No obvious similarity could be seen between this ribozyme and any of the known natural self-cleaving RNAs.

Our experiments firmly suggest that the hammerhead ribozyme is the simplest RNA motif capable of catalysing RNA strand cleavage at rates between 0.1 and 1.0 min<sup>-1</sup>, because this single motif dominated the selected population once those rates had been attained. Other RNA motifs that catalyse self-cleavage with similar rates must either be more complex (and thus less common in samples of random-sequence RNA), must amplify significantly less



**Figure 3** Alignments of unique hammerhead-type clones from rounds 12 and 16. **a**, A standard hammerhead ribozyme<sup>16</sup>. X is A, U or C. **b**, **c**, Alignment of DNA sequences derived from the random region of the library, for clones from rounds 12 and 16, respectively, with substrate cleavage sites at the top. No., clone number; Act., self-

cleavage rates (min<sup>-1</sup>); Pos., position of first base shown from 5' end of transcript. Colours represent the following: magenta, putative core hammerhead motifs; yellow, potential base pairing with the cleavage region shown at the top; red, potential base pairing with the alternative cleavage region; blue, potential stem II base pairing.

**Table 1 Appearance of hammerhead clones during selection**

| Round | Frequency* | Activity† (min <sup>-1</sup> ) | No.‡ | Progenitors§ |
|-------|------------|--------------------------------|------|--------------|
| 5     | 0.02       | 0.0026 ± 0.000057              | 40   | 1/39         |
| 7     | 0.17       | 0.026 ± 0.0021                 | 29   | 20/5         |
| 9     | 0.67       | 0.086 ± 0.0022                 | 33   | 23/8         |
| 11    | 1.00       | 0.166 ± 0.0053                 | 30   | 17/0         |
| 12    | 0.98       | 0.322 ± 0.014                  | 55   | 22/1         |
| 13    | —          | 0.655 ± 0.022                  | —    | —            |
|       |            | 0.0381 ± 0.0031                |      |              |
| 14    | 0.93       | 0.835 ± 0.054                  | 30   | 19/1         |
|       |            | 0.050 ± 0.00098                |      |              |
| 16    | 0.77       | 0.289 ± 0.031                  | 30   | 7/2          |
|       |            | 0.0649 ± 0.0069                |      |              |

\* Observed frequency of hammerhead-containing clones.

† Estimated pool activity (mean ± s.d.).

‡ Number of sequenced clones.

§ Number of distinct, independent sequences with/without hammerhead motif.

|| Activity in 0.5 mM MgCl<sub>2</sub> at pH 7.2.

efficiently by the RT-PCR procedure we used, or must be restricted to specific substrate sequences not included in our selection. Our results are consistent with those of Tang and Breaker<sup>7</sup>, who observed many distinct ribozymes with activities lower than 0.1 min<sup>-1</sup>. In contrast, another selection experiment<sup>10</sup>, in which the random-sequence region contained only 18 nucleotides, yielded only hammerhead ribozymes at both low and high activities. However, the stringent size constraint on the random-sequence region, and the constraints placed on cleavage site recognition, may have ruled out the appearance of other ribozymes. There are several natural classes of self-cleaving ribozymes other than the hammerhead, including the hairpin ribozyme<sup>11</sup>, the hepatitis delta ribozyme (HDV)<sup>12,13</sup>, and the *Neurospora* VS motif<sup>14</sup>. We wondered why none of these alternative highly active ribozymes was observed in our selection. These ribozymes are all significantly more complex than the hammerhead ribozyme, and are thus expected to occur much less frequently in selection experiments, in the absence of selective factors other than rate of cleavage. Thus, if selective pressure requires only that self-cleavage activity be in the range of 0.1–1.0 min<sup>-1</sup>, the expected result, both *in vitro* and in nature, is the emergence of the hammerhead ribozyme as the most probable solution. Our results show that, despite the dominance of contingency (historical accident) in some recent discussions of evolutionary mechanisms<sup>15</sup>, purely chemical constraints (that is, the ability of only certain sequences to carry out particular functions) can lead to the repeated evolution of the same macromolecular structures. □

**Methods**

**Selection procedure**

Pool RNA dissolved in 10 mM Tris-HCl buffer at pH 8.0 and 1 mM EDTA (TE) was heated to 85 °C then cooled to ambient temperature. Cleavage reactions were started by adding pool RNA to reaction buffer such that the final composition was 1 μM RNA, 5 mM MgCl<sub>2</sub>, 50 mM Tris-HCl at pH 7.8, 10 mM DTT and 1 U μl<sup>-1</sup> RNasin. Reaction times were 17–20 h at 26 °C in rounds 1–5 and in the subsequent rounds were adjusted to produce 1% cleavage. Reactions were stopped by addition of two volumes of 8 M urea, 2 mM Tris-HCl at pH 7.5 and 75 mM EDTA. Samples were resolved by electrophoresis in 8% denaturing polyacrylamide gels alongside a marker RNA to identify the position of the cleaved products; this region was excised and the RNA was isolated. RNAs were reverse transcribed using SuperScript II (Gibco-BRL), amplified by PCR and transcribed with T7 RNA polymerase (0.5 mM nucleoside triphosphates, 40 mM Tris-HCl at pH 7.9, 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM dithiothreitol and 1 U μl<sup>-1</sup> RNasin) in the presence of 50 μM of the blocking oligonucleotide (5'-TACTGACTGGGTCCAG-3'). After transcription, RNAs were purified by denaturing polyacrylamide gels, eluted, ethanol precipitated, dissolved in TE, and stored at -80 °C.

**Cleavage site determination**

RNAs were denatured with heat, cooled to ambient temperature, incubated in cleavage buffer, and then reverse transcribed with a 5'-<sup>32</sup>P-labelled primer (5'-CGTTGTGGTCTATC-3'). Samples mixed with formamide stop buffer were analysed by electrophoresis in 12% denaturing polyacrylamide gels.

Received 25 May; accepted 7 August 2001.

- Pabon-Pena, L. M., Zhang, Y. & Epstein, L. M. New satellite 2 transcripts self-cleave by using an extended hammerhead structure. *Mol. Cell. Biol.* **11**, 6109–6115 (1991).
- Zhang, Y. & Epstein, L. M. Cloning and characterization of extended hammerheads from a diverse set of caudate amphibians. *Gene* **172**, 183–190 (1996).
- Ferbeyre, G., Smith, J. M. & Cedergren, R. Schistosome satellite DNA encodes active hammerhead ribozymes. *Mol. Cell. Biol.* **18**, 3880–3888 (1998).
- Rojas, A. A. *et al.* Hammerhead-mediated processing of satellite pDo500 family transcripts from *Dolichopoda* cave crickets. *Nucleic Acids Res.* **28**, 4037–4043 (2000).
- Diener, T. O. Circular RNAs: relics of precellular evolution? *Proc. Natl Acad. Sci. USA* **86**, 9370–9374 (1989).
- Williams, K. P., Ciafre, S. & Tocchini-Valentini, G. P. Selection of novel Mg<sup>2+</sup>-dependent self-cleaving ribozymes. *EMBO J.* **14**, 4551–4557 (1995).
- Tang, J. & Breaker, R. R. Structural diversity of self-cleaving ribozymes. *Proc. Natl Acad. Sci. USA* **97**, 5784–5789 (2000).
- Thomson, J. B., Tuschl, T. & Eckstein, F. in *Catalytic RNA* (eds Eckstein, F. & Lilley, D. M. J.) 173–196 (Springer, Heidelberg, 1997).
- Ruffner, D. E., Stormo, G. D. & Uhlenbeck, O. C. Sequence requirements of the hammerhead RNA self-cleavage reaction. *Biochemistry* **29**, 10695–10702 (1990).
- Conaty, J., Hendry, P. & Lockett, T. Selected classes of minimised hammerhead ribozyme have very high cleavage rates at low Mg<sup>2+</sup> concentration. *Nucleic Acids Res.* **27**, 2400–2407 (1999).
- Hampel, N. in *Progress in Nucleic Acids Research and Molecular Biology* (ed. Moldave, K.) 1–38 (Academic, London, 1998).
- Been, M. D. *Cis-* and *trans-*acting ribozymes from a human pathogen, hepatitis delta virus. *Trends Biochem. Sci.* **19**, 251–256 (1994).
- Thill, G., Vasseur, M. & Tanner, N. K. Structural and sequence elements required for the self-cleavage activity of the hepatitis delta virus ribozyme. *Biochemistry* **32**, 4254–4262 (1993).
- Beattie, T. L., Olive, J. E. & Collins, R. A. A secondary-structure model for the self-cleaving region of *Neurospora* VS RNA. *Proc. Natl Acad. Sci. USA* **92**, 4686–4690 (1995).
- Gould, S. J. *Wonderful Life: The Burgess Shale and the Nature of History* 292–323 (Norton, New York, 1989).
- Long, D. M. & Uhlenbeck, O. C. Self-cleaving catalytic RNA. *FASEB J.* **7**, 25–30 (1993).

**Acknowledgements**

We thank P. Svec for assistance in cloning and sequencing, and members of the laboratory for their comments on the manuscript. This work was supported in part by a grant from the National Institutes of Health. J.W.S. is an investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to J.W.S. (e-mail: szostak@molbio.mgh.harvard.edu).

**corrections**

**Retrotransposition of a bacterial group II intron**

**Benoit Cousineau, Stacey Lawrence, Dorie Smith & Marlene Belfort**

*Nature* **404**, 1018–1021 (2000).

In this Letter, we demonstrated active group II intron transposition to multiple ectopic sites and showed that movement to these new sites occurs via an RNA intermediate in a process termed retrotransposition. Of interest was whether retrotransposition occurs by reverse splicing of the intron RNA into RNA, single-stranded DNA and/or double-stranded DNA targets. Our results suggested that retrotransposition occurs predominantly by the intron reverse-splicing into an RNA target. However, subsequent experiments using a different intron donor have indicated that the selection method for retrotransposition events can strongly bias the interpretation of pathway (K. Ichyanagi, A. Beauregard and M.B., unpublished results). The new data do not point to RNA-targeting as the principal means for intron retrotransposition, and DNA-based pathways must be considered as playing an important role. □