Direct *in vitro* selection of *trans*-acting ribozymes for posttranscriptional, site-specific, and covalent fluorescent labeling of RNA

Mohammad Ghaem Maghami,[†‡] Carolin P.M. Scheitl,[†] Claudia Höbartner*[†‡]

[†]Institute of Organic Chemistry, University of Würzburg, Am Hubland, 97074 Würzburg, Germany.
[‡] International Max Planck Research School Molecular Biology, University Göttingen, Germany.

Supporting Information Placeholder

ABSTRACT: General and efficient tools for site-specific fluorescent or bioorthogonal labeling of RNA are in high demand. Here, we report direct *in vitro* selection, characterization, and application of versatile *trans*-acting 2'-5' adenylyl transferase ribozymes for covalent and site-specific RNA labeling. The design of our partially structured RNA pool allowed for *in vitro* evolution of ribozymes that modify a predetermined nucleotide in *cis* (i.e. intramolecular reaction), and were then easily engineered for applications in *trans* (i.e. in an intermolecular setup). The resulting ribozymes are readily designed for specific target sites in small and large RNAs and accept a wide variety of N^6-modified ATP analogues as small molecule substrates. The most efficient new ribozyme (FH14) shows excellent specificity towards its target sequence also in the context of total cellular RNA.

RNA labeling and visualization is a critical pre-requisite to study RNA localization, folding and structural dynamics. Chemical synthesis of modified RNA is well established but limited to short oligonucleotides. A number of chemical and chemoenzymatic methods have been reported for post-synthetic, site-specific, and covalent labeling of RNA.[1-5] RNA-methyltransferases in combination with synthetic *S*-adenosyl methionine analogues have been repurposed for postsynthetic RNA labeling, including trimethylguanosine synthases for modification of the mRNA cap structure,[6-7] tRNA modifying enzymes[8] and guide RNA dependent archaeal C/D box methyltransferases.[9] Additional tRNA modifying enzymes have been used for covalent labeling of RNA, including the tRNA[Ile2]-agmatidine synthetase (Tias) with synthetic agmatine analogs,[10] and a bacterial tRNA guanine transglycosylase using synthetic preQ1 substrates to site-specifically install functional nucleobase analogs on sequence-defined hairpin RNA tags.[11-12] However, the need for inserting exogenous sequences into the target RNA may have undesired implications on RNA structures and functions.

A different approach for RNA labeling is the use of artificial nucleic acid based catalysts. For example, the epoxide reactive ribozyme derived from *A. prenix* genomic RNA,[13] and the fluorescein-iodoacetamide reactive ribozyme[14] have been developed for RNA labeling, although the exact site of reaction is not known in all cases. These ribozymes act via RNA self-alkylation and therefore need to be inserted into the RNA of interest. In contrast, the engineered twin-ribozyme cleaves a segment from the target RNA and replaces it with a labeled RNA patch.[15] Based on sequence complementarity, the twin-ribozyme can be adopted for various target RNAs, but the limited efficiency of the multistep reaction hampered widespread applications. Another intermolecular labeling approach with modified nucleotides was recently established with the engineered RNA polymerase ribozyme 24-3.[16] This primer extension-based labeling tool is limited to RNA 3'-end labeling. Covalent labeling at internal RNA nucleotides was achieved with DNA as a catalyst.[17-20] The original RNA-ligating 10DM24 DNA enzyme[19] was engineered to accept 2'/3'-labeled GTP substrates,[17] and catalyzes ligation to an internal adenosine via a 2'-5' branched phosphodiester bond.[18] However, the three-helix junction setup[18] and the fact that 10DM24 needs heavy metal ions (manganese, terbium) for optimal activity,[17, 20-21] limit its applications for sensitive RNA substrates and prospective *in vivo* applications. Cellular RNA labeling with nucleic acid catalysts could be more feasible with genetically encodable ribozymes.

In this work, we report versatile RNA catalysts (ribozymes) for general and efficient site-specific labeling of RNA targets *in vitro*. The ribozymes were identified by *in vitro* selection from a random RNA library and catalyze the attachment of a labeled nucleotide at an internal position under mild reaction conditions and with high yields. Such RNA catalysts were identified from a partially structured random RNA pool (N_{40}) using N^6-biotinylated ATP as a small molecule selection substrate (Figure 1). The constant region of the RNA pool contained the hypothetical substrate sequence and a pair of recognition arms complementary to the substrate sequence. An unpaired adenosine was chosen as the target site for the labeling reaction, flanked by two Watson-Crick base-paired guanosines in the binding arms (see Figure S1 for sequence). After each incubation step, the unreacted N^6-biotinylated ATP was removed by ethanol precipitation and the biotinylated RNA molecules (i.e. the active ribozymes) were immobilized on neutravidin- or streptavidin-coated magnetic beads. The beads were washed extensively before the bound RNA was eluted. The recovered RNA was reverse transcribed and the cDNA was amplified by PCR to generate the dsDNA template from which the enriched RNA pool entering the next selection round was transcribed by T7 RNA polymerase. (Experimental details and schematic of pool assembly and PCR amplification steps see Figure S1).
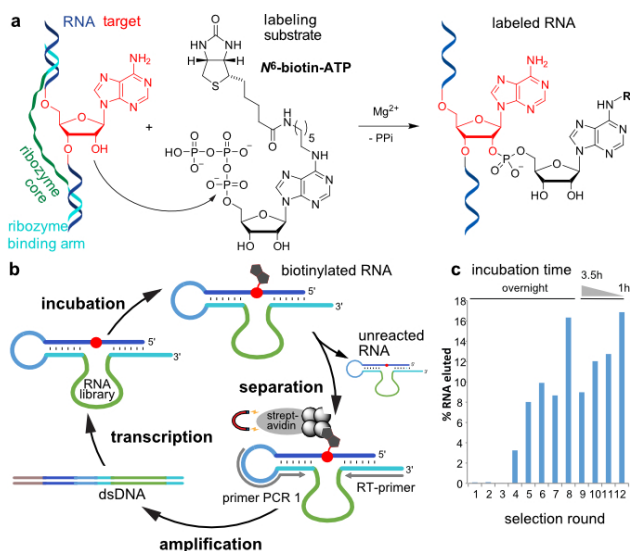
Figure 1. a) Ribozyme-catalyzed RNA labeling via 2'-5' phosphodiester at internal adenosine (red). b) Design of the RNA pool and *in vitro* selection cycle. Red sphere: the target adenosine; dark blue: substrate RNA; light blue: connecting loop; cyan: binding arms; green: N_{40} random region. Incubation: 200 μM N^6-biotin-ATP, 40 mM Mg^{2+}; 37°C, pH 7.5. Separation: streptavidin/neutravidin beads. Amplification: RT-PCR followed by PCR to re-install T7 promoter and substrate sequence in the dsDNA template. c) Progress of *in vitro* selection, monitored by fluorescence of eluted RNA.

Figure 2. a) Ribozyme core sequences of FH14, FH20 and FH31. b) *trans*-activity assay with parent RNA transcript (3'-fluorescein) and N^6-biotin-ATP. c) T1 digestion and alkaline hydrolysis of 5'-fluorescein-RNA substrate (ref) and biotin-labeled RNA product. The gap and shift identify the 2'-OH of A10 as the labeling site. d) Labeling of transition (TM) and transversion mutants (TV-1 and TV-2) of the parent RNA by complementary FH14 ribozymes (200 μM N^6-biotin-ATP, timepoints up to 4 h).

In each round, a fraction of the pool was fluorescently labeled at the 3'-end to allow monitoring of the selection progress (Figure 1c). The first sign of enrichment appeared at round 4, when ~3 % of the RNA was recovered after the capture step. The amount of recovered RNA rose steadily to ~16% RNA at round 8. To promote selection of faster reacting variants, the incubation time was decreased from overnight to 3.5 h at round 9. This resulted in a significant reduction in the yield of recovered RNA to ~9%. Despite further reduced incubation time, the level of eluted RNA increased in the following rounds; an incubation of 1 h led to ~17% recovery at round 12. A streptavidin gel-shift assay of the round 12 pool, revealed up to 70% biotinylated RNA after 6 h, implying that a major fraction of the pool was catalytically active at this stage (Figure S2). A preliminary test also suggested that some ribozymes in the enriched pool are able to catalyze the biotinylation reaction *in trans* (Figure S2).

The round 13 RNA pool was cloned, and individual clones were tested by the streptavidin gel shift assay (Figure S3). Plasmids from 22 clones were then sequenced (Table S1). Three major sequence groups were identified, and the representative sequences of these groups were denoted FH14, FH20 and FH31 (Figure 2). Single turnover activity assays of these variants revealed that all three ribozymes possessed the desired *trans*-activity, i.e. that the modification occurred in the substrate sequence and not on the ribozyme strand. The FH14 ribozyme turned out as the most efficient ribozyme with >60% labeled RNA after 2 h reaction time, followed by FH31 (> 50% in 4 h), FH20 (~ 20% in 4 h) (Figure 2b). In order to identify the site at which the biotinylated adenosine was attached to the target RNA, the product of the RNA-catalyzed labeling reaction was isolated and subjected to RNase T1 probing and alkaline
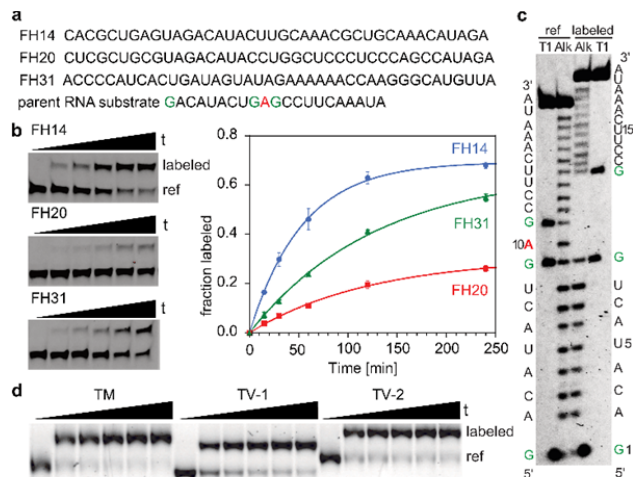
hydrolysis (Figure 2c). The unmodified substrate sequence was used as a reference and showed an unperturbed hydrolysis ladder. In contrast, the probing experiments of the biotin-labeled product revealed a gap in the hydrolysis pattern. The band corresponding to cleavage at A10 was missing, and the bands resulting from cleavage after A10 were shifted (slower mobility compared to reference). These results demonstrate that the modification was localized to the 2'-OH of A10, which corresponds to the designed labeling site in the structured RNA pool. Analogous results were obtained for the FH31 ribozyme (Figure S4). The attachment via a phosphodiester linkage was confirmed by mass spectrometry (Figure S5), and the concentration dependence (N^6-biotin-ATP and Mg^{2+}) was analyzed (Figure S6).

To establish the new ribozymes as general RNA labeling tools, we examined whether FH14 and FH31 can tolerate mutations in the RNA substrate. Three mutants of the selection substrate sequence and the ribozymes with corresponding complementary recognition arms were synthesized. For the initial tests, the GAG sequence context was left intact, while changing the other nucleotides to transition (TM) and transversion (TV1 and TV2) mutations. The single-turnover kinetic experiments revealed a high degree of flexibility regarding the positions outside the GAG context for both FH14 (Figure 2d) and FH31 ribozymes (Figure S7). The labeled RNAs were obtained with up to 80% yield after 30 min reaction time at 37°C. Surprisingly, the FH14 ribozyme was least efficient for labeling the parent RNA substrate, while this difference was less pronounced for FH31 (Figure S7).

Next, we introduced mutations to the GAG modification context, and tested the RNA substrate variants with the corresponding FH14 and FH31 ribozymes (Figure S8). When the reactive adenosine in the RNA substrate was mutated to guanosine, no detectable labeling activity was observed (Figure S9). The kinetic assays with other mutant sequences revealed that the FH14 ribozymes tolerated individual substitutions of both flanking Gs, while the

FH31 ribozyme demonstrated a stronger preference for maintaining the G nucleotide 3' to the modification site, while mutations to the 5'G were better tolerated (Figure S7).

To examine the small molecule substrate scope of the FH14 and FH31 ribozymes, we tested N^6-(6-aminohexyl)-ATP, which lacks the biotin moiety, and unmodified ATP in comparison to the biotinylated selection substrate. The results of the single turnover kinetic experiments revealed that biotin is not an essential substrate requirement, since the ATP analog with the free aminohexyl group gave similar yields with both FH14 and FH31 (Figure 3a, S10, S11). The reaction rate and efficiency with unmodified ATP were drastically lower than with N^6-modified ATP analogues. When ATP was replaced by GTP, no labeled product was detectable over the time course tested (Figure S9). These results suggest that the ribozymes can be used with a wide variety of N^6-labeled ATP derivatives. As demonstrated in Figure 3b, three different fluorophores were used, including fluorescein, ATTO550 and Cy3, which were attached via N^6-aminohexyl-ATP (structures shown in Figure S12). The FH14-catalyzed reaction resulted in >80% labeled RNA with 5FAM and ATTO550, while the yield of the Cy3-labeled RNA was ~50% after 5h (Figure 3b).
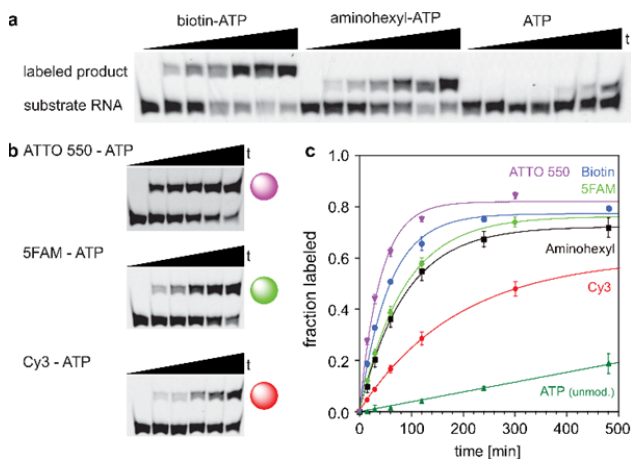

Figure 3. Substrate scope of FH14 ribozyme. a) FH14-catalyzed RNA labeling with N^6-biotin ATP, N^6-(6-aminohexyl)-ATP, and unmodified ATP on 5'-FAM-RNA target; timepoints up to 8 h. b) reactions with N^6-fluorophore-labeled ATP substrates (up to 5 h). c) single turnover kinetics of reactions shown in a) and b); k_{obs} values are in Table S2.

To demonstrate the application of FH14 for labeling biologically relevant RNA, we chose *E.coli* 5S rRNA as an abundant RNA of cellular origin. Three adenosines in different sequence and structural contexts of the 120-nt RNA were chosen as candidate sites for labeling (Figure 4a), and the corresponding FH14 variants were designed with binding arms of 8-10 nt flanking each labeling site (Figure S13). *In vitro* transcribed 3'-fluorescein labeled 5S rRNA was first used to test the labeling potential of the FH14 ribozymes. The time course experiment showed that ATTO550-ATP was attached at all three chosen sites after 2 h of incubation (Figure 4a). By combining two or three ribozymes, multiple chromophores were installed to enhance the brightness of the labeled product (Figure 4b). Moreover, *E.coli* 5S rRNA was selectively labeled by each of the tested FH14 ribozymes in the presence of total cellular RNA (Figure 4c). A fluorescein-labeled 5S rRNA *in vitro* transcript was used as a size standard. After imaging the ATTO550-labeled RNA, the gel was stained with SYBR Gold to visualize all RNA species present in the sample (right gel image).
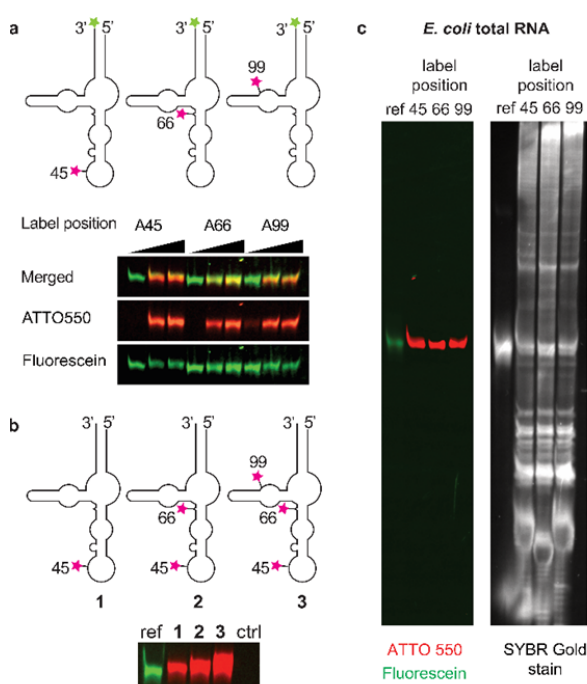

Figure 4. FH14-catalyzed labeling of 5S rRNA with ATTO550-ATP. a) Individual target sites addressed in a 3'-fluorescein-labeled *in vitro* transcript. b) Single, double and triple labeling of unmodified transcript. As a control (ctrl), the parent FH14 ribozyme was used, with binding arms that are not complementary to 5S rRNA. c) Labeling of *E.coli* 5S rRNA in total cellular RNA extract. Left: fluorescent image (overlay of fluorescein and ATTO550 channels), right: same gel stained with SYBR Gold.

In summary, the ribozymes reported in this study are highly versatile and efficient *trans*-acting ribozymes for site-specific labeling of RNA. They were identified directly by *in vitro* selection from a structured RNA pool and a selection procedure that allowed the enrichment of catalysts that target a pre-determined nucleotide as labeling position. This setup is advantageous compared to other *in vitro* selection efforts using unstructured RNA pools that resulted in *cis* reacting ribozymes that could not easily be engineered for intermolecular applications.[13-14] Both FH14 and FH31 ribozymes can easily be adjusted to target various RNA sequences simply by programming the binding arms complementary to the target RNA, and they show good scope for a broad range of N^6-labeled ATP derivatives. Using bioorthogonal nucleotide analogs, this selection design holds promise for the development of analogous ribozymes for RNA labeling *in vivo*.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website, and contains experimental details, sequences of DNA and RNA oligonucleotides, Supporting Tables and Supporting Figures S1-S13.

AUTHOR INFORMATION

Corresponding Author

claudia.hoebartner@uni-wuerzburg.de

Notes
The authors declare no competing financial interests.

REFERENCES

1. Schulz, D.; Rentmeister, A., Current approaches for RNA labeling in vitro and in cells based on click reactions. *ChemBioChem* **2014,** *15*, 2342-2347.
2. Gillingham, D.; Shahid, R., Catalysts for RNA and DNA modification. *Curr. Opin. Chem. Biol.* **2015,** *25*, 110-114.
3. Tomkuvienė, M.; Mickutė, M.; Vilkaitis, G.; Klimašauskas, S., Repurposing enzymatic transferase reactions for targeted labeling and analysis of DNA and RNA. *Curr. Opin. Biotech.* **2019,** *55*, 114-123.
4. Liu, Y.; Holmstrom, E.; Zhang, J.; Yu, P.; Wang, J.; Dyba, M. A.; Chen, D.; Ying, J.; Lockett, S.; Nesbitt, D. J.; Ferre-D'Amare, A. R.; Sousa, R.; Stagno, J. R.; Wang, Y. X., Synthesis and applications of RNAs with position-selective labelling and mosaic composition. *Nature* **2015,** *522*, 368-372.
5. Zhao, M.; Steffen, F. D.; Borner, R.; Schaffer, M. F.; Sigel, R. K. O.; Freisinger, E., Site-specific dual-color labeling of long RNAs for single-molecule spectroscopy. *Nucl. Acids Res.* **2018,** *46*, e13.
6. Schulz, D.; Holstein, J. M.; Rentmeister, A., A chemo-enzymatic approach for site-specific modification of the RNA cap. *Angew. Chem. Int. Ed.* **2013,** *52*, 7874-7878.
7. Holstein, J. M.; Schulz, D.; Rentmeister, A., Bioorthogonal site-specific labeling of the 5'-cap structure in eukaryotic mRNAs. *Chem. Commun.* **2014,** *50*, 4478-4481.
8. Motorin, Y.; Burhenne, J.; Teimer, R.; Koynov, K.; Willnow, S.; Weinhold, E.; Helm, M., Expanding the chemical scope of RNA: methyltransferases to site-specific alkynylation of RNA for click labeling. *Nucl. Acids Res.* **2010,** *39*, 1943-1952.
9. Tomkuvienė, M.; Clouet-d'Orval, B.; Černiauskas, I.; Weinhold, E.; Klimašauskas, S., Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. *Nucl. Acids Res.* **2012,** *40*, 6765-6773.
10. Li, F.; Dong, J.; Hu, X.; Gong, W.; Li, J.; Shen, J.; Tian, H.; Wang, J., A covalent approach for site-specific RNA labeling in Mammalian cells. *Angew. Chem. Int. Ed.* **2015,** *54*, 4597-4602.
11. Alexander, S. C.; Busby, K. N.; Cole, C. M.; Zhou, C. Y.; Devaraj, N. K., Site-Specific Covalent Labeling of RNA by Enzymatic Transglycosylation. *J. Am. Chem. Soc.* **2015,** *137*, 12756-12759.
12. Zhang, D.; Zhou, C. Y.; Busby, K. N.; Alexander, S. C.; Devaraj, N. K., Light-Activated Control of Translation by Enzymatic Covalent mRNA Labeling. *Angew. Chem. Int. Ed.* **2018,** *57*, 2822-2826.
13. McDonald, R. I.; Guilinger, J. P.; Mukherji, S.; Curtis, E. A.; Lee, W. I.; Liu, D. R., Electrophilic activity-based RNA probes reveal a self-alkylating RNA for RNA labeling. *Nat. Chem. Biol.* **2014,** *10*, 1049.
14. Sharma, A. K.; Plant, J. J.; Rangel, A. E.; Meek, K. N.; Anamisis, A. J.; Hollien, J.; Heemstra, J. M., Fluorescent RNA labeling using self-alkylating ribozymes. *ACS Chem. Biol.* **2014,** *9*, 1680-1684.
15. Welz, R.; Bossmann, K.; Klug, C.; Schmidt, C.; Fritz, H. J.; Müller, S., Site directed alteration of RNA sequence mediated by an engineered twin ribozyme. *Angew. Chem. Int. Ed.* **2003,** *42*, 2424-2427.
16. Samanta, B.; Horning, D. P.; Joyce, G. F., 3′-End labeling of nucleic acids by a polymerase ribozyme. *Nucl. Acids Res.* **2018,** *46*, e103.
17. Büttner, L.; Javadi-Zarnaghi, F.; Höbartner, C., Site-specific labeling of RNA at internal ribose hydroxyl groups: terbium-assisted deoxyribozymes at work. *J. Am. Chem. Soc.* **2014,** *136*, 8131-8137.
18. Höbartner, C.; Silverman, S. K., Engineering a selective small-molecule substrate binding site into a deoxyribozyme. *Angew. Chem. Int. Ed.* **2007,** *46*, 7420-7424.
19. Baum, D. A.; Silverman, S. K., Deoxyribozyme-catalyzed labeling of RNA. *Angew. Chem. Int. Ed.* **2007,** *46*, 3502-3504.
20. Carrocci, T. J.; Lohe, L.; Ashton, M. J.; Höbartner, C.; Hoskins, A. A., Debranchase-resistant labeling of RNA using the 10DM24 deoxyribozyme and fluorescent modified nucleotides. *Chem. Commun.* **2017,** *53*, 11992-11995.
21. Javadi-Zarnaghi, F.; Höbartner, C., Lanthanide cofactors accelerate DNA-catalyzed synthesis of branched RNA. *J. Am. Chem. Soc.* **2013,** *135*, 12839-12848.

Table of contents artwork