Electronic Supporting Information

An RNA catalyst that reacts with a mechanistic inhibitor of serine proteases

Sandeep Ameta and Andres Jäschke*

**Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, Heidelberg, 69120, Germany*
E-mail: jaeschke@uni-hd.de

General materials and methods

All reagents were purchased from Sigma-Aldrich (unless specified otherwise) and used without further purification. Reversed-phase HPLC purifications were performed on an Agilent 1100 Series HPLC system equipped with a diode-array detector. Denaturing polyacrylamide gels were prepared using gel stock solution from Roth and was always ran at 1X TBE (Tris-Borate EDTA, prepared from 10X TBE from Roth). For all the purposes water was used from a MilliQ water purifier system (Millipore). Non-radioactive RNA concentrations were measured on a NanoDrop-1000 UV-spectrophotometer (peqlab). Radioactive gels were scanned on a Typhoon 9400 Imager (Amersham Biosciences) using storage phosphor screens. Analysis and calculation of conversions were carried out with Image Quant software (Molecular Dynamics, Version 5.2).

Substrate for the selection

The substrate for the selection was synthesized by standard NHS-ester coupling chemistry between NHS-PEG_4-biotin and a known thrombin protease peptidyl-chloromethyl ketone inhibitor.[1] 5 mg of PPACK, NH_2-D-Phe-Pro-Arg-chloromethyl ketone (Molecular Innovations & LOXO GmBH) was added to 12 mg of Biotin-PEG_4-NHS (Thermo Scientific) with 600 µL of phosphate buffer, *p*H 7.0, and incubated at room temperature for 4 h. The coupled product was purified *via* HPLC on a reversed-phase C18 column (Phenomenex, 250 × 4.60 mm, 5 micron). A flow rate of 1 mL/min with water/acetonitrile (both with 0.1 % TFA) as solvent system was used under the following gradient: starting with 2 % of acetonitrile going to 60 % in 55 min then 98 % of acetonitrile for 7 min as cleaning step and finally to starting gradient. The product peak was collected, lyophilized (3 times) and dissolved in water for further use in selection. Product identity was confirmed by HR-ESI mass spectroscopy; Calculated (M): m/z 923.4415 g/mol, Found (M+H): m/z 924.4383 g/mol and purity was analyzed by reversed-phase HPLC (Figure S1).

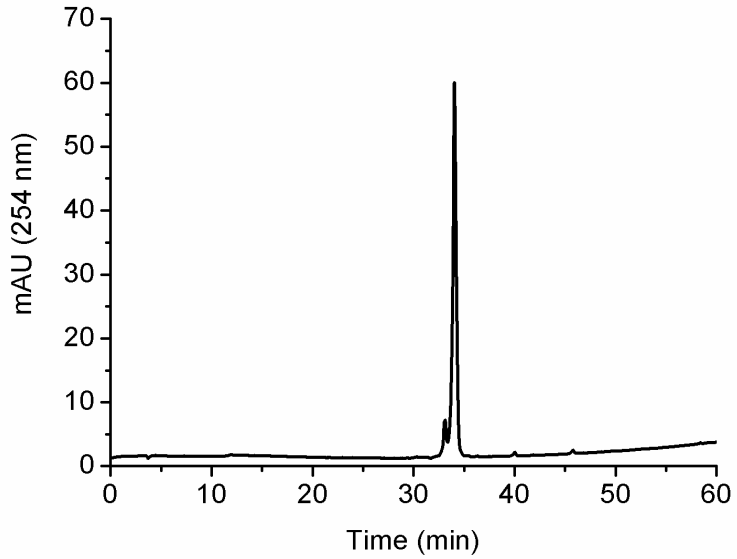

Figure S1. HPLC chromatogram of the substrate. Reversed-phase (C18) HPLC chromatogram of the substrate (Biotin-PEG_4-PPACK) after coupling and purification. A C18 column was used (Phenomenex, 250 × 4.60 mm, 5 micron) with the same gradient as used for purification.

DNA pool used for *in vitro* selection

For selection, a 252 nucleotide long dsDNA pool (N252) was constructed by assembling two single stranded DNAs, having 100 and 70 randomized positions respectively (sequence 1 & sequence 2, chemically synthesized by NOXXON AG, Berlin, Figure S2) *via* their inverse complementary sequences and subsequent 5 cycles of PCR amplification. Both sequences were mixed in an equimolar ratio along with 1X PCR buffer (Rapidozyme; 67 mM Tris, *p*H 8.8, 16 mM $(NH_4)_2SO_4$, 0.01 % Tween 20), 6 mM $MgCl_2$, 4 µM of each forward and reverse primer (IBA), 0.5 mM of each dNTP, 0.1 U/µL *Taq* polymerase (Rapidozyme) and subjected to the following conditions: strand filling 92°C/1min, 52°C/1 min, 72°C/1 min and amplification: 92°C/1 min, 62°C/1 min, 72°C/1 min with 4 additional cycles from step 2 and final amplification at 72°C/3 min. After PCR, amplified dsDNA pool was purified by NAP-25 gel filtration columns (GE Healthcare) and then ethanol precipitated. The purity of the dsDNA was checked on a 2 % agarose gel (with ethidium bromide, under standard electrophoresis conditions; 1X TBE, 180V, 40 min).

(a) Pool design

(b) Pool sequence

Sequence 1: 5'-GGAGCTCAGCCTTCACTGC- **N1**-GATGTTCGCATGATGTGCACTAGC-3'
Sequence 2: 5'-GTGGATCCGACCGTGGTGCC-**N2**-GCTAGTGCACATCATGCGAACATC-3'

(c) *In vitro* assembly of sequences to generate the 252 bp long pool (N252)

5'-GGAGCTCAGCCTTCACTGC- **N1**-GATGTTCGCATGATGTGCACTAGC-3'
3'-CTACAAGCGTACTACACGTGATCG-**N2**-CCGTGGTGCCAGCCTAGGTG-5'

(d) Primers used for amplification
Forward primer (Primer A): 5'-TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGC-3'
Reverse primer (Primer B): 5'-GTGGATCCGACCGTGGTGCC-3'

Figure S2. The schematic of the DNA pool and its assembly. a) Pool schematic showing the lengths of constant and randomized regions. b) Single stranded DNA sequences with randomized regions to construct the dsN252 pool, where N1 and N2 are random regions of 100 and 70 nucleotides, respectively. c) Schematic showing the assembly of the ssDNA sequences with the help of in-between constant region. d) Sequences of the primers used for the amplification. The forward primer introduces the promoter sequence (underlined) for T7 RNA polymerase.

Apparent *k*

Apparent *k* was calculated from the % immobilization of reacted RNAs from each round of the selection using the following:

$$k = \frac{1}{t(a-b)} \times ln \frac{b \times CA}{a \times CB}$$

where,

k = apparent k (k_{app}), t = time, a = substrate concentration, b = RNA concentration used for the selection reaction, CA = concentration of a at time t, CB = concentration of b at time t

Cloning, sequencing and activity determination of clones

The PCR product from round 13 was cloned into a TA-vector for Sanger sequencing using the TA-cloning kit (Invitrogen, Cat.No. K2020-20). The respective PCR products were freshly appended with 3‘ A-overhangs to insure better ligation into the TA-vector and then subjected to ligation overnight at 14°C. 3 µL of the ligation reaction was used for transformation in *E. Coli* DH5α cells. Cells were plated on LB-agar plates (containing 100 µg/mL of ampicillin, pre-incubated with X-Gal and IPTG) and incubated at 37°C for overnight. Colonies were picked using blue-white screening, positive clones were checked for the presence of an insert of the expected size by colony PCR, and plasmids were prepared from the respective colonies. 300-600 ng of each plasmid was mixed with 20 pmol of M13 reverse primer and sequenced at Seqlab (Seqlab, Göttingen, Germany).

From each clone’s plasmid, the insert was PCR amplified and *in vitro* transcribed (using same protocol as was used during selections) to determine their catalytic activity. 1-2 ng of plasmid was used for each PCR in a 30 µL reaction. *In vitro* transcriptions were performed using 4 mM of each NTP. The RNAs were purified over an 8 % denaturing polyacrylamide gel using standard electrophoresis conditions (1X TBE, 2W, 1.5h run). Bands were visualized by UV shadowing and excised. For activity assays, 10 µM of purified RNA was mixed with 300 mM NaCl, 200 mM KCl, 5 mM $MgCl_2$, 5 µM of $MnCl_2$ and 5 µM of $CaCl_2$, *p*H 5.5 in 10 µL reaction volumes. The reaction mixture was heated to 80°C for 3 min and cooled down to 10°C at a rate of 0.1°C/sec. Then it was kept at 10°C for 5 min and the substrate was added in 75-fold excess over RNA. The reaction mixture was incubated at 25°C. After 2h, the reaction was ethanol-precipitated and subjected to streptavidin-based gel shifts (see below). The gels were first visualized by UV-shadowing and then stained with SYBR Gold stain (Invitrogen). The stained gels were scanned using a Typhoon Scanner (GE Healthcare). The clonal RNAs were assigned to four categories according to their activity. Exceptionally active; all those whose streptavidin-shifted product band was already visible by UV-shadowing. Highly active; where a prominent streptavidin-shifted product was only visible after SYBR Gold staining. Less active; where only a faint product band was visible after SYBR Gold staining. Inactive; No streptavidin-shifted product band was observed. Out of 55 clones, 4 were exceptionally active, 22 were highly active, 12 were less active and 17 were inactive. One of the exceptionally active clones, RNA 19 was chosen for further analysis (Figure S3).

Sequence analysis

All the RNA sequences derived from clones were folded using Mfold web server (http://mfold.rna.albany.edu/?q=mfold/). The RNA structures were refolded with the help of VARNA 3-7 software (http://varna.lri.fr/) using minimum energy output from Mfold.

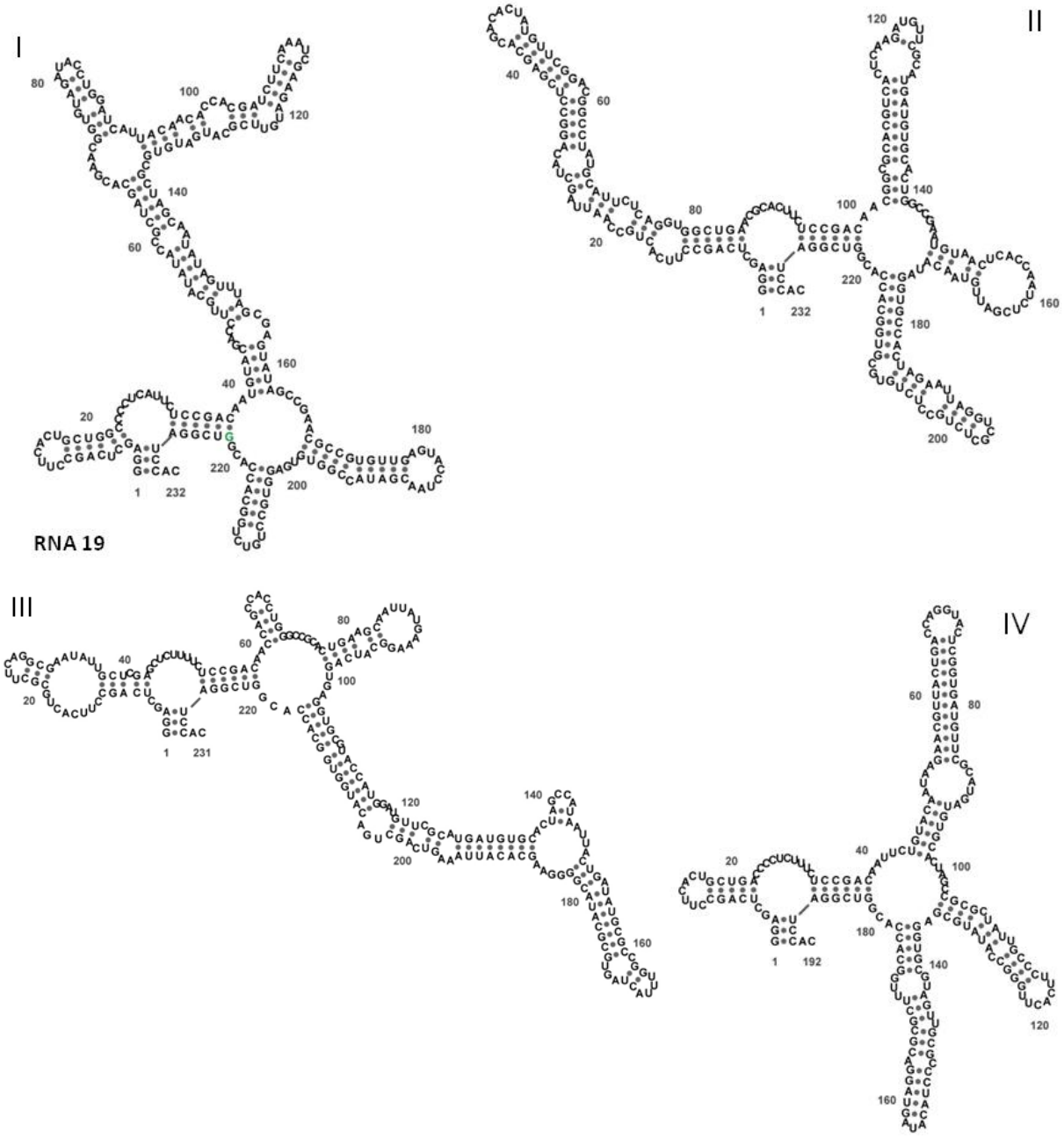

Figure S3. The minimum-energy predicted secondary structures for representative members of all four sequences families. They all share a common folded four-way junction structure (Figure 3). The RNA 19 from family I was characterized thoroughly in this study. The marked G nucleotide (G222, green) is identified as the reaction site in RNA 19 (Figure 6 and ESI page S9).

A 11mer sequence stretch, 5'-UUCUCCGACAA-3', was found to be conserved in all exceptionally and several highly active sequences. All sequences containing this motif can be assigned to four major families using multiple sequence alignments from a web server

application MultAlign (http://multalin.toulouse.inra.fr/multalin/) and subsequent secondary structure predictions from Mfold. Family I, II and IV were identified in the Sanger sequencing data, whereas family III was additionally identified when some sequences carrying this motif from next-generation sequencing data were analyzed. A detailed analysis of the evolution of catalytically active sequences by next-generation sequencing will be published elsewhere.[2] The representative members from each of these families can be folded into a common four stem-loop junction structure (Figure S3). The 11mer motif is found to base-pair with the conserved 3'-primer region in all the four families.

3' labeling of RNA

Unless specified otherwise, in all the analytical procedures, 3'-labeled ^{32}P-pCp RNA 19 was doped into the unlabeled RNAs (typically 0.2 to 0.6 pmol of pCp-labeled RNA per 100 pmol of unlabeled RNA). For 3' pCp labeling, 30 pmol of *in vitro* transcribed RNA 19 was mixed with ~ 34 pmol of ^{32}P α-pCp (10 µCi/µL, Hartmann Analytics, Germany) in 1X buffer (T4 RNA Ligase buffer, Fermentas), 0.1 mg/mL of BSA, 10 % of DMSO and 0.4 U/µL of T4 RNA Ligase (Fermentas). The reaction mixture was incubated at 16°C for overnight. The labeled RNA was purified over a 12 % polyacrylamide gel using standard electrophoresis conditions (1X TBE, 25W, 2 h run). The respective bands were excised, eluted and ethanol precipitated. The labeled RNA pellets were dissolved in water to give a final concentration of ~0.1 -0.3 pmol/µL.

Streptavidin-based gel shifts

To facilitate the separation between reacted and non-reacted RNAs, streptavidin-based gel shifts assays were used. After each reaction, RNA was ethanol precipitated and pellets were dissolved in water to give a final concentration ~2.6 pmol/µL. Then 2 equivalents of streptavidin (New England Biolabs) were added, mixed and the solution kept at room temperature for 1 min. After addition of gel loading buffer, the mixture was separated on a 12 % denaturing polyacrylamide gel using standard electrophoresis conditions (1X TBE, 25W, run for 2 h 30-40 min). Due to the streptavidin-biotin interaction, the product bands migrated slower than the rest of the RNAs.

Identification of the minimal structure

To identify the minimal active structure, smaller fragments were PCR-amplified from the RNA 19 template using region-specific primers. These fragments were used as templates for *in vitro* transcriptions and several smaller RNA fragments were generated. Subsequently, activity assays were performed using these fragments either alone or in combination with each other. The minimized active structure, RNA 19a, was obtained by shortening of the stem II in RNA 19 (Figure S3 and S4). The catalytic efficiency of RNA 19a was analyzed by performing activity assays from this 128nt long RNA (*in vitro* transcribed from a commercially ordered DNA template, Biomers) and comparing it with the full-length RNA 19. RNA 19a was found to retain 19% of the catalytic activity from the wild-type full-

length RNA 19. The minimized RNA 19a structure was found to contain the proposed conserved four-way junction stem-loop features (Figure 3 and S4). When 5'-3' base-paring in RNA 19a was perturbed by shortening the 3'end by 6nt (to red arrow, Figure S4), no self-modifying activity was observed. Similarly, when the complete 5' stem loop region was removed (from nucleotide 1 to 33, to green arrow, Figure S4), the activity of the resulting construct was also abolished.

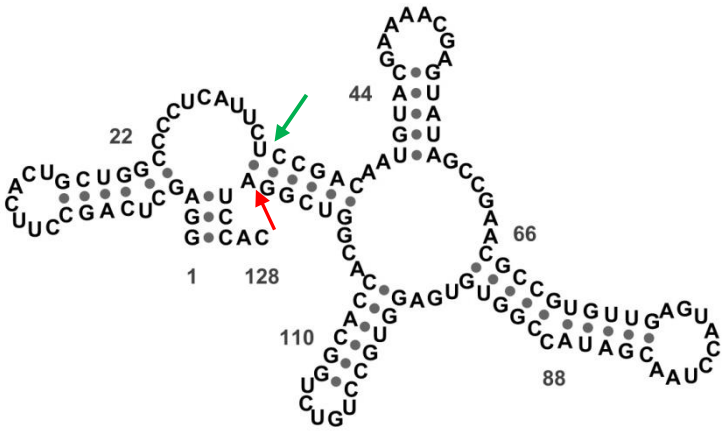

Figure S4. Sequence and predicted secondary structure of RNA 19a. For details, see text above the figure.

Model substrate catalysis

RNA 19 was also subjected to different assays involving model protease substrates (Table S1) but no measureable catalysis over the background could be measured. All measurements were performed in 384-well plates (Greiner, black wells, transparent flat bottom) using a TECAN safire^2 multi-plate reader (Tecan Group Ltd.). 10 µM of RNA 19 was mixed with 1X of reaction buffer in PCR tubes, incubated at 80°C for 3 min, and subsequently cooled down to 10°C (at a rate of 0.1°C/sec). Then the respective substrates were added to a final concentration of 0.75-2 mM and the reactions were incubated at 25°C for 14-16 h. The complete reaction mixture was then transferred to a 384 multi-well plate and fluorescence (FL) or absorbance (Abs) were measured as described.[3, 4]

Table S1: Catalysis of model protease substrates by RNA 19

Substrate	**FL/ Abs (AU)**	
	(-) control	***Selected RNA***
AFC: NH_2-D-Phe-Pro-Arginine-4-trifluoromethylcoumarin-7-amide	19559	18431
PPA-pNA: NH_2-D-Phe-Pro-Arginine-p-nitroanilide	0.89	0.92
Arg-pNA: Z-Arginine-p-nitroanilide	2.3	3
Arg-pNE: Z-Arginine-p-nitrobenzyl ester	0.4	0.2

Supplier's information: AFC (Anaspec), PPA-pNA (Thermo Scientific), Arg-pNA (Bachem), Arg-pNE (Bachem)

Substrate recognition characteristics

a) Competition assay

To probe the substrate recognition characteristics of RNA 19, competition assays with close analogues of arginine were performed. 10 µM of RNA 19 was mixed with 1X reaction buffer, *p*H 5.0, heated to 80°C for 3 min and cooled down to 10°C at the rate of 0.1°C/sec. Then competitor (Figure S5) was added to the final concentration of 3.5 mM and incubated at 10°C for 5 min. The reaction was then supplemented with 800 µM of the substrate (PPACK) and incubated at 25°C for 90 min.

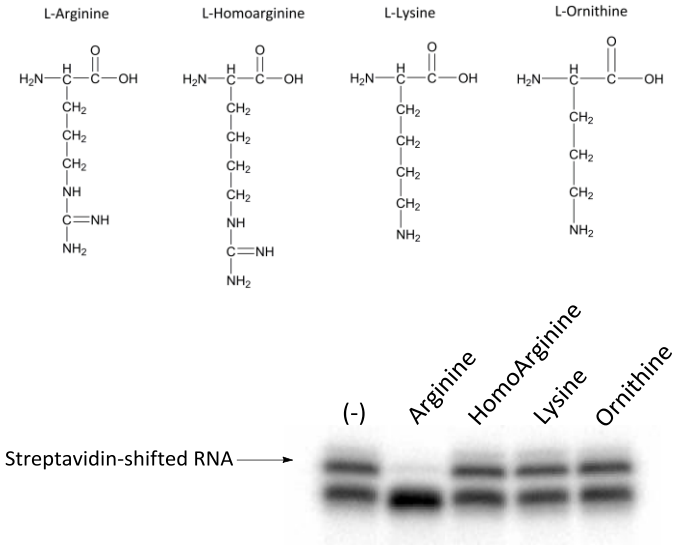

Figure S5. Competition assay. A streptavidin-based gel shift analysis of the competition assay was conducted where RNA 19 was pre-incubated with different amino acids (top panel) and then reacted with the PPACK substrate.

The RNAs were ethanol precipitated and subjected to streptavidin-based gel analysis. Only arginine was found to efficiently block the substrate binding site in RNA 19 (Figure S5).

b) Requirement for biotin and the linker

The requirement for biotin and the tetra (ethylene glycol) linker was also probed using α-^{32}P-GTP-labeled RNA 19 and nuclease P1. The protocol used here is the same as that used in the identification of the reacting nucleotide (see below). *In vitro* transcribed RNA 19 was reacted with PPACK, both with and without carrying biotin (Figure S6) and digested to mono-nucleotides using nuclease P1. The digested products were analyzed by 15 % denaturing PAGE under standard electrophoresis conditions (1X TBE, 25 W, 2 h run). Due to the attached substrate peptide, the reacted nucleotide migrates slowly. The slowly migrating reaction product was observed with both types of PPACK (with and without biotin and tetra-glycol spacer), confirming that both features are not required for the reaction (Figure S6).

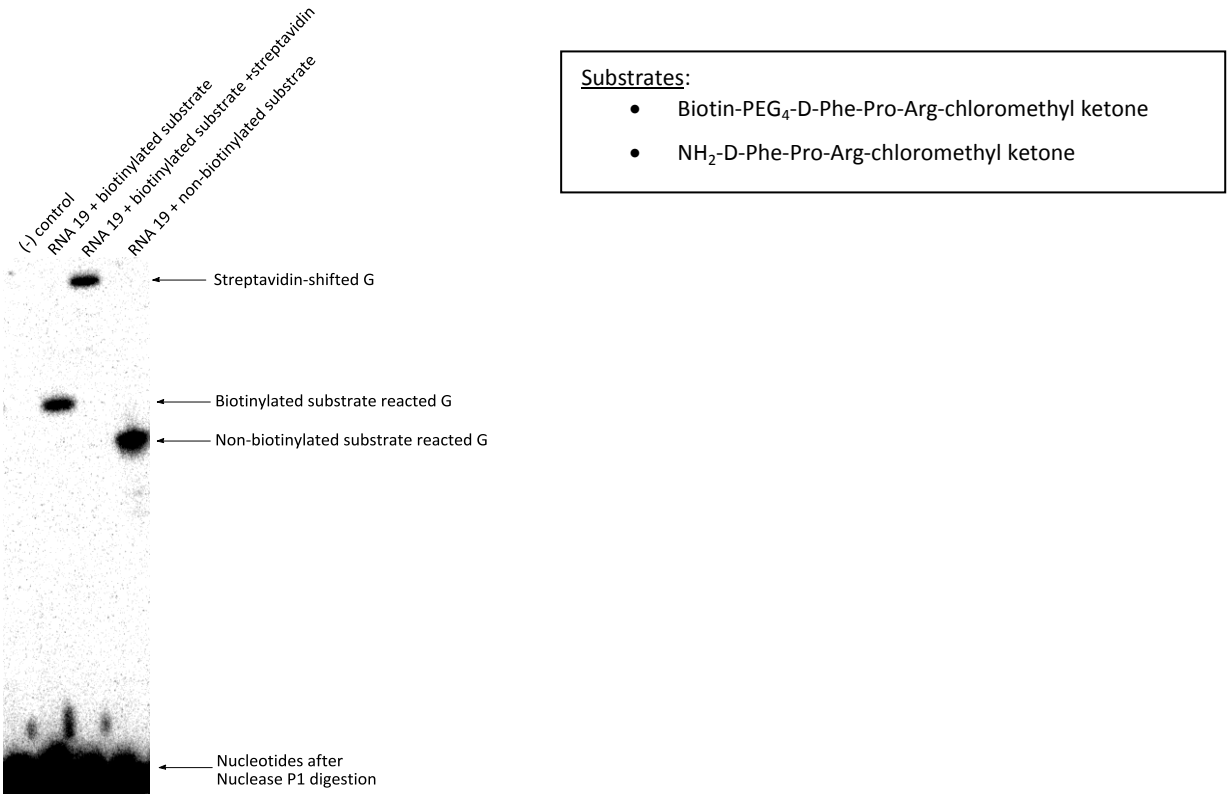

Figure S6. Requirement for biotin and the linker. 15 % denaturing PAGE analysis for the reaction of RNA 19 with PPACKs (with and without carrying biotin and tetra (ethylene glycol) linker). The product band was observed in both cases. The product band in the reaction with biotinylated PPACK could be further shifted upon addition of streptavidin.

Identification of the reaction site

a) Type of nucleotide

To learn at which nucleotide the substrate is reacting, RNA 19 was *in vitro* transcribed where either one of the four NTPs was supplied as a α-^{32}P-NTP along with the other non-radioactive nucleotides. *In vitro* transcription was done following the standard protocol (as mentioned earlier, see *in vitro* selection) in a 25 µL scale using 100 µCi of radioactive nucleotide (Hartmann Analytics) supplemented with 0.4 mM of the respective non-radioactive nucleotide. The other three non-radioactive NTPs were provided in a final concentration of 4 mM each. A master mix was prepared using NTPs, 1X transcription buffer, 0.01 mg/mL of BSA, 10 mM of DTT, 10 U/µL of T7 RNA polymerase and then 100 µCi of the respective radioactive NTP was added in each reaction. The reaction was incubated for 4h at 37°C and transcripts were purified over an 8 % denaturing polyacrylamide gel. 10 µM of radioactive RNA from each NTP transcripts were reacted with 750 µM of the peptidyl chloromethyl ketone substrate in a 10 µL volume applying the protocol described above. After incubation for 2h, the reaction mixture was ethanol-precipitated and pellets were resuspended in 16 µL of water. An 8 µL sample was withdrawn and mixed with 20 mM of NH_4-Acetate *p*H 4.3, 0.5 mM of $ZnCl_2$, 1.2 U of Nuclease P1 (Roche) (20 µL final volume) and incubated at 55°C for 3h. The reaction was stopped by adding 20 µL of gel loading buffer. The reacted and unreacted nucleotides were separated on a 15 % polyacrylamide gel. Gel analysis revealed

that the substrate reacted only at a guanine and not at any of the other three types of nucleotides in RNA 19 (Figure S7).

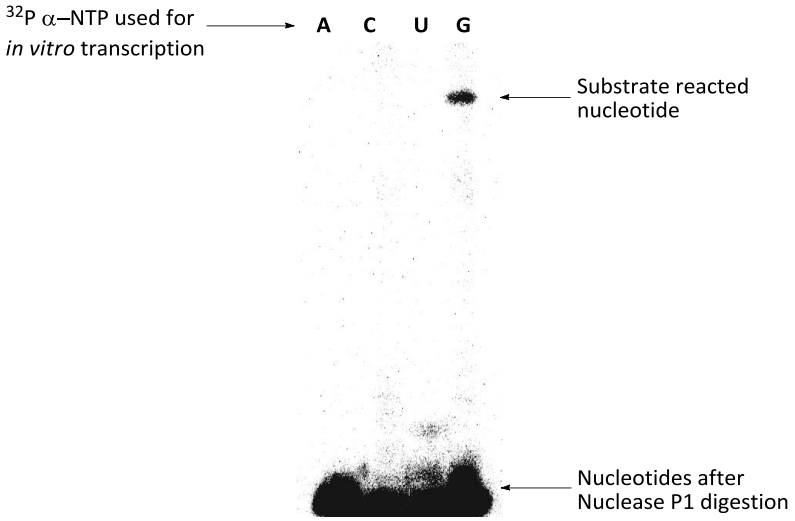

Figure S7. Nucleotide of the reaction. 15 % denaturing PAGE analysis of the nucleotide at which the substrate has reacted. Due to the presence of the peptidyl substrate, the nucleotide at which substrate is attached migrates much slower, compared to other nucleotides. Lanes; Lane1-4, each respective radioactive NTP transcript was subjected to P1 nuclease degradation after the reaction substrate. A, C, U, G represents each respective NTPs reaction. Gel: Gel was under standard gel electrophoresis conditions (1X-TBE, 25W, 2-3h).

b) DNAzyme cleavage of a 36mer RNA fragment from RNA 19

To generate a smaller RNA fragment for the identification of the reaction site, a 10-23 DNAzyme[5] was designed to cleave off the last 36 nucleotides from the the 3'-terminus of RNA 19 (Figure S8).

a)

RNA 19 sequence

5'-GGAGCUCAGCCUUCACUGCUGGCCCCUCAUUCUCCGACAAUGUACGACCUUGCAUAUACCGC
UAGCACGAACGGUGUAGAUACCUGGAUCAUUACAACACCACGAUCUUCAAAUCGAAGAUGUUC
GCAUGAUGUGCGCUAGCAAUAUAGUUUAGCGAGUAUAGCCGAACGCCGUGUUGAGUACCUAAC
GAUACCG**GU**GUGAGGUGCCUGUCUGGCACCACGGUCGGAUCCAC-3'

b)

10-23 DNAzyme sequence

5'-GGCACCTCACAGGCTAGCTACAACGACGGTATCGTTAGG-3'

Figure S8. RNA 19 sequence and the designed 10-23 DNAzyme. a) Complete sequence of RNA 19 with highlighted bases (bold and underlined) as the site of cleavage for the 10-23 DNAzyme. Cleavage will generate a 36 nucleotides long RNA fragment starting with a 5'U. b) 10-23 DNAzyme sequence designed for the site-specific cleavage of RNA 19. The highlighted parts are the two RNA binding arms (11 and 13 nucleotides long) and in between is the 15 nucleotides long catalytic core.

After several optimizations, the final conditions for the DNAzyme cleavage of RNA 19 were found to be as follows: 4 µM of RNA (with and without reaction with the substrate), 1X DNAzyme buffer (50 mM Tris-HCl *p*H 7.5, 10 mM $MgCl_2$, 100 mM KCl), 10-fold excess of 10-23 DNAzyme, 0.5- 1 pmol of ^{32}P α-pCp labeled RNA 19 were added in a PCR tube (20 µL reaction volume). The reaction was heated to 80°C for 30 sec, subsequently cooled down to 37°C at the rate of 0.5°C/sec and then incubated at 37°C for 3 min. To increase the amount of cleaved product, this process was cycled 20 times.[6] The reaction was stopped by adding gel loading buffer and the cleaved product was separated from non-cleaved RNA on a 18% denaturing polyacrylamide gel under standard electrophoresis conditions (1X TBE, 25W, 3h run) (Figure S9). Bands of 36mer RNA with and without modification were excised, eluted overnight in 0.3M Na-Ac, *p*H 5.5, and iso-propanol precipitated. The pellets were dissolved in water and used for further analysis.

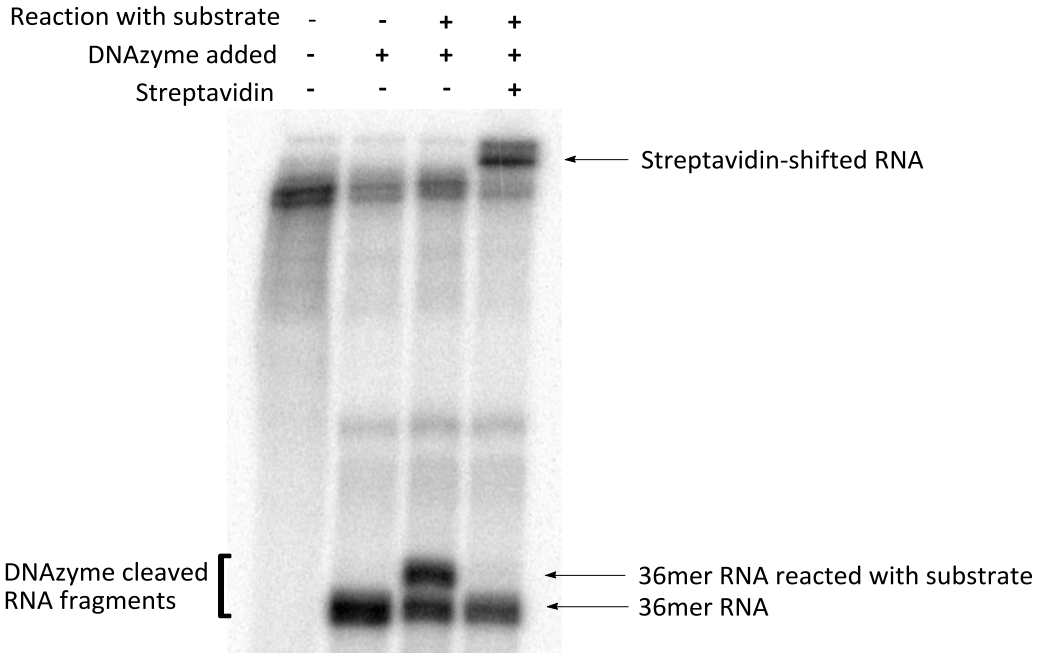

Figure S9. Gel analysis of the DNAzyme cleavage of RNA 19. Both the reacted and non-reacted RNAs were subjected to DNAzyme cleavage. Lanes; Lane1: (-) control only RNA 19 without DNAzyme. Lane2: RNA 19 without reaction with the substrate and subjected to DNAzyme cleavage. It yields a 36mer RNA without substrate attached. Lane3: RNA 19 reacted with the substrate and subjected to DNAzyme cleavage. It yields 36mer RNAs both with and without the substrate attached, as reaction efficiency for this sample was 50-60 %. Lane4: same sample as Lane3 but with streptavidin added to shift the substrate-attached 36mer, as only this contains the biotin group.

d) Site-directed mutagenesis

To support further that the identified reaction site, G222, is the only reaction site in RNA 19, G222 was systematically replaced with A and U and the catalytic activities of these mutant RNAs were compared with the wild-type RNA 19. This was done *via* primer based site-directed mutagenesis (Figure S10). The site-directed mutagenesis was done with and without co-substitution of the complementary base-pairing nucleotide in order to rule out any structural distortions (Figure S10 and S3). First, dsDNA templates were generated from the

wild-type RNA 19 plasmid *via* PCR using primers with the respective mutations (Figure S11). 18 pg of wild-type RNA 19 plasmid was mixed with 1X Hybrid-PCR buffer (10 X Hybrid-X-Polymerase buffer), 0.5 µM of each respective forward & reverse primer (Biomers), 0.2 mM dNTP, 0.02 U/µL of Hybrid-X-Polymerase in a 200 µL PCR reaction. The amplification was done using following protocol; Step1: 98°C for 30 sec, Step2: 98°C for 30 sec, Step3: 52°C for 30 sec, Step4: 72°C for 30sec, Step5: go to Step2 for 25 times, Step6: 72°C for 7 min, Step7: end. The PCR products were analyzed on a 2 % agarose gel and subsequently purified by a Qiagen-PCR purification kit. For the introduction of single mutations (G222U & G222A), only one PCR was performed with Primer A and mutated reverse primer, whereas in the case of double mutations (C38U, G222A and C38A, G222U), the first PCR was done with both forward and reverse mutated primer (as mentioned above) and then 1 ng of these purified PCR products was further PCR amplified using primer A and mutated reverse primer in order to add the T7 promoter region. After PCR, products were analyzed on a 2 % agarose gel (with ethidium bromide, under standard electrophoresis conditions; 1X TBE, 180V, 40 min) and used directly for *in vitro* transcriptions without any further purifications. Mutant transcripts were produced using the standard *in vitro* transcription protocol and doped with the respective 3'-pCp labeled RNAs.

a) RNA 19 template sequence (as a insert in RNA 19 plasmid)

TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGCTGGCCCCTCATTCTCCGA**C**AATGTACGA
CCTTGCATATACCGCTAGCACGAACGGTGTAGATACCTGGATCATTACAACACCACGATCTTCAAA
TCGAAGATGTTCGCATGATGTGCGCTAGCAATATAGTTTAGCGAGTATAGCCGAACGCCGTGTTGA
GTACCTAACGATACCGGTGTGAGGTGCCTGTCTGGCACCACG**G**TCGGATCCAC

b) Primers used

Mutated forward primers:

5'-GGAGCTCAGCCTTCACTGCTGGCCCCTCATTCTCCGA**X**AATGTACG-3'

Mutated reverse primers:

5'-GTGGATCCGA**X**CGTGGTGCCAGACAGGCACCTCACACCGG-3'

X = **A or T** (RNA will have U or A at this position)

Primer A:

5'-TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGC-3'

Figure S10. RNA 19 template sequence and primers used for site-directed mutagenesis. a) RNA 19 sequence as an insert in RNA 19 wild-type plasmid highlighted nucleotides are changed to A or U *via* mutated primers during PCR. b) Primers used for the site-directed mutagenesis. The T7 promoter region is highlighted (in gray) in primer A.

Both single mutant (G222A and G222U) and double mutant (C38A, G222U and C38U, G222A) RNAs were reacted with the substrate and their reactivity was compared with the wild-type RNA 19. For the activity assay, 10 µM of RNA was reacted with an 80-fold excess of substrate in 1X reaction buffer of *p*H 5.0, following the protocol reported above.

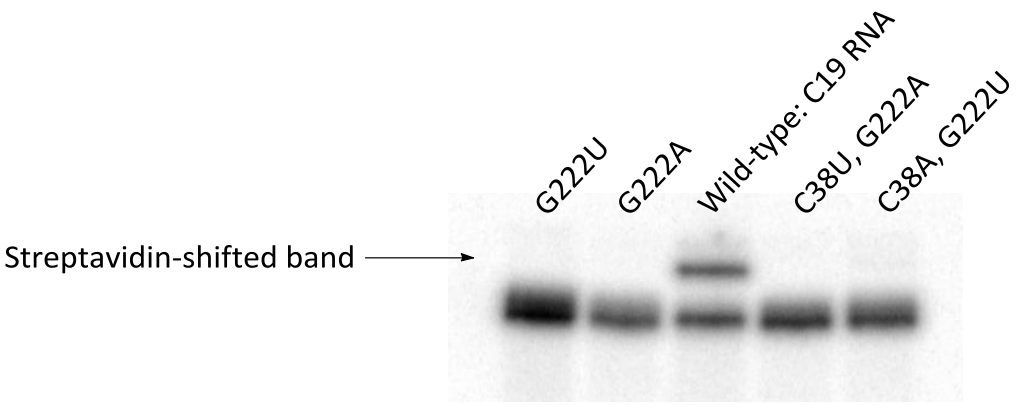

Figure S11. Streptavidin-based gel shift assay to analyze the activities of mutants. The RNA 19 reaction site G222 was mutated to A and U *via* PCR-based site-directed mutagenesis. Activities of the mutants were compared with wild type RNA 19 using a streptavidin-based gel shift assay. Gel; 12 % denaturing polyacrylamide gel was run under standard electrophoresis conditions (1X TBE, 25 W, 2h 40 min).

After reaction, reacted RNAs were analyzed using streptavidin-based gel shift assays (Figure S11). The mutated RNAs were found to be completely inactive, thereby supporting G222 as the single reaction site in RNA 19.

Characterization of the reaction product

To characterize the nature of the product formed between the substrate and RNA 19, several attempts were made. RNA 19 is too large to obtain a good resolution in mass spectrometry, with and without reaction, as it forms several salts adducts and the mass difference of the reaction product will not be sufficient to characterize the nature of the product.
First, 36mer RNA fragments (generated from DNAzyme experiments, see above) were subjected to MALDI-TOF mass spectrometry. While we could obtain a mass of both reacted and non-reacted RNA fragments but resolution was not sufficient to resolve the isotopic patterns (data not shown).

RNA 19 sequence

5’-GGAGCUCAGCCUUCACUGCUGGCCCCUCAUUCUCCGACAAUGUACGACCUUGCAUAUACCGC
UAGCACGAACGGUGUAGAUACCUGGAUCAUUACAACACCACGAUCUUCAAAUCGAAGAUGUUC
GCAUGAUGUGCGCUAGCAAUAUAGUUUAGCGAGUAUAGCCGAACGCCGUGUUGAGUACCUAAC
GAUACCGGUGUGAGGUGCCUGUCUGGCACCACGGUCGGAUCCAC-3’

rDNA3

5’-CCGTGGTGCC-3’

Figure S12. RNA 19 and ssDNA sequences used for ribonuclease H cleavage. The highlighted region (in gray) is the hybridization region for rDNA3 and the underlined nucleotides are the most prominent cleavage site under our reaction conditions. RNase H mainly cleaves between A and C at 3’ end of A. G222 is highlighted in red.

To improve the resolution, shorter RNAs were generated using ribonuclease H (RNase H). Ribonuclease H is known to cleave the RNA strand in RNA-DNA hybrids.[7] Several ssDNAs

were designed to hybridize to the 3'-end of RNA 19, in order to obtain RNA fragments as small as possible, and carrying the modification. One of the ssDNA, rDNA3, turned out to induce efficient cleavage of RNA 19, both with and without substrate attached (Figure S12 & S13). For cleavage, 2 µM of RNA 19 (with and without prior reaction with the substrate) was mixed with 1X ribonuclease H reaction buffer (10X RNase H reaction buffer, Fermentas) and 20 µM of rDNA3 (10-fold excess to the RNA) in a 40 µL of reaction volume. To visualize the cleaved fragments, the reaction was doped with 0.3 pmol of 3'-^{32}P-labeled RNA 19. The reaction mixture was then heated to 80°C for 3 min and subsequently incubated on ice for another 3 min. Then 0.4-0.5 U/µL of ribonuclease H (Fermentas) was added and the reaction mixture was incubated at 37°C for 1h with subsequent heat inactivation at 70°C for 10 min. The rDNA3 was digested using DNase I treatment, therefore adding 1X DNase I buffer (10X DNase I reaction buffer, Fermentas) and 2U/µL DNase I enzyme (Fermentas) to the reaction mixture. The reaction mixture was incubated at 37°C for 30 min and stopped by adding gel loading buffer. The cleaved fragments of RNA 19 were analyzed on a 18 % denaturing polyacrylamide gel using standard gel electrophoresis conditions (1X TBE, 25 W, 2h 30 min). As expected, several cleaved fragments were generated from RNA 19 using rDNA3 and ribonuclease H. One of the fragments carried the substrate modification, as it was shifted upon addition of streptavidin (Figure S13). There was no cleavage observed in the absence of rDNA3. One of the cleaved fragments was identified as 13mer (both with and without the substrate modification) by MALDI-TOF mass spectrometry. These 13mer fragments were further analyzed by high-resolution ESI-TOF mass spectrometry.

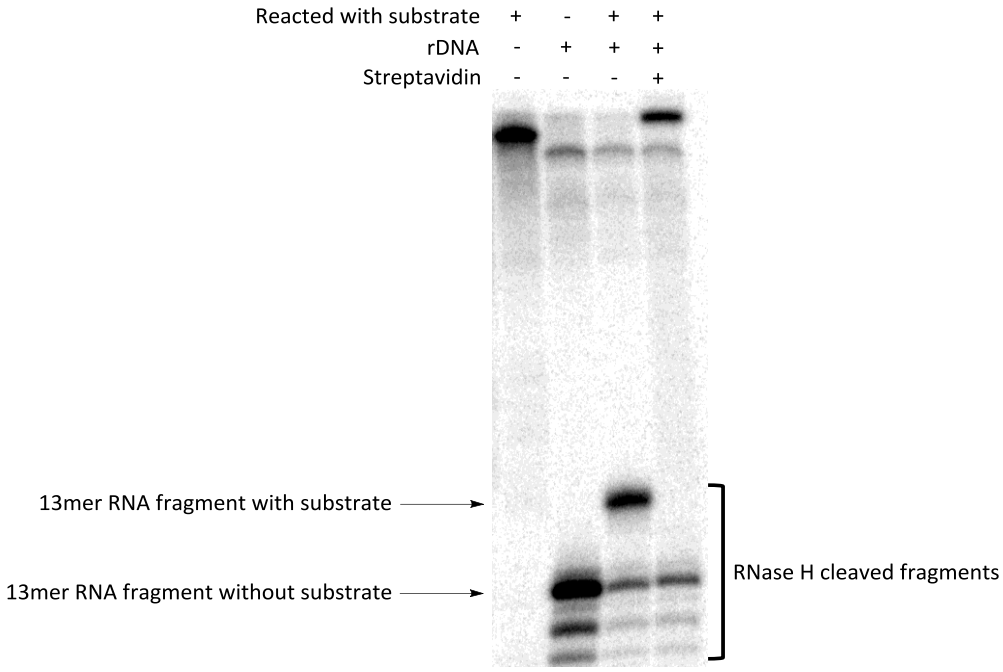

Figure S13. Ribonuclease H cleavage of RNA 19. Both the substrate-reacted and non-reacted RNA 19 was subjected to ribonuclease H cleavage in the presence of rDNA3. The cleaved fragment carrying the biotinylated substrate is shifted by streptavidin. Lanes, Lane1: (-) control, ribonuclease H treatment of the substrate-reacted RNA 19 in the absence of rDNA3. Lane2: RNA 19 (without substrate reaction) subjected to ribonuclease H cleavage in the presence of rDNA3. Lane3: Same as Lane2 but with the substrate-reacted RNA 19. Lane4: same sample as Lane3 but with streptavidin added.

Table S2: Calculated and found masses of the substrate and product

Sample	Information	Found/ Observed	Calculated	Deviation
Substrate	Biotin-PEG_4-PPACK	923.4383	923.4415	7.2 ppm
Control RNA fragment	5'-PO_4-CGGUCGGAUCCAC-OH-3'	4194.5818	4194.5651	3.9 ppm
Reacted RNA fragment	5'-PO_4-CG**G**UCGGAUCCAC-OH-3' (Reacted at **G**)			
1) Alkylation	N7 alkylation	5083.0021	5083.0263	4.9 ppm
2) Hydration	Additional hydration of ketone	5101.0180	5101.0368	3.7 ppm

Note: Substrate was measured in the positive mode and RNA fragments were measured in negative mode in ESI.

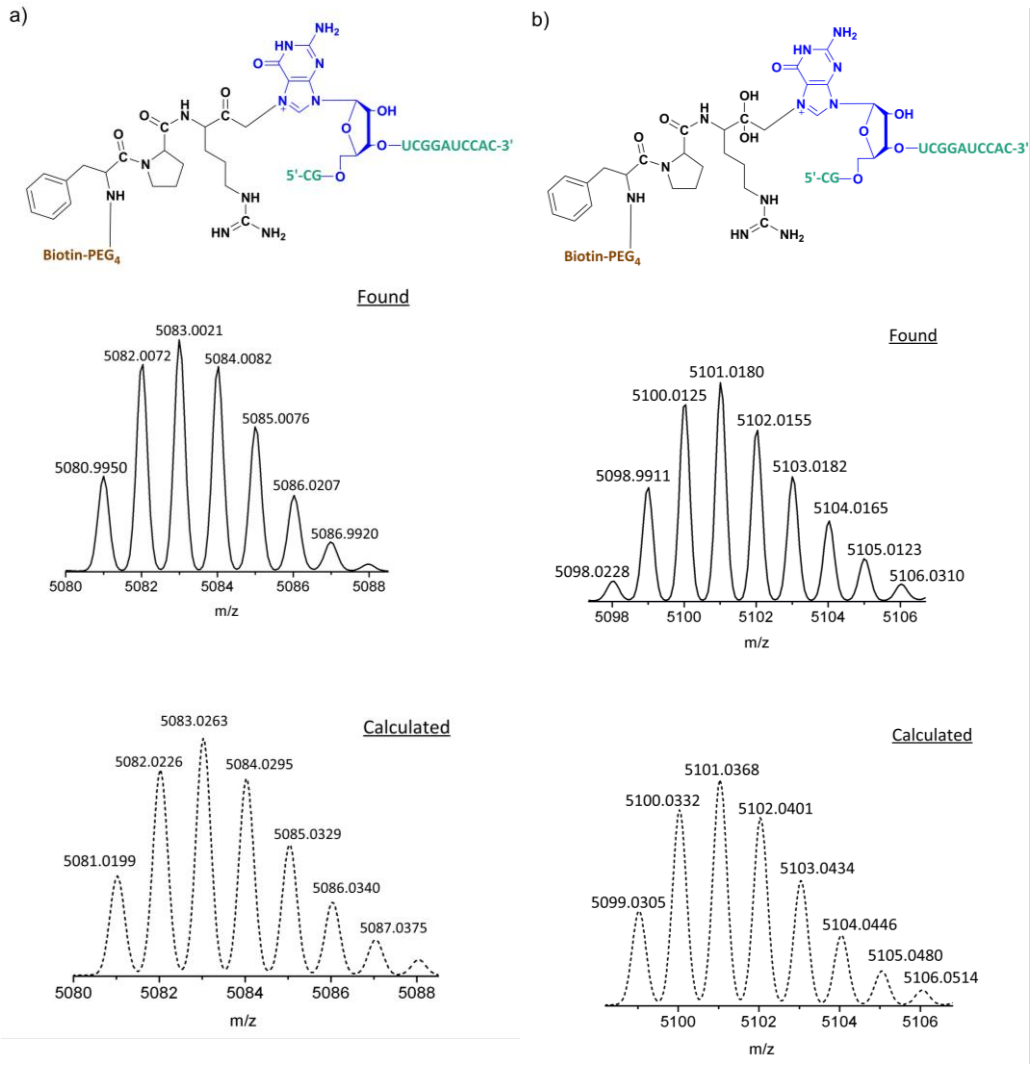

Figure S14. Mass analysis of ribonuclease H-cleaved 13mer RNA fragment. 13mer RNA fragment from RNA 19 both with and without substrate modification was characterized by ESI-TOF. The found masses are shown in continuous lines (top spectra) and calculated ones are shown with the dotted lines (bottom spectra). a) Structure and of the product arising from N7 alkylation. b) The product of the hydration of the ketone. Samples were measured in the negative mode.

Mass analysis confirmed the formation of both, the alkylation product and its keto hydrate. The mass spectra of the both compounds matched their respective calculated masses and isotopic patterns (Figure S14, Table 2).

References

1. J. C. Powers, J. L. Asgian, O. D. Ekici and K. E. James, *Chem. Rev.*, 2002, **102**, 4639-4750.
2. Ameta S., Winz, M.-L., Jäschke, A., *Unpublished work*.
3. K. Higuchi, A. Kajiki, M. Nakamura, S. Harada, P. J. Pula, A. L. Scott and A. M. Dannenberg, Jr., *Inflammation*, 1988, **12**, 311-334.
4. P. Schack and N. C. Kaarsholm, *Biochem. J.*, 1984, **219**, 727-733.
5. M. J. Cairns, A. King and L. Q. Sun, *Nucleic Acids Res.*, 2003, **31**, 2883-2889.
6. M. Hengesbach, M. Meusburger, F. Lyko and M. Helm, *RNA*, 2008, **14**, 180-187.
7. H. Donis-Keller, *Nucleic Acids Res.*, 1979, **7**, 179-192.