Published in final edited form as:

Science. 2010 August 13; 329(5993): 845–848. doi:10.1126/science.1190713.

An Allosteric Self-splicing Ribozyme Triggered by a Bacterial Second Messenger

Elaine R. Lee^{#1}, Jenny L. Baker^{#2}, Zasha Weinberg^{1,3}, Narasimhan Sudarsan^{1,3}, and Ronald R. Breaker^{1,3,4,*}

¹Department of Molecular, Cellular and Developmental Biology, Yale University, Box 208103, New Haven, CT 06520-8103, USA.

²Department of Chemistry, Yale University, Box 208103, New Haven, CT 06520-8103, USA.

³Howard Hughes Medical Institute, Yale University, Box 208103, New Haven, CT 06520-8103, USA.

⁴Department of Molecular Biophysics and Biochemistry, Yale University, Box 208103, New Haven, CT 06520-8103, USA.

[#] These authors contributed equally to this work.

Abstract

Group I self-splicing ribozymes commonly function as components of selfish mobile genetic elements. We identified an allosteric group I ribozyme, wherein self-splicing is regulated by a distinct riboswitch class that senses the bacterial second messenger c-di-GMP. The tandem RNA sensory system resides in the 5' untranslated region of the messenger RNA for a putative virulence gene in the pathogenic bacterium *Clostridium difficile*. c-di-GMP binding by the riboswitch induces folding changes at atypical splice site junctions to modulate alternative RNA processing. Our findings indicate that some self-splicing ribozymes are not selfish elements, but are harnessed by cells as metabolite sensors and genetic regulators.

Metabolite-sensing riboswitches control gene expression in most bacteria, plants and fungi (1-3). Computer-assisted searches (4) have recently been used to identify additional riboswitches (5), including an RNA motif (Fig. 1A) that associates with genes involved in c-di-GMP (cyclic di-guanosyl-5'-monophosphate) production, degradation, and signaling (fig. S1) (6). The structural characteristics of this motif are distinct from a previously validated c-di-GMP-I riboswitch class (fig. S2) for this second messenger (7-9), and therefore likely represents a distinct class called c-di-GMP-II.

An 84-nucleotide c-di-GMP-II RNA from *C. difficile* strain 630 (termed 84 Cd; Fig. 1B), was 5' ³²P-labeled and subjected to in-line probing (10, 11). The pattern of spontaneous RNA cleavage products generated without c-di-GMP (Fig. 1C) is consistent with our secondary structure model. Furthermore, 11 linkages exhibit reduced strand scission in 10 μ M c-di-GMP, which indicates these nucleotides are structurally stabilized on second

*Correspondence should be addressed to R.R.B. Dr. Ronald R. Breaker Tel: (203) 432-9389 Fax: (203) 432-0753
ronald.breaker@yale.edu.

messenger binding. Probing using a range of c-di-GMP concentrations reveals 1-to-1 binding with an apparent dissociation constant (K_D) of ~200 pM (Fig. 1D, fig. S3). K_D values for analogs are more than three orders of magnitude poorer (Fig. 1E), revealing molecular discrimination comparable to that exhibited by c-di-GMP-I riboswitches (7).

Riboswitch function by other c-di-GMP-II RNAs was validated *in vivo* using riboswitch-reporter gene fusions (fig. S4A, B) and *in vitro* using transcription termination assays (fig. S4C, D). However, the 84 Cd aptamer lacks normal riboswitch control structures and is located ~600 nucleotides upstream of the putative start codon for its associated ORF (Fig. 2A). This unusually large gap has all the hallmarks of a group I self-splicing ribozyme (12, 13) (fig. S5), therefore we speculated that the RNAs may collaborate to function as an allosteric ribozyme wherein splicing is controlled by c-di-GMP.

Allosteric ribozyme function was assessed by using an 864-nucleotide internally ^{32}P -labeled RNA (864 Cd Tandem) carrying the aptamer and ribozyme domains, and part of the *C. difficile* CD3246 ORF (Fig. 2B, fig. S5). Incubation of wild-type RNA with GTP yields characteristic group I ribozyme products (Fig. 2C, fig. S6). However, the amount of 5'E-3'E spliced exons increases substantially when both GTP and c-di-GMP are present (Fig. 3A, B). The dose-response curve for c-di-GMP induction of splicing fits a 1-to-1 binding curve with half-maximal 5'E-3'E production occurring with 30 nM second messenger (fig. S7).

Analysis of *in vitro* splice products (figs. S8, S9) reveals an atypical 5' SS. Total RNA isolated from cultured *C. difficile* 630 cells produced a dominant RT-PCR product (fig. S10) with a splice junction determined by sequencing that corresponding to that found in our *in vitro* assays. Furthermore, the extent of splicing increases with culture age. This latter finding is consistent with allosteric activation of splicing by c-di-GMP, whose concentrations should increase with increasing cell density.

Another RNA product (3'E*), determined to be a fragment of the 3' exon caused by GTP attack far from the 5' SS (fig. S8), is reduced in the presence of c-di-GMP (Figs. 2C, 3D). Production of both 5'E-3'E and 3'E* RNAs were monitored in subsequent assays. Aptamer mutants M1 and M2 that disrupt c-di-GMP binding no longer respond to the second messenger in splicing reactions (Fig. 2C, fig. S11). In contrast, mutant M3 that carries nucleotide changes to M2 to restore aptamer P3 stem formation recovers c-di-GMP-mediated splicing control (Fig. 2C). Thus, c-di-GMP binding to the aptamer triggers increased production of spliced exons.

We noticed two alternative base pairing structures, the anti-5' SS stem (Fig. 2B, blue shading) and the alternative ribozyme P1 stem (green shading), that may explain c-di-GMP splicing control. Anti-5' SS stem formation disrupts two base pairs of the ribozyme P1 stem, which subsequently favors formation of the alternative ribozyme P1 stem. Since c-di-GMP binding stabilizes aptamer substructures (Fig. 1C), low c-di-GMP concentrations should favor anti-5' SS stem formation and prevent GTP attack at the 5' SS. Indeed, a 132-nucleotide construct encompassing the aptamer through nucleotide 130 of the tandem RNA yields an in-line probing pattern consistent with anti-5' SS stem formation in the absence of

second messenger (fig. S12A). Additional mutants were used to validate this c-di-GMP-mediated structural rearrangement (fig. S12B-D).

The role of the anti-5' SS stem was further evaluated by establishing the effects of mutants M4, M5 and M6. Consistent with our model, mutants M4 and M5 disrupt regulation (Fig. 2C, fig. S11), while M6 combines these disruptive mutations to yield a construct that restores c-di-GMP binding and regulation. Mutations M7 through M11 were created to assess the importance of the ribozyme P1 stem and the alternative ribozyme P1 stem for allosteric control. Mutations M7 through M10 disrupt specific structures and exhibit expected reductions in particular RNA products. Most notably, normal c-di-GMP regulation is observed for M11, which combines several disruptive mutations to form an RNA that restores alternative base-pairing potential. Finally, M12 lacks any splicing activity due to disruption of the ribozyme's GTP binding site. These findings are consistent with an allosteric ribozyme mechanism wherein c-di-GMP regulates alternative base pairing between distal regions of the precursor mRNA.

Reactions with unlabeled 864 Cd Tandem RNA and [α - 32 P]GTP (Fig. 3A) were conducted to estimate the observed rate constant (k_{obs}) values for GTP attack (Fig. 3B). A near 13-fold increase in initial rate constant (first 15 minutes) was observed for GTP attack at the 5' SS when both GTP and c-di-GMP are present ($2.3 \times 10^{-2} \text{ min}^{-1}$). In contrast, there is only a modest reduction in the initial rate constant for GTP attack at the alternative GTP₂ site. This same pattern of ribozyme modulation is observed when internally labeled RNAs are used (Fig. 3C). A 12-fold increase in the initial k_{obs} for spliced product formation in the absence of c-di-GMP.

Similar results are obtained by evaluating the yields of spliced exon and alternative attack products after incubating for 120 minutes (Fig. 3C, E). Without c-di-GMP, the ribozyme produces greater than 20-fold more alternative attack product than spliced exons (Fig. 3E, left). Near equal amounts are produced when c-di-GMP is present. This change is largely due to an ~8-fold increase in spliced exon production, as the ratio of alternative attack product in the absence versus the presence of c-di-GMP drops only modestly (Fig. 3E, right). Thus, c-di-GMP enhances the production of fully spliced exons largely by favoring formation of the ribozyme P1.

The sequences and structures of the precursor mRNA and processed RNAs suggest a mechanism for translation control of the CD3246 ORF (Fig. 4). Although the UUG start codon predicted in annotated *C. difficile* genomes is atypical, this translation start site and initial polypeptide sequence is consistent with similar genes in related organisms. In the precursor RNA, the start codon resides in the right shoulder of the P10 stem (Fig. 4A). This arrangement should restrict ribosome access and preclude translation, which is a common mechanism for translation control by riboswitches (3). With sufficient c-di-GMP and GTP, ribozyme action yields a processed mRNA wherein the spliced junction resides in a perfect ribosome binding site (AGGAGG) located an optimal distance upstream of the start codon (Fig. 4B). Thus, allosteric activation of ribozyme self-splicing by c-di-GMP should favor translation. In contrast, ribozyme action in the absence of c-di-GMP favors GTP attack only four nucleotides upstream of the start codon. This product is trimmed of nucleotides that

otherwise could serve as a ribosome binding site (Fig. 4C), which should disfavor translation.

Bacteria naturally exploit tandem riboswitch architectures to create more complex gene control elements (14-17). Our findings add to this complexity by demonstrating how two ligand-responsive RNAs, a self-splicing ribozyme and a riboswitch aptamer, collaborate to function as an allosteric RNA requiring two RNA compounds (GTP and c-di-GMP) to promote splicing. This conjoined riboswitch-ribozyme system validates a prediction made more than 20 years ago (18) that some group I ribozymes could be controlled by nucleotide-derived alarmones. Also, this RNA is a natural mimic of an engineered ribozyme that controls splicing and gene expression in response to theophylline binding (19).

Some group I ribozymes may independently function as riboswitches that sense GTP because sufficient levels must be present for splicing to occur. If true, then the tandem riboswitch-ribozyme examined in this study may constitute a two-input control system that reads GTP and c-di-GMP concentrations to trigger splicing accordingly. Notably, GTP is the immediate biosynthetic precursor of c-di-GMP, and overproduction of this second messenger causes substantial decreases in GTP concentrations *in vivo* (20).

The harnessing of group I ribozymes by some organisms is suggested by the observation that bacteriophage ribozyme splicing is diminished in a bacterial host when ribosome function is inhibited (21). Moreover, the gene associations of some examples are conserved in evolutionarily distant species (13), implying useful rather than purely selfish functions. Of at least ten group I ribozymes present in *C. difficile* 630, nine are associated with a transposase gene (fig. S12), which facilitates mobility of selfish genetic elements. Only the allosteric ribozyme described in this work lacks a transposase gene, suggesting that this representative is not a selfish RNA, but has a location and a function that benefits the host.

Acknowledgments

We thank Jessica Davis for assistance with genetic experiments and members of the Breaker laboratory for helpful discussions. We also thank Linc Sonenshein for supplying *C. difficile* 630 spores and for training N.S. in *C. difficile* techniques. This work was supported by NIH grant PO1 GM022778 and by the Howard Hughes Medical Institute.

References and Notes

1. Roth A, Breaker RR. Annu. Rev. Biochem. 2009; 78:305. [PubMed: 19298181]
2. Dambach MD, Winkler WC. Curr. Opin. Microbiol. 2009; 12:161. [PubMed: 19250859]
3. Barrick JE, Breaker RR. Genome Biol. 2007; 8:R239. [PubMed: 17997835]
4. Tseng HH, Weinberg Z, Gore J, Breaker RR, Ruzzo WL. J. Bioinform. Comput. Biol. 2009; 7:373. [PubMed: 19340921]
5. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, Breaker RR. Genome Biol. 2010; 11:R31. [PubMed: 20230605]
6. Hengge R. Nat. Rev. Microbiol. 2009; 7:263. [PubMed: 19287449]
7. Sudarsan N, E. R. et al. Science. 2008; 321:411. [PubMed: 18635805]
8. Kulshina N, Baird NJ, FerréD'Amaré AR. Nat. Struct. Mol. Biol. 2009; 16:1212. [PubMed: 19898478]
9. Smith KD, Lipchock SV, Ames TD, Wang J, Breaker RR, Strobel SA. Nat. Struct. Mol. Biol. 2009; 16:1218. [PubMed: 19898477]

10. Soukup GA, Breaker RR. RNA. 1999; 5:1308. [PubMed: 10573122]
11. Regulski EE, Breaker RR. Methods Mol. Biol. 2008; 419:53. [PubMed: 18369975]
12. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. Cell. 1982; 31:147. [PubMed: 6297745]
13. Nielsen H, Johansen SD. RNA Biol. 2009; 6:375. [PubMed: 19667762]
14. Mandal M, et al. Science. 2004; 306:275. [PubMed: 15472076]
15. Welz R, Breaker RR. RNA. 2007; 13:573. [PubMed: 17307816]
16. Sudarsan N, et al. Science. 2006; 13:300. [PubMed: 17038623]
17. Poiata E, Meyer MM, Ames TD, Breaker RR. RNA. 2009; 15:2046. [PubMed: 19776155]
18. Shub DA, Xu M-Q, Gott JM, Zeeh A, Wilson LD. Cold Spring Harbor Symp. Quant. Biol. LII. 1987:193–200.
19. Thompson KM, Syrett HA, Knudsen SM, Ellington AD. BMC Biotechnol. 2002; 2:21. [PubMed: 12466025]
20. Simm R, Morr M, Kader A, Nimtz M, Römling U. Mol. Microbiol. 2004; 53:1123. [PubMed: 15306016]
21. Sandegren L, Sjöberg BM. J. Bacteriol. 2007; 189:980–990. [PubMed: 17122344]

Fig. 1. c-di-GMP-II riboswitches. **(A)** Consensus sequence and secondary structure model for c-di-GMP-II riboswitch aptamers. Nucleotides in red are conserved in >97% of the representatives. Other annotations are described in fig. S2. **(B)** Wild-type (WT) 84 Cd RNA encompassing the motif upstream of a possible virulence gene. Disruptive (M2) and restorative (M3) mutations are depicted, and sites of spontaneous cleavage are derived from C. Asterisks identify nucleotides added to facilitate *in vitro* transcription. **(C)** Denaturing polyacrylamide gel electrophoresis (PAGE) of WT 84 Cd RNA in-line probing cleavage products. Notations: no reaction (NR); partial digestion with RNase T1 (T1, cleaves after G residues); alkali (- OH, cleaves at all linkages); incubation in the absence (-) or presence (+) of 10 μ M c-di-GMP. Selected RNase T1 digestion product bands are identified. **(D)** Plot of the fraction of RNA remaining unbound versus the logarithm of the concentration of c-di-GMP present during in-line probing (fig. S3). **(E)** K_D values for 84 Cd binding to c-di-GMP and various analogs.

Fig. 3.

Rate constant modulation by c-di-GMP. (A) Time course of $[\alpha\text{-}^{32}\text{P}]\text{GTP}$ attack at sites GTP₁ or GTP₂ in the absence or presence of c-di-GMP. (B) Plot of the natural logarithm of the fraction of unprocessed or differently processed RNA ("pre-dp RNA") versus time for the reaction in A. Values for fraction processed were corrected for ~50% of the precursor remaining after exhaustive incubation. (C) Time course of the production of spliced exons (5'E-3'E) or alternative GTP₂ site fragment (3'E*) in the absence or presence of c-di-GMP. Annotations are as described for Fig. 2C. (D) Plot of the natural logarithm of the fraction of unprocessed or differently processed RNA versus time for the reaction in C, corrected as described in B. (E) Changes in splice product yields on introduction of c-di-GMP. Left: ratio of the number of 3'E* molecules versus the number of 5'E-3'E spliced exon molecules in the absence or presence of c-di-GMP, respectively. Right: ratio of the numbers of 3'E* molecules and the ratio of the numbers of 5'E-3'E spliced exon molecules in the absence or presence of c-di-GMP, respectively.

Fig. 4. Proposed mechanism for allosteric ribozyme-mediated gene control. **(A)** Precursor mRNA with the start codon sequestered by the ribozyme P10 stem. **(B)** RNA processed in the presence of GTP and c-di-GMP unmask the start codon and creates a perfect ribosome binding site. **(C)** RNA processed in the presence of GTP alone lacks a ribosome binding site.