



Published as: *Nat Chem Biol.* 2009 February ; 5(2): 97–99.

A natural ribozyme with 3',5' RNA ligase activity

Quentin Vicens and Thomas R Cech

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, UCB 215, Boulder, Colorado 80309-0215, USA

Abstract

Using electrophoresis, sequencing and enzymatic digestion, we show that the group I intron from the cyanobacterium *Anabaena* sp. *PCC 7120* catalyzes phosphodiester bond formation using a triphosphate on the 5'-terminal nucleotide, much like protein polymerases and engineered ribozymes. In the process, this ribozyme forms a unique circular RNA that incorporates the exogenous guanosine cofactor added during self-splicing. This finding may have relevance to a prebiotic RNA world and to modern biology.

Group I introns self-splice in two chemical steps upon addition of a cofactor that can be guanosine (**1**) or one of its 5'-phosphorylated forms—guanosine 5'-monophosphate (GMP or pG, **2**), guanosine 5'-diphosphate (GDP, **3**) or guanosine 5'-triphosphate (GTP, **4**)¹. During the first step, this exogenous guanosine (here called ' α G') binds to a guanosine binding site inherent to the folded RNA structure; it then attacks the phosphorus atom at the 5' splice site, becoming covalently attached to the 5' end of the intron through a 3',5' linkage (Fig. 1a). A conformational change then brings the conserved guanosine at the 3' end of the intron (called ' ω G') into the guanosine binding site in place of the α G. During the second chemical step, the free 3' hydroxyl of the 5' exon attacks the 3' splice site, resulting in the formation of a phosphodiester bond that ligates the two exons. The released intron RNA is a linear RNA molecule having a guanosine at both the 5' end (α G) and the 3' end (ω G). This intron can undergo further transesterification reactions, thereby forming circles that contain fewer nucleotides than the linear intron² (Fig. 1a). A parallel pathway that requires hydrolysis at the 3' splice site leads to full-length circles (Fig. 1b). The formation of full-length circles is a general property of group I introns that may be linked to intron mobility³.

We recently assayed for self-splicing activity of 12 group I introns under 6 reaction conditions, in the presence of [α -³²P]GTP (**5**) as the cofactor⁴. While screening for catalytic activity, an unexpected product was observed for a 249-nucleotide group I intron embedded in a tRNA^{Leu} gene from the cyanobacterium *Anabaena* sp. *PCC 7120*^{5,6} (Fig. 2a). This ³²P-labeled product had an apparent size around 550–600 nucleotides—larger than the initial unspliced precursor RNA (334 nucleotides) and much larger than the self-spliced linear intron, which was observed (as expected) to be around 250 nucleotides in size. No labeled product of that size could be accounted for by known self-splicing mechanisms.

The low electrophoretic mobility of this RNA species suggested that it was either multimeric, circular or branched. In order to test for this, a time-course reaction over two

Correspondence should be addressed to Q.V. (quentin.vicens@colorado.edu).

Note: Supplementary information and chemical compound information is available on the Nature Chemical Biology website.

AUTHOR CONTRIBUTIONS

Q.V. and T.R.C. designed the study and wrote the manuscript; Q.V. performed the experiments.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

hours was split and loaded onto two separate polyacrylamide gels, which were run using TBE buffer (1× TBE is 100 mM Tris-base, 83 mM boric acid, 1.0 mM EDTA) at a concentration of either 0.5× or 1× (see Supplementary Methods online) (Fig. 2b,c). This experiment first revealed that the formation of the low-mobility product followed the release of the linear intron, starting between 3 and 13 min after RNA splicing was initiated. Second, the low-mobility band had different relative mobilities in the two TBE concentrations, which indicates that it corresponded to either a branched or a circular species⁷. A similar time-course experiment using the gel-purified linear intron as the starting material (which contained the αG, but in the absence of any additional free [α -³²P]GTP) confirmed a precursor-product relationship between the excised linear intron and the new RNA (Supplementary Fig. 1a online). The reverse reaction was also demonstrated: the linear intron could be obtained after incubation of the gel-purified low-mobility product under similar reaction conditions (Supplementary Fig. 1a).

Sequencing the gel-purified low-mobility product by reverse transcription using a primer binding to a region near the 5' end of the *Anabaena* intron revealed that this product was in fact a covalently closed circle (Fig. 2d). This circle contained an additional residue between the ωG and the first residue at the 5' end of the intron. The presence of this additional nucleotide was also apparent after sequencing of the circularization product obtained from the precursor RNA instead of the purified excised linear intron, which in fact corresponded to both this circle and the full-length circle previously observed⁶ (Supplementary Fig. 1b). The pausing of the reverse transcriptase observed across the sequencing gel at the position of this extra residue—likely due to circle reopening (Supplementary Fig. 1a)—masked its identity. We anticipated this nucleotide to be the αG and confirmed this prediction by performing a complete digestion of the circle using nuclease P1. Nuclease P1 cleaves any sequence and leaves a phosphate group on the 5' end of the mononucleotide products (Supplementary Fig. 2a online). A single spot was observed next to the guanosine marker on a TLC plate, which is consistent with GMP having been incorporated into the circle (the label at the α position of the αG would be the only radioactive label in the circle) (Fig. 2e).

When splicing was performed in the presence of [γ -³²P]GTP (**6**), the intron was excised but no labeled circle was observed, which suggests that circle formation was accompanied by inorganic pyrophosphate release (Fig. 2f). Further supporting this mechanism, the excised intron derived from body-labeled precursor RNA self-spliced using GMP was unable to circularize (Supplementary Fig. 3 online). Such a reaction mechanism is similar to that used by RNA and DNA polymerases, as well as several *in vitro*-evolved ribozymes (ref. 8; reviewed in ref. 9). However, certain ribozymes catalyze the formation of a 2',5' bond instead of the 3',5' bond found in most natural RNA molecules¹⁰. In order to test the regiospecificity of the linkage at the circularization site, we performed a complete digestion of the circle using a combination of RNases T1 and A, which specifically cleave 3',5' bonds (Supplementary Methods) after guanosine and uracil or cytosine, respectively (Supplementary Fig. 2b). In the event of a 2',5' bond at the circularization site, the product would have been a GpGp dinucleotide (**7**), whereas a 3',5' bond at the circularization site would have been cleaved by RNase T1, resulting in guanosine 3'-monophosphate (Gp, **8**) as the final product. The product observed after migration on a TLC plate was the mononucleotide, which indicates that the circle was linked by a 3',5' bond (the only label in the circle came from the phosphorus atom connecting the αG and the ωG; Fig. 2g).

To further characterize the catalytic properties of this natural RNA ligase, we measured the rate of circle formation in the presence of 15 mM MgCl₂, 25 mM NaCl at pH 7.5 and 32 °C. The first-order rate constant was $2.0 \times 10^{-3} \text{ min}^{-1}$, which is the same order of magnitude as that found for a ribozyme *in vitro* evolved from the P4–P6 domain of the *Tetrahymena thermophila* group I intron¹¹ ($3.0 \times 10^{-3} \text{ min}^{-1}$; Fig. 2h,i). This rate is modest compared to

that achieved by some protein enzymes and *in vitro*-evolved ribozymes under optimal conditions ($>0.1 \text{ min}^{-1}$)⁹. Nevertheless, the rate enhancement over the uncatalyzed reaction is still 10^4 (using $2 \times 10^{-7} \text{ min}^{-1}$ as an estimate of the rate of the uncatalyzed reaction⁹) (Supplementary Fig. 4 online). Finally, this rate constant was unchanged between 2.0 and 50 nM RNA, which indicates that the reaction is unimolecular (Supplementary Fig. 5 online).

These results expand on growing observations that a common RNA scaffold can support various catalytic mechanisms^{12,13}. Yet the catalytic site for this circularization reaction remains to be pinpointed. The simplest explanation would be that the guanosine binding site is being used, and the free 3' hydroxyl group of the ω G (still bound to the guanosine binding site after the second step of self-splicing) attacks the α phosphate of the (α -³²P)GTP at the 5' end. Although we have no evidence that such circle formation occurs *in vivo*, having an additional nucleotide could be an advantage for the subsequent reinsertion of these introns into genomic sequences with no loss of information. Such a process would be analogous to a recombination event during which the extra nucleotide could serve as a leaving group for the insertion and could thus be swapped for the host sequence. Consequently, other ribozymes able to catalyze a similar reaction could be found among the numerous instances in the literature where [α -³²P]GTP has been used to identify introns in cellular RNA extracts.

In fact, the original discoverers of the *Anabaena* group I intron had also observed a product with low electrophoretic mobility. At the time, they related it to a highly structured product of the first step of self-splicing, because the corresponding product was resistant to alkaline phosphatase⁵. This band may now be attributed to the new circular species characterized here. Subsequent studies of the reactions catalyzed by the *Anabaena* intron—see for example references^{6,14}—used guanosine in place of GTP. Hence, the reaction reported herein would not have occurred and thereby remained unidentified for 18 years.

Whether this reaction is a relic of a prebiotic RNA world, or simply just another illustration of the plasticity of the group I intron scaffold, is unclear. The features of using a nucleoside triphosphate and extending an RNA chain by formation of a 3',5' bond are expected characteristics of a prebiotic ribozyme, if it is to transition to modern biochemistry¹⁵. However, obtaining the 3',5' regiospecificity is difficult, as prebiotic simulations of RNA synthesis often yield 2',5' linkages^{16,17}. Even though the first catalytic RNAs discovered (group I introns and the RNA component of RNase P) were considered to provide the much-anticipated evidence that an RNA world could have existed, these ribozymes only performed RNA cleavage or transesterification reactions from substrates that contained at least two nucleotides connected by phosphodiester linkages¹. Only later was it discovered that ribozymes can use the 5'-terminal triphosphate of an RNA molecule for RNA ligation, in the case of a modified group II intron¹⁸ and ribozymes identified by *in vitro* selection (ref. 8; reviewed in ref. 9). The results presented herein illustrate how a natural group I intron is also able to perform 3',5' RNA ligation. Together with the earlier findings that ribozymes selected by *in vitro* evolution from group I introns can catalyze RNA ligation^{11,19}, this provides additional evidence that prebiotic replication could have been catalyzed by ribozymes related to group I introns²⁰.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

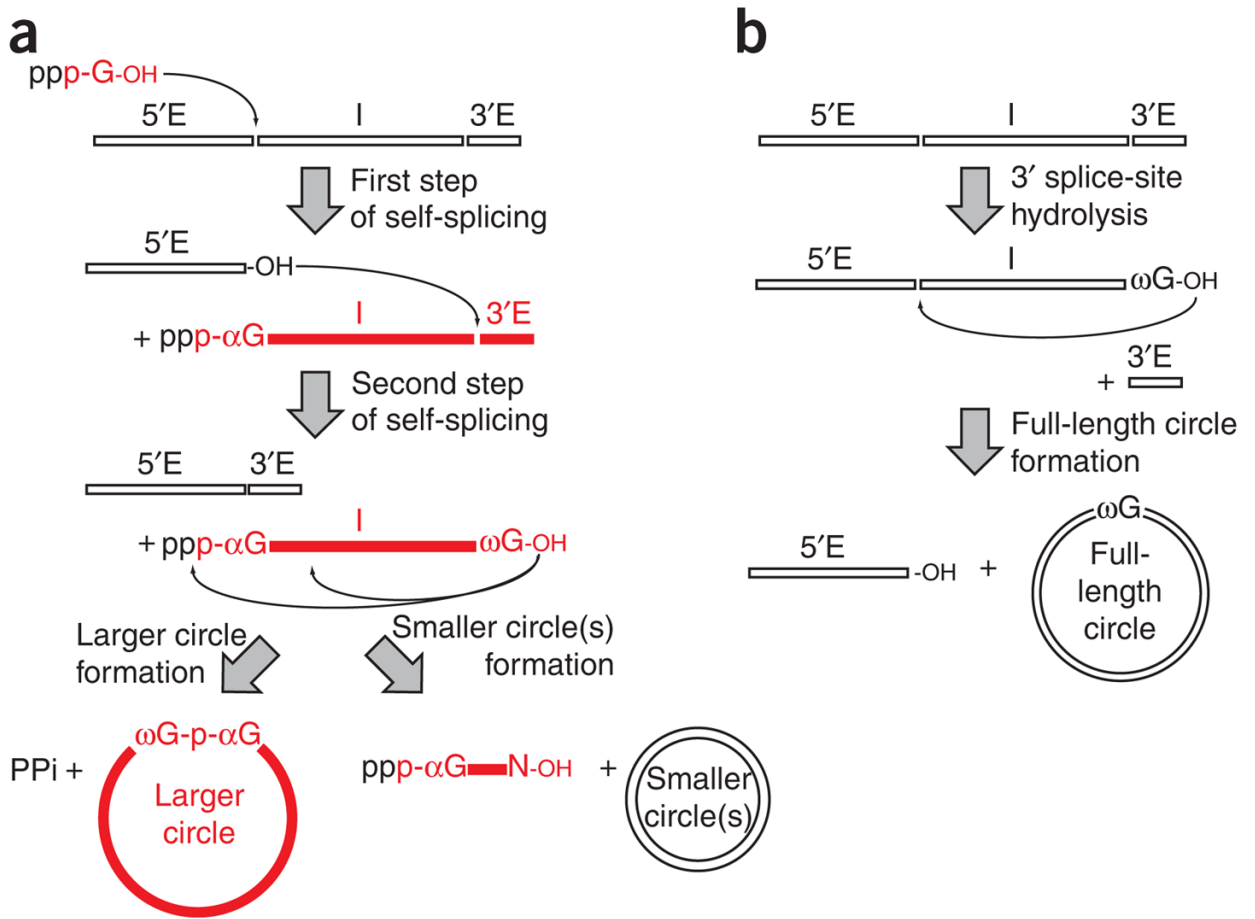
Acknowledgments

We thank A. Zaugg for experimental suggestions and insightful discussions; A. Gooding (Howard Hughes Medical Institute) for preparing the T7 RNA polymerase; D. Zappulla (Johns Hopkins University) for the gift of RNase A; the remaining members of the Cech laboratory, as well as H. Nielsen, O. Uhlenbeck, S. Silverman and M.

Caruthers, for helpful discussions; and H. Nielsen, S. Silverman, A. Gooding and A. Zaug for careful reading of the manuscript.

References

1. Cech, TR. The RNA World. I. Gesteland, RF.; Atkins, JF., editors. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York, USA: 1993. p. 239-269.
2. Zaug AJ, Grabowski PJ, Cech TR. *Nature* 1983;301:578–583. [PubMed: 6186917]
3. Nielsen H, et al. *RNA* 2003;9:1464–1475. [PubMed: 14624003]
4. Vicens Q, Paukstelis PJ, Westhof E, Lambowitz AM, Cech TR. *RNA* 2008;14:2013–2029. [PubMed: 18768647]
5. Xu MQ, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA. *Science* 1990;250:1566–1570. [PubMed: 2125747]
6. Zaug AJ, McEvoy MM, Cech TR. *Biochemistry* 1993;32:7946–7953. [PubMed: 8347600]
7. Kjems, J.; Egebjerg, J.; Christiansen, J. *Laboratory Techniques in Biochemistry and Molecular Biology: Analysis of RNA-Protein Complexes In Vitro*. van der Vliet, PC., editor. Elsevier; Amsterdam, The Netherlands: 1998. p. 181-232.
8. Bartel DP, Szostak JW. *Science* 1993;261:1411–1418. [PubMed: 7690155]
9. Joyce GF. *Angew Chem Int Edn Engl* 2007;46:6420–6436.
10. Murray HL, et al. *Mol Cell* 2001;8:201–211. [PubMed: 11511373]
11. Yoshioka W, Ikawa Y, Jaeger L, Shiraishi H, Inoue T. *RNA* 2004;10:1900–1906. [PubMed: 15525711]
12. Hammann C, Westhof E. *Genome Biol* 2007;8:210. [PubMed: 17472738]
13. Beckert B, et al. *EMBO J* 2008;27:667–678. [PubMed: 18219270]
14. Golden BL, Cech TR. *Biochemistry* 1996;35:3754–3763. [PubMed: 8619996]
15. Orgel LE. *Trends Biochem Sci* 1998;23:491–495. [PubMed: 9868373]
16. Ferris, JP.; Usher, DA. *Biochemistry*. Zubay, G., editor. Macmillan; New York: 1988. p. 1120-1152.
17. Lohrmann R, Orgel LE. *Tetrahedron* 1978;34:853–855.
18. Mörl M, Niemer I, Schmelzer C. *Cell* 1992;70:803–810. [PubMed: 1381286]
19. Jaeger L, Wright MC, Joyce GF. *Proc Natl Acad Sci USA* 1999;96:14712–14717. [PubMed: 10611278]
20. Cech TR. *Proc Natl Acad Sci USA* 1986;83:4360–4363. [PubMed: 2424025]
21. Grabowski PJ, Zaug AJ, Cech TR. *Cell* 1981;23:467–476. [PubMed: 6162571]

**Figure 1.**

Reaction pathways catalyzed by group I introns. **(a)** Self-splicing is followed by the formation of a larger circle (left; this work) or smaller circles (right). The exogenous [α - 32 P]GTP used here is shown in red. The subsequently radioactively labeled products are colored in red. **(b)** Full-length circle formation as described previously^{3,21}. 5'E, 5' exon; I, intron; 3'E, 3' exon.

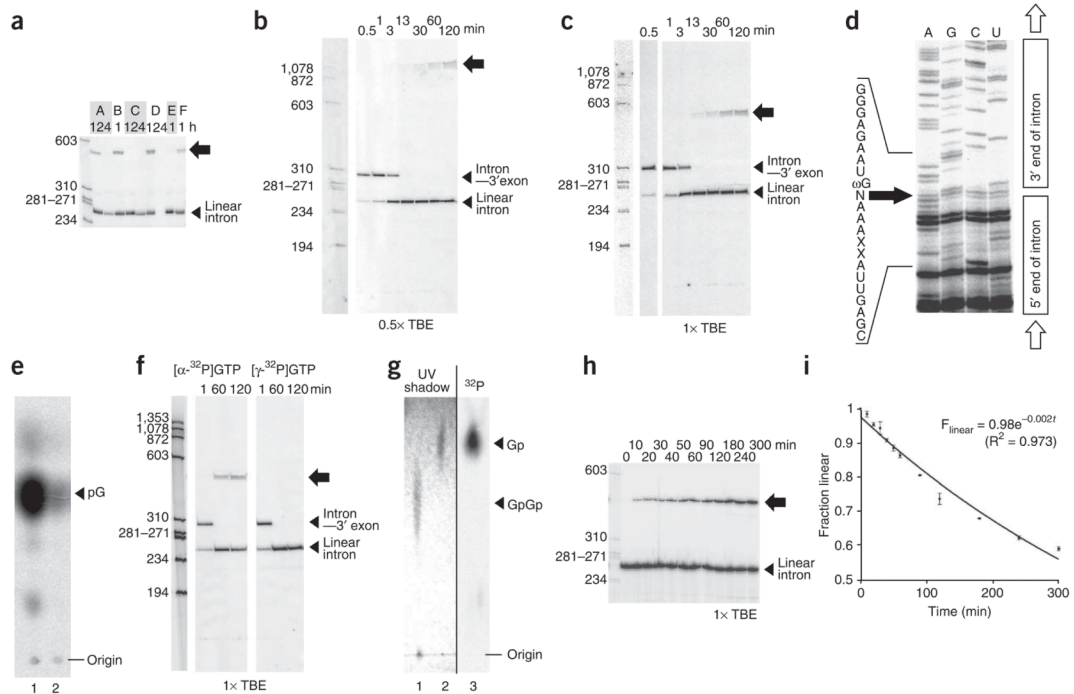


Figure 2.

New circular RNA. **(a)** *Anabaena* precursor RNA containing the group I intron after incubation in 0.2 mM GTP, $1.0 \mu\text{Ci ml}^{-1}$ [α - ^{32}P]GTP, under six reaction conditions (A: 15 mM MgCl_2 , 25 mM NaCl, 25 mM HEPES pH 7.5, 32 °C; B: same as A except 42 °C; C: same as A except 1.0 M NaCl; D: same as A except 200 mM MgCl_2 ; E: 100 mM $\text{Mg}(\text{OAc})_2$, 1.0 M NH_4OAc , 25 mM HEPES pH 7.5, 37 °C; F: 15 mM MgCl_2 , 0.5 M KCl, 5 mM DTT, 2 mM spermidine, 40 mM Tris pH 7.5, 50 °C) (from ref. 4). Block arrow, new species. **(b)** Reaction time-course (condition A), electrophoresis in 0.5 \times TBE. **(c)** Same except electrophoresis in 1.0 \times TBE. **(d)** Sequencing across the circularization site. N, incorporated residue. **(e)** Complete nuclease P1 digestion followed by TLC identifies αG incorporated into the circle (lane 1: P4–P6 RNA body-labeled with [α - ^{32}P]GTP and P1-digested; lane 2: gel-purified circle incorporating a single label at the circularization site and P1-digested). **(f)** Side-by-side time-course experiment as in **b** using [α - ^{32}P]GTP (left) or [γ - ^{32}P]GTP (right). **(g)** Complete RNase A plus RNase T1 digestion to determine the regiospecificity at the circularization site, visualized by TLC (lane 1: dinucleotide GpGp control, unlabeled; lane 2: mononucleotide Gp control, unlabeled; lane 3: gel-purified circle incorporating a single label at the circularization site, digested with both RNases). **(h)** Reaction time-course starting with purified linear intron (condition: A), electrophoresis in 1.0 \times TBE. **(i)** Determination of first-order rate constant of circularization. Error bars, s.d. of four experiments (10, 30, 60, 120 min) or two experiments (other data points). Full-length versions of these gels and repeat experiments are shown in Supplementary Figs. 6–9 online.