Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Polymerase-Mediated Site-Specific Incorporation of a Synthetic Fluorescent Isomorphic G Surrogate into RNA

*Yao Li, Andrea Fin, Lisa McCoy, and Yitzhak Tor**

anie_201609327_sm_miscellaneous_information.pdf

Supplementary Information

1. Materials and methods

Reagents, buffers and salts were purchased from Sigma-Aldrich, Fluka, TCI, and Acros and were used without further purification unless otherwise specified. NTPs were purchased from Fisher, the enzymes were purchased from New England Biolabs or Promega. The oligonucleotides were purchased from IDT and further purified by gel electrophoreris and subjected to standard desalting protocols.

Solvents were purchased from Sigma-Aldrich and Fisher Scientific, and dried by standard techniques. NMR solvents were purchased from Cambridge Isotope Laboratories (Andover, MA). All reactions were monitored with analytical TLC (Merck Kieselgel 60 F254). Column chromatography was carried out with Teledyne ISCO Combiflash Rf with silica gel particle size 40-63 μm. NMR spectra were obtained on Varian Mercury 400 MHz and Varian VX 500 MHz. Mass spectra were obtained on an Agilent 6230 HR-ESI-TOF MS at the Molecular Mass Spectrometry Facility at the UCSD Chemistry and Biochemistry Department.

2. Synthetic procedures

Ribonucleoside, th**G**, was synthesized as previously published. [S1]

2.1. RNA oligonucleotides synthesis

Transcripts **E1**, **S1**, **S2**, **S3** as well as donor strands **1b**, **2b**, **3b**, **4b** and **5b** were transcribed using T7 RNA polymerase (Figure S2). Each single strand DNA template was annealed to a consensus 18-mer T7 RNA polymerase promoter in TE buffer (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.8) by heating a 1:1 mixture (10 μM) at 90 °C for 3 min and cooling the solution slowly down to room temperature. Large scale transcription reactions were performed in buffer (40 mM Tris-HCl, pH 7.9) containing annealed templates (500 nM), dithiothreitol (10 mM DTT), NaCl (10 mM), spermidine (2 mM), RNase inhibitor (RiboLock, 1 U/μL), and T7 RNA polymerase (0.1μg / μL) in a total volume of 250 μL. For the native transcripts **E1**, **S1** and **S2**, ATP (2 mM), GTP (2 mM), CTP (2 mM), UTP (2 mM) and $MgCl_2$ (20 mM) were used.

Various concentrations of NTPs and th**G** as well as $MgCl_2$ were used for different templates to prepare 5'-thG transcripts, as indicated below. For template **1**: ATP (1 mM), GTP (1 mM), CTP (1 mM), UTP (1 mM), th**G** (5 mM) and $MgCl_2$ (10 mM); template **2**: ATP (2.5 mM), GTP (2.5 mM), CTP (2.5 mM), UTP (2.5 mM), th**G** (18.75 mM) and $MgCl_2$ (20 mM); template **3** and **4**: ATP (2 mM), GTP (2 mM), CTP (2 mM), UTP (2 mM), th**G** (15 mM) and $MgCl_2$ (20 mM); template **5**: ATP (1 mM), GTP (1 mM), CTP (1 mM), UTP (1 mM), th**G** (7.5 mM) and $MgCl_2$ (10 mM).

The reaction was incubated for 4 hours at 37 °C. The precipitated magnesium pyrophosphate was removed by centrifugation. Loading buffer (125 μL) was added after the reaction was concentrated to half its volume. The mixture was heated up for 3 minutes at 75 °C and loaded onto a preparative 20% denaturing polyacrylamide gel. The gel was UV shadowed; appropriate bands were excised, extracted overnight with ammonium acetate (0.5 M), and desalted on a Sep-Pak C18 column. Concentrations of the RNA transcripts

were determined using UV absorption spectroscopy at 260 nm, using the following molar extinction coefficients: C (ε_{260} = 7200 L·mol^{-1}cm^{-1}), U (ε_{260} = 9900 L·mol^{-1}cm^{-1}), G (ε_{260} = 11500 L·mol^{-1}cm^{-1}), A (ε_{260} = 15400 L·mol^{-1}cm^{-1}) and **thG** (ε_{260} = 5517 L·mol^{-1}cm^{-1}) The relative yield of each thG-initiated full-length transcript was determined to be 0.88 ± 0.01 (for transcript **2b**/**2a**), 1.16 ± 0.12 (transcript **3b**/**3a**), 0.82 ± 0.06 (transcript **4b**/**4a**), and 0.36 ± 0.06 (for transcript **5b**/**5a**).

2.2. 5' Phosphorylation

RNA oligonucleotides containing thG at their 5'-end (transcripts **2b**, **3b**, **4b**, **5b** and donor **S3**) were phosphorylated with T4 polynucleotide kinase. Large scale phosphorylation reactions containing thG-initialed RNA oligonucleotides (1-2 nmol, 6 μM) were performed in kinase buffer (1X New England Biolabs), additional DTT (5 mM), ATP (1 mM) and T4 polynucleotide kinase (0.2 U/ μL New England Biolabs). The reaction was incubated for 2 hours at 37 °C and precipitated with ammonium acetate (0.4 M), Glycoblue (100 μg/mL), and cold ethanol (2.5−3 v/v) in dry ice bath for 1 hour, followed by centrifugation (14,000 rpm) for 20 min and removal of the supernatant. The pellet was washed with cold ethanol (200 μL, 70%) and air-dried for 30 minutes before getting dissolved in gel loading buffer (50 μL). The mixture was heated for 3 minutes at 75 °C and loaded onto a preparative 20% denaturing polyacrylamide gel. The gel was UV shadowed; appropriate bands were excised, extracted overnight with ammonium acetate (0.5 M), and desalted on a Sep-Pak C18 column. Concentrations of the RNA transcripts were determined using UV absorption spectroscopy at 260 nm as described above.

2.3. Splinted ligation

DNA strands, complementing 10 to 15 nucleotides in the donor/acceptor were used as splint to create a short nicked duplex. Each 5'-phosphorylated donor strand (10 μM) was mixed with the corresponding acceptor strands (10 μM) and splint and annealed in Tris-HCl buffer (40 mM, pH 7.8) by heating up the solution for 3 minutes at 90 °C and cooling down slowly to room temperature. $MgCl_2$ (10 mM), DTT (10 mM), PEG 4000 (0.5 mM, 0.1 v/v of 50%) and T4 DNA ligase (0.7-1 U/μL, Fermentas) were then added to the mixture. The reaction was incubated at 37 °C for 2 hours before precipitated as described above. The pellet was washed with cold ethanol (200 μL, 70%) and air-dried for 30 minutes before getting dissolved in the gel loading buffer (50 μL). The RNA was resolved by gel electrophoresis and isolated and desalted as described above.

2.4. 5'-^{32}P-Labeling

Native transcript **S1** was first dephosphorylated with calf intestinal alkaline phosphatase (CIP). Transcript **S1** (26 pmol) was mixed with dephosphorylation buffer (6 μl, 10X) and CIP (2 μl) in a total volume of 60 μL and

incubated at 37 °C for 2 hours. Water (70 μL) was added and the reaction mixture was extracted with phenol:chloroform ($CHCl_3$):isoamyl alcohol (iAA) (100 μL, 25:24:1). The water layer was then extracted with chloroform (100 μL). The RNA in the aqueous layer was precipitated with ammonium acetate (20 μL, 10M), glycoblue (5 μL), EtOH (400 μL) and put in dry ice bath for 1 hour, followed by centrifugation (14,000 rpm) for 20 minutes and removal of the supernatant. The pellet was washed with cold ethanol (4 × 50 μL, 70%). The pellet was air-dried for 30 minutes and then dissolved in water (38 μL). Kinase buffer (5 μL, 10X), γ-^{32}P ATP (5 μL), of DTT (1 μL), and T4 polynucleotide kinase (1 μL), were added and the reaction was heated to 37 °C for 2 hours. The RNA was then precipitated with ammonium acetate (10 μL, 10 M), glycoblue (2 μL), ethanol (200 μL) and washed with cold ethanol (25 μl, 70%). The pellet was dissolved in loading buffer (1X TBE, 7M urea), and then the RNA was resolved by gel electrophoresis on a denaturing 20% polyacrylamide gel. The RNA was cut out and extracted with water overnight, filtered, and then concentrated using a speed vac.

3. RNA oligonucleotides characterization

MALDI spectra of the oligonucleotides were run either on Applied Biosystems Voyager-DETM Pro or Bruker biflex IV MALDI-TOFMS. The corresponding measuring matrices are described in the figure caption for each oligonucleotide. ESI of the oligonucleotides were run on ThermoFinnigan LCQ DECA XP for ESI.

3.1. Digestion with S1 nuclease

E5 (1.5 mmol) was incubated in S1 reaction buffer (1X, Promega) with S1 nuclease for 2 hours at 37 °C, followed with incubation in dephosphorylation buffer (Promega) with alkaline phosphatase for 2 hours at 37 °C. The ribonucleoside mixture obtained was analyzed by reverse-phase analytical HPLC. HPLC analysis was carried out with an Agilent 1200 series system with an Eclipse XDB-C18 5μm, 4.5 × 150 mm column. 0.1% formic acid stock solutions were prepared by dissolving 1 ml of formic acid (Acros, 99%) in 999 ml HPLC grade acetonitrile (Sigma) or MilliQ water and filtered using Millipore type GNWP 0.2 μm filters before use. The injection was subjected to a gradient (12 minutes, from 0 to 6% acetonitrile 0.1% formic acid in water 0.1% formic acid). A flow rate of 1 ml / minute was used and the run was carried out at 25.00 ± 0.10 °C. Each run was monitored at 260 and 321 nm with calibrated references at 650 nm and slit set at 1 nm.

3.2. Digestion with T1 nuclease

Each ^{32}P-labeled **E1**, **E2**, **E3**, **E4**, **E5**, **S2** and **S3** was digested with T1 nuclease. Each RNA (2 μL, about 60,000 cpm) was first denatured in T1 denaturing buffer (pH 5) containing of sodium citrate (20 mM), urea (7M) and 1 mM EDTA (1 mM) by heating up the solution at 55 °C for 3 minutes. Yeast tRNA (2 μg) and T1 nuclease (1 U) were added after the mixture was cooled down to room temperature, resulting in a total volume of 10 μL.

The reaction mixture was incubated at room temperature for 20 minutes before precipitated and washed with ethanol as described above. The pellet was dissolved in stop buffer (8 μL) containing formamidine (95%), EDTA (20 mM), bromphenol blue (0.05%) and xylene cyanol (0.05%) and then loaded on 15% polyacrylamide with urea (7M) gel. Corresponding bands were imaged with Personal Molecular Imager.

3.3. Alkaline hydrolysis

Each ^{32}P-labeled RNA (2uL, about 60,000 cpm) was incubated in alkaline hydrolysis buffer (pH 9.2) containing sodium carbonate (50 mM) and EDTA (1 mM) in a total volume of 15 μL at 90 °C for 20 minutes before precipitated and washed with ethanol as described above. The pellet was dissolved in stop buffer (8 μL) containing formamide (97%), EDTA (20 mM), bromphenol blue (0.05%) and xylene cyanol (0.05%) and then loaded on 15% polyacrylamide with urea (7M) gel. Corresponding bands were imaged with Personal Molecular Imager.

4. Ribozyme cleavage reaction conditions

Cleavage reactions were conducted in a total reaction volume of 34 μL for **E1**, **E4**, **E5** and 22 μL for **E2**, **E3** with **S1**, and 34 μL for **E6** with **S2** or **S3** for radiography. For the fluorescence-based experiments, a total volume of 125 μL was used. The reactions were carried out at 31 °C in a buffer containing Tris-HCl (50 mM, pH 7.0) and NaCl (200 mM). Buffered solutions of the substrate **S3** (0.6 μM) and enzyme **E6** (6 μM) were denatured separately by heating to 90 °C for 3 minutes and cooled down to room temperature over 10 minutes to allow for refolding. $MgCl_2$ was added to both the enzyme and substrate to make a final concentration of 10 mM, and both were equilibrated at 31 °C for 10 minutes. The cleavage reaction was then initiated by manually mixing equal volumes of the modified or natural substrate (0.6 μM) with the enzyme or modified enzyme (6 μM) in a heat block at 31 °C, to give final concentrations of 0.3 μM of the substrate and 3 μM of the enzyme and 10 mM $MgCl_2$.

4.1. Ribozyme cleavage assay

For initial data points (time = 0), an aliquot of the substrate **S1**, **S2** or **S3** (2 μL) was removed immediately prior to starting the reaction. Following initiation of the reaction, aliquots (4 μL) were removed at designated time periods and quenched with urea containing loading buffer (15 μL, 7 M urea, 1× TBE buffer, 0.05% bromophenol blue, and 0.05% xylene cyanol). Each tube was heated up to 90 °C for 1.5 minutes and loaded on a 20% polyacrylamide with urea (7M) gel. Corresponding bands were quantified on a Personal Molecular Imager and analyzed with Quantity One software (Biorad).

4.2. Ribozyme cleavage data analysis

Enzymatic process rate constants (k_2) were calculated as the slope of the semi-logarithmic plot of ln (1 − S/S_0) versus time, where S/S_0 is the fraction of cleaved substrate. For experiments utilizing a radioactively labeled substrate, S/S_0 was determined by dividing the amount of cleaved substrate by the sum of the full length and cleaved substrates.

5. Absorption and emission spectroscopy assays and data

Absorption spectra were measured on a Shimadzu UV-2450 spectrophotometer setting the slit at 1 nm and using a resolution of 0.5 nm. All the spectra were corrected for the blank. Steady state emission spectra were measured on a Horiba Fluoromax-4 equipped with a cuvette holder with a stirring system setting both the excitation and the emission slits at 3 nm, the resolution at 1 nm and the integration time 0.1 s if not otherwise described. The probe was excited just before the absorbance maxima to provide enough energy to excite all the fluorophore population, and maximize the recorded emission signal. The equation below was used to calculate the excitation wavelength for the steady state fluorescence measurements.

$$\lambda_{exc}(nm) = \lambda_{abs}(nm) - 5\ (nm)$$

All the spectra were corrected for the blank. Both instruments were equipped with a thermostat controlled ethylene glycol-water bath fitted to specially designed cuvette holder and the temperature was kept at 25.00 ± 0.10 °C if not otherwise described. **thG** was dissolved in DMSO (3.47mM) to prepare highly concentrated stock solution.

5.1 pK_a determination

Aqueous stock solutions (100 ml) were prepared by mixing aqueous sodium phosphate monobasic (0.5 M), aqueous sodium phosphate dibasic (0.5 M) and aqueous sodium chloride (2 M) to have a final concentration of 100 mM NaCl and 10 mM phosphate ions. The pH of each solution was adjusted to the desire value by adding aliquots of 2 M aqueous HCl or 2 M aqueous NaOH prior to spectral measurements. **thG** was dissolved in DMSO (3.5 mM) to prepare highly concentrated stock solution.

In a typical experiment, aliquots (10 μl) of the concentrated DMSO solution were diluted with air-saturated solvents (3 ml). The solution was mixed with a pipette for 10 seconds and placed in the cuvette holder at 25.00 ± 0.10 °C for 3 minutes before spectra were recorded. All sample contain 0.3 v/v % of DMSO.

The absorption (λ_{abs}) and the emission (λ_{em}) maxima were plotted versus the pH and fitted using a Boltzmann sigmoidal curve using Kaleidagraph 3.5. The pK_a values were determined by interpolation of the fitting curves. The reported pK_a values represent the average of three independent sets of measurements. The values and

the relative standard deviations are reported in table S1.

5.2. Ribozyme cleavage emission assay

The ribozyme cleavage experiment was carried out on a total volume of 125 μL in a 125 μL cuvette. The reactions were carried out at 31 °C in a buffer containing Tris-HCl (50 mM, pH 7.0) and NaCl (200 mM). Buffered solutions of the substrate **S3** (0.6 μM) and enzyme **E6** (6 μM) were denatured separately by heating to 90 °C for 3 minutes and cooled down to room temperature over 10 minutes to allow for refolding. $MgCl_2$ was added to both the enzyme and substrate to make a final concentration of 10 mM, and both were equilibrated at 31 °C for 10 minutes. The cleavage reaction was then initiated by manually mixing equal volumes of the modified or natural substrate (0.6 μM) with the enzyme or modified enzyme (6 μM) in the cuvette 31 ± 0.10 °C, to give final concentrations of 0.3 μM of the substrate and 3 μM of the enzyme and 10 mM $MgCl_2$. The reaction was followed by recording emission spectra (430 –480 nm) upon excitation at 360 nm over time. The excitation and the emission slits were set at 10 nm, the resolution at 1 nm and the integration time 0.1 s.

6. Supplementary figures

6.1. Polyacrylamide gels

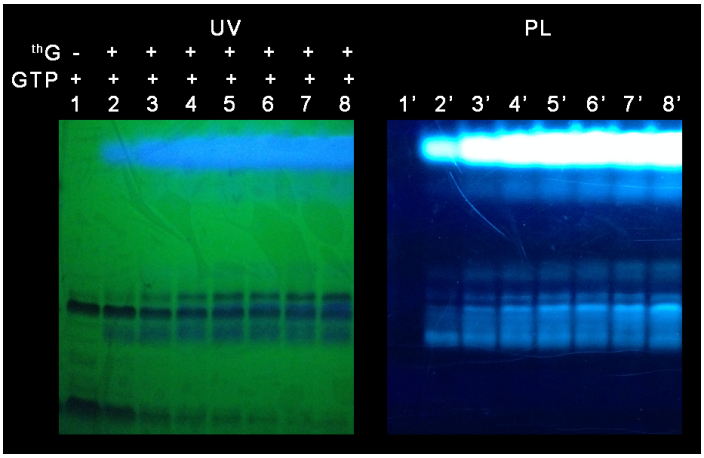

Figure S1 Transcription reactions of template **1** with various thG/GTP ratios. Lane 1, 1 mM of each NTP (ATP, GTP, CTP, UTP); lane 2, 1 mM of each NTP, 1 mM thG; lane 3, 1 mM each NTP, 3 mM thG; lane 4, 1 mM each NTP, 5 mM thG; lane 5, 1 mM each NTP, 7 mM thG; lane 6, 1 mM each NTP, 9 mM thG; lane 7, 1 mM each NTP, 11 mM thG; lane 8, 1 mM each NTP, 13 mM thG.

T7 Promoter 5′-TAA TAC GAC TCA CTA TAG-3′

Template **2** 3′-ATT ATG CTG AGT GAT ATC TCC GGC TTT CCG GCT TTG CAA GCG-5′

Template **3** 3′-ATT ATG CTG AGT GAT ATC CGG CTT TCC GGC TTT GCA AGC G-5′

Template **4** 3′-ATT ATG CTG AGT GAT ATC CGG CTT TGC AAG CG-5′

Template **5** 3′-ATT ATG CTG AGT GAT ATC TTT GCA AGC G-5′

T7 RNA Pol | thG, NTPs

Transcripts

2a X = pppG 5′-XAG GCC GAA AGG CCG AAA CGU UCG C-3′
2b X= thG
3a X = pppG 5′-XGC CGA AAG GCC GAA ACG UUC GC-3′
3b X= thG
4a X = pppG **4b** X= thG 5′-XGC CGA AAC GUU CGC-3′
5a X = pppG **5b** X= thG 5′-XAA ACG UUC GC-3′

Figure S2. Transcription reaction of different 5'-thG terminated oligonucleotides (**2b**, **3b**, **4b** and **5b**) starting from template **2**–**5**.

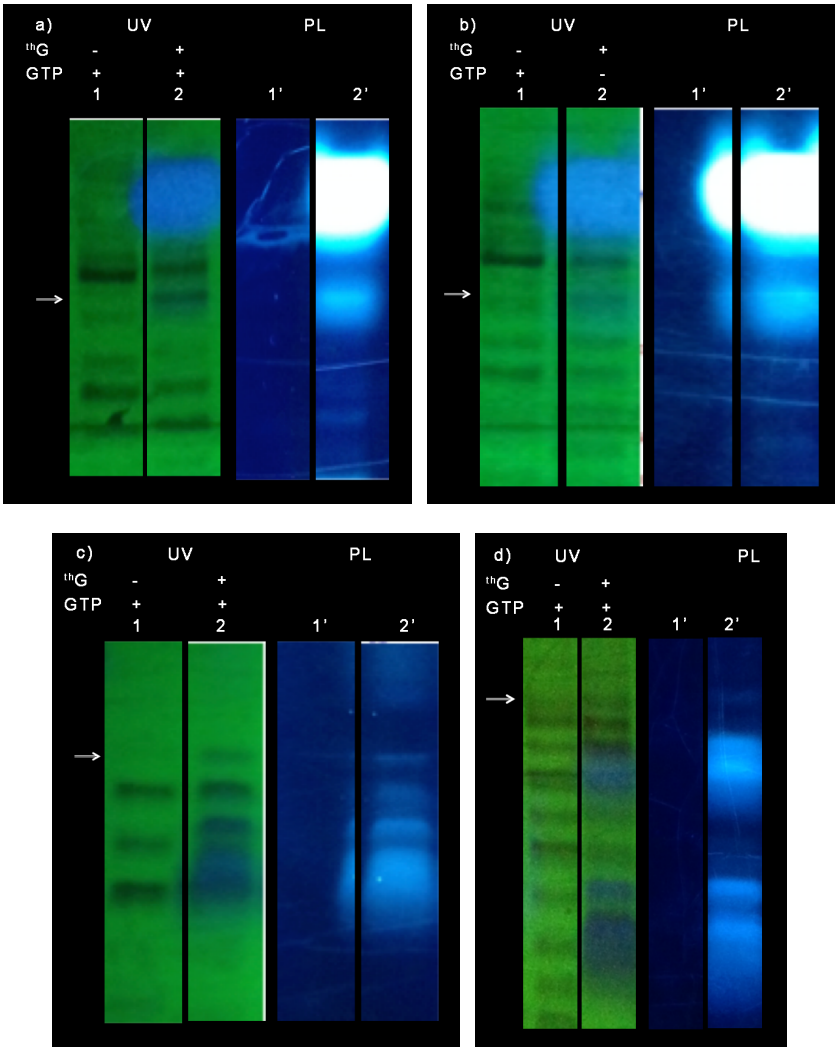

Figure S3. Transcription reactions of HH Ribozyme donor strands. The white arrows indicate positions of full-length 5'-thG transcripts. a) transcription reactions of template **3**; b) transcription reactions of template **2**; c) transcription reactions of template **5**; d) transcription reactions of template **4**.

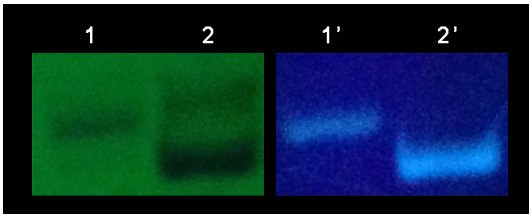

Figure S4. Phosphorylation of transcript **3b**. Lane 1, non-treated transcript 3b; lane 2, phosphorylation of transcript **3b**.

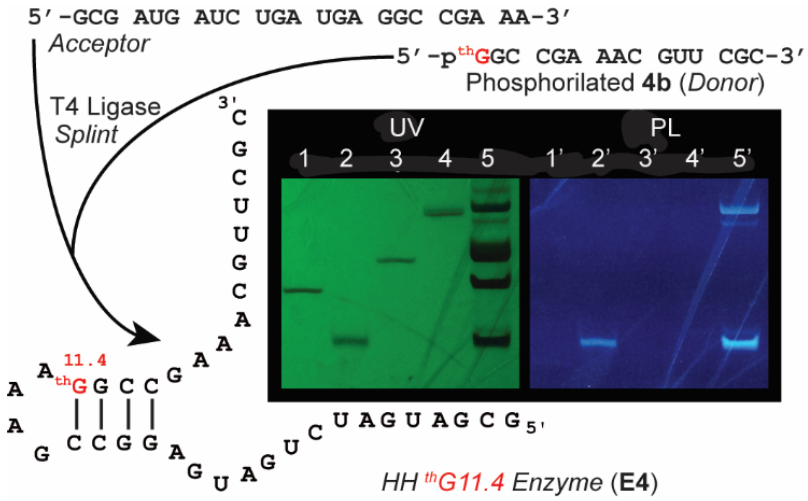

Figure S5. Splinted ligation of HH16 thG11.4 E donor (**4b**) and acceptor performed with T4 DNA ligase. HH16 thG11.4 E acceptor (lane 1 and 1'), phosphorylated **4b** (lane 2 and 2'), splint (lane 3 and 3'), native HH16 E (lane 4 and 4') and ligation reaction (lane 5 and 5').

Acceptor	5′-G CGA UGA UCU GAU-3′
Donor (phosphorylated **2b**)	5′-pthGAG GCC GAA AGG CCG AAA CGU UCG C-3′
Splint	3′-C GCT ACT AGA CTA CTC CGG CTT TCC GGC TTT GC-5′
E2 (24%)	5′-G CGA UGA UCU GAU thGAG GCC GAA AGG CCG AAA CGU UCG C-3′
Acceptor	5′-GCG AUG AUC UGA UGA-3′
Donor (phosphorylated **3b**)	5′-pthGGC CGA AAG GCC GAA ACG UUC GC-3′
Splint	3′-GC TAC TAG ACT ACT CCG GCT TTC CGG CTT TGC-5′
E3 (35%)	5′-GCG AUG AUC UGA UGA thGGC CGA AAG GCC GAA ACG UUC GC-3′
Acceptor	5′-GC GAU GAU CUG AUG AGG CCG AAA-3′
Donor (phosphorylated **4b**)	5′-pthGGC CGA AAC GUU CGC-3′
Splint	3′-CTA GAC TAC TCC GGC TTT CCG GCT TTG CAA GCG-5′
E4 (18%)	5′-GC GAU GAU CUG AUG AGG CCG AAA thGGC CGA AAC GUU CGC-3′
Acceptor	5′-GCG AUG AUC UGA UGA GGC CGA AAG GCC-3′
Donor (phosphorylated **5b**)	5′-pthG AAA CGU UCG C-3′
Splint	3′-C TAG ACT ACT CCG GCT TTC CGG C TTT GCA AGC G-5′
E5 (40%)	5′-GCG AUG AUC UGA UGA GGC CGA AAG GCC thGAA ACG UUC GC-3′

Figure S6. Design of splinted ligation of **E2**, **E3**, **E4** and **E5**.

6.2. MALDI and ESI Spectra

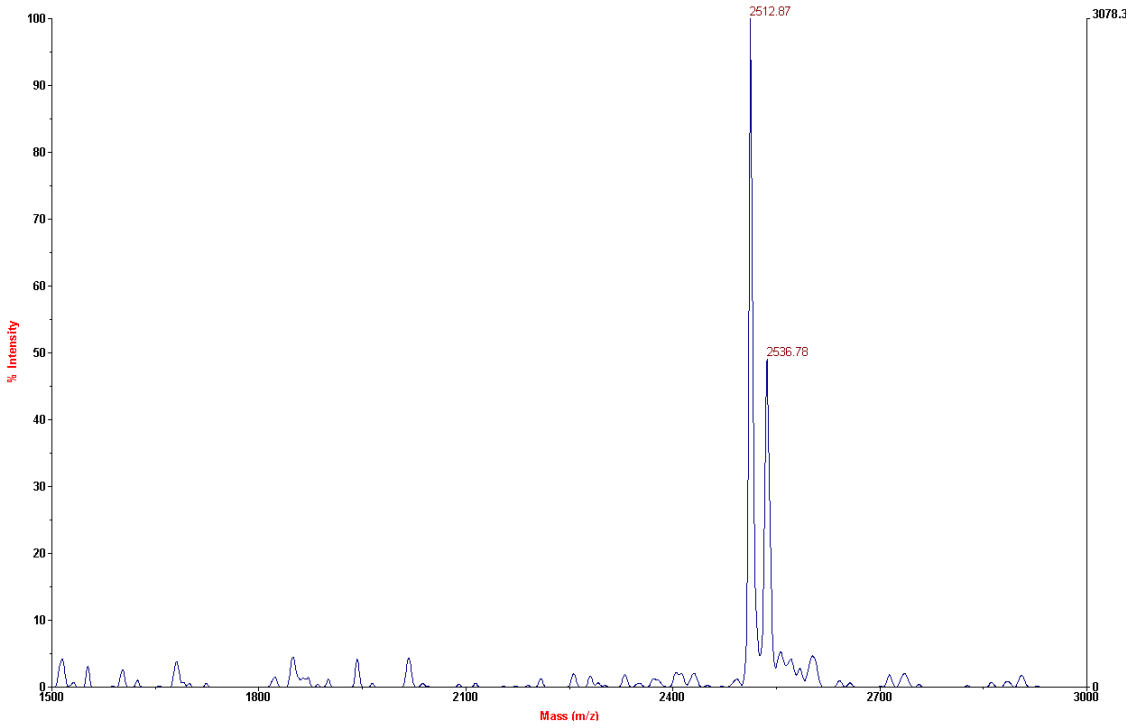

Figure S7. MALDI-TOF mass spectrum of donor strand of **S3**. Expected molar mass: 2517.56 Da. Taken by MALDI-TOF mass spectrometer in megetive ion mode with Voyager™ Biospectrometry™ Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

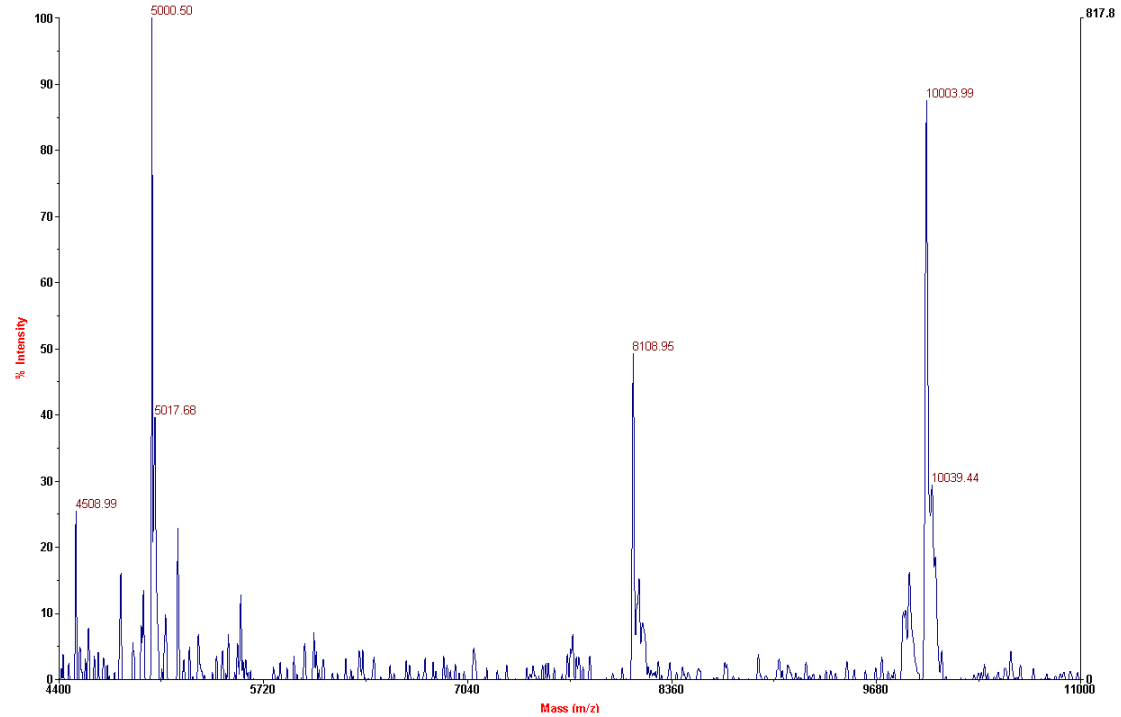

Figure S8. MALDI-TOF mass spectrum of transcript **2b** with standard. Expected molar mass of transcript **2b** was 8113.95 Da, and expected molar mass of the standard oligonucleotide was 10007.90 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager™ Biospectrometry™ Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

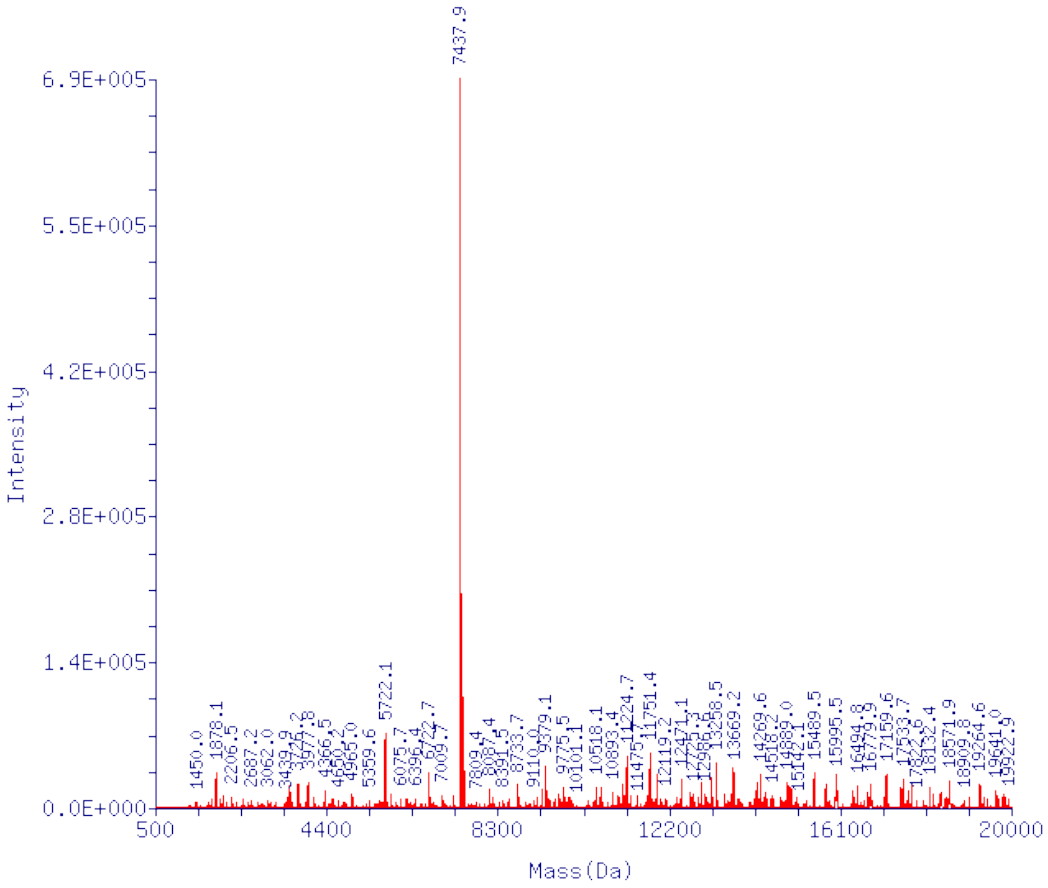

Figure S9. ESI mass spectrum in negative ion mode transcript **3b** with Xcalibur software version 1.3, and the raw ESI-MS m/z data were deconvoluted by ProMass for Xcalibur Version 2.5 SR-1. The running buffer was 10 mM *tert*-butylamine in 70% acetonitrile in water.

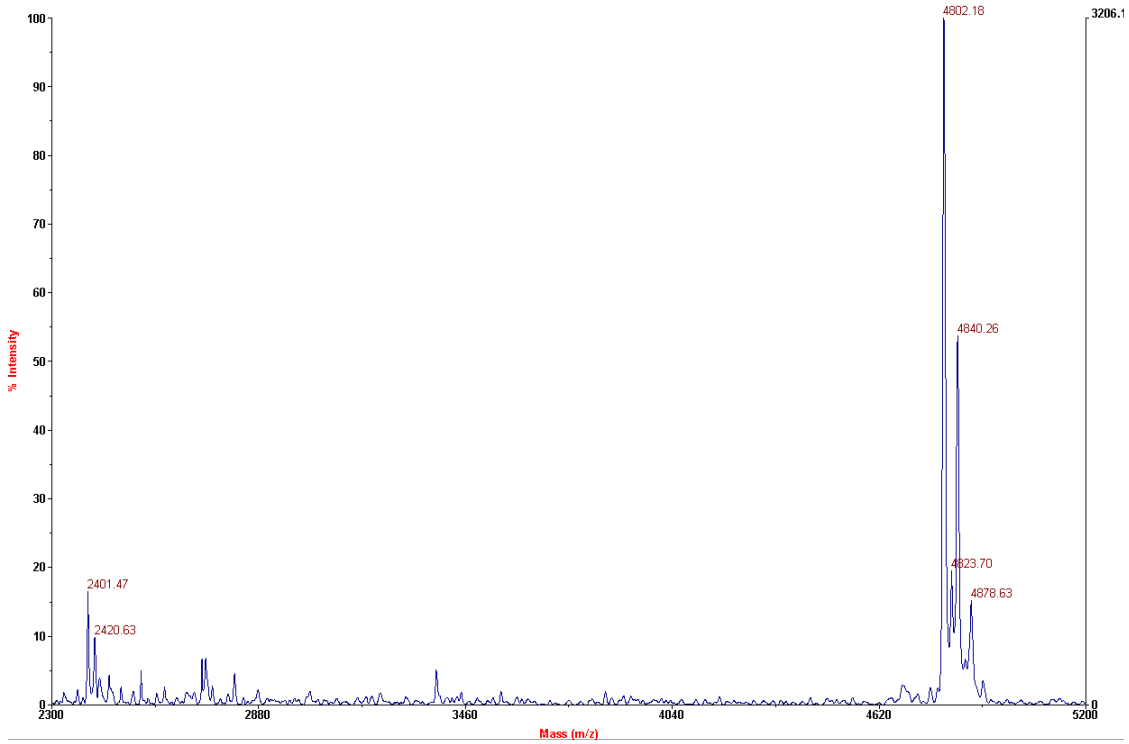

Figure S10. MALDI-TOF mass spectrum of transcript **4b**. Expected molar mass: 4805.95 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager™ Biospectrometry™ Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate (dibasic) in 50% ACN.

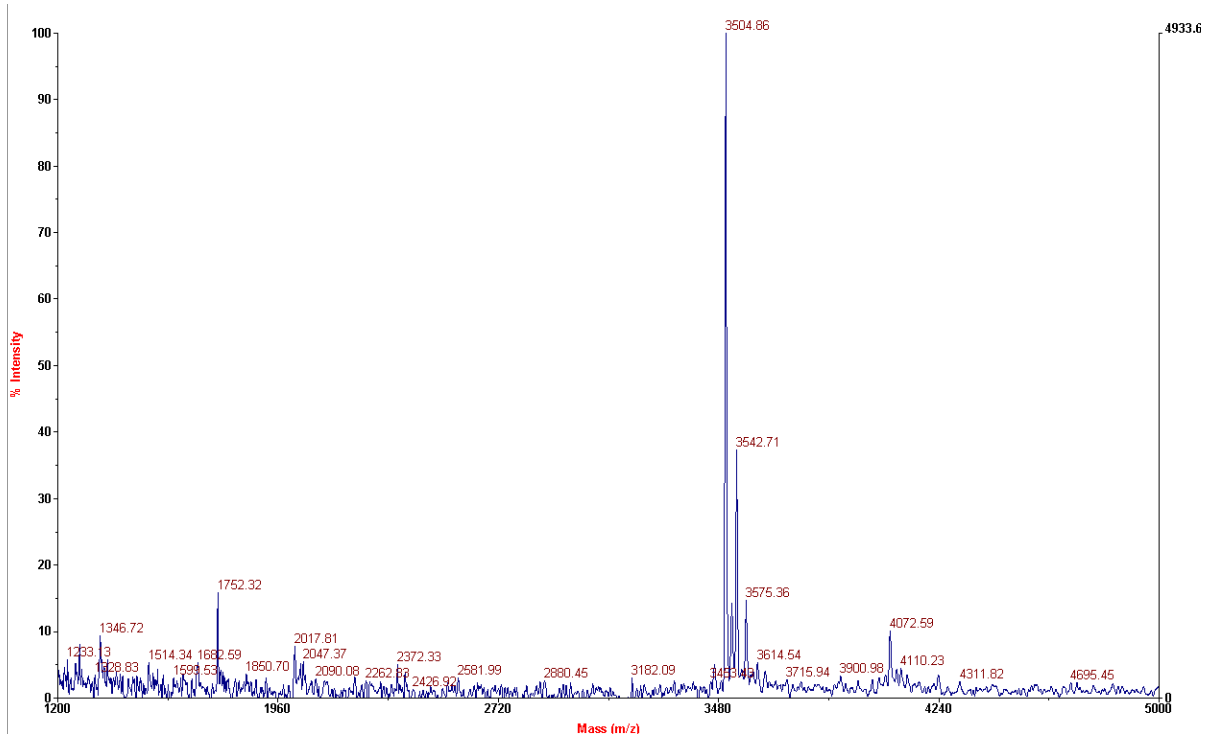

Figure S11. MALDI-TOF mass spectrum of transcript **5b**. Expected molar mass: 3505.19 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager™ Biospectrometry™ Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

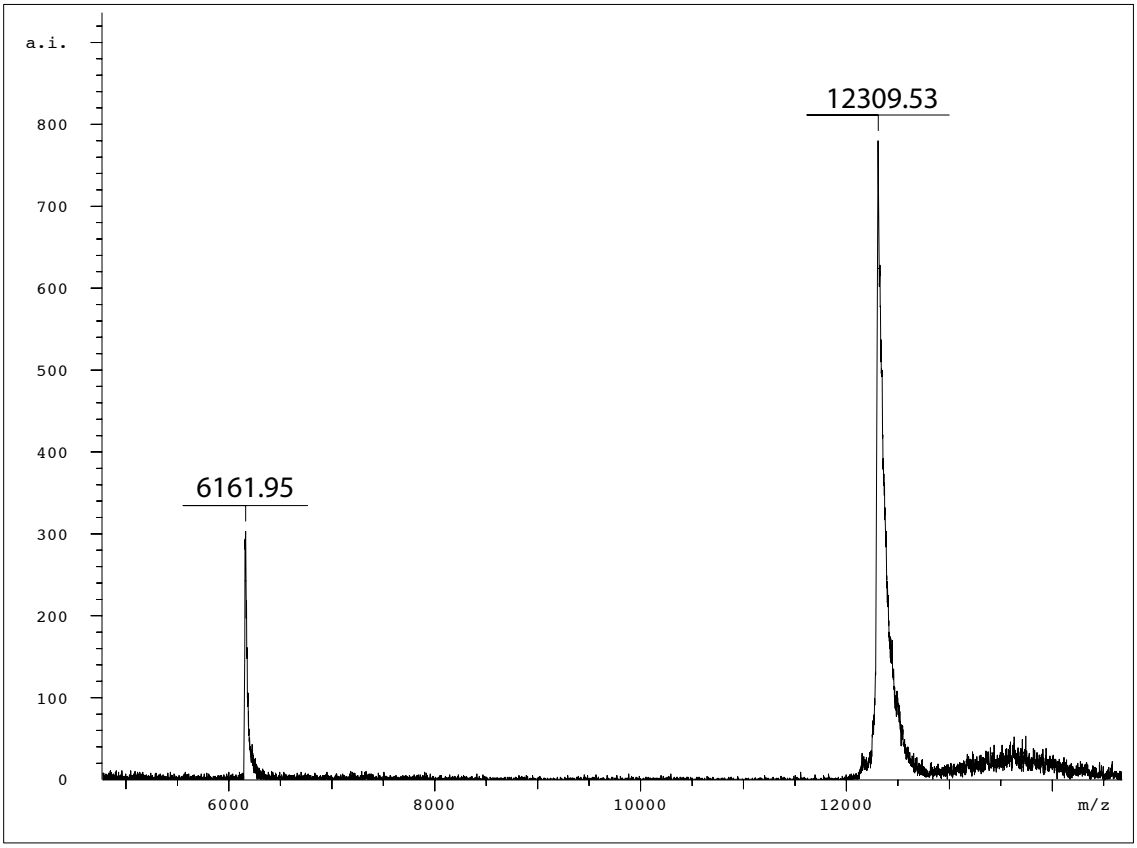

Figure S12. MALDI-TOF mass spectrum of **E2**. Expected molar mass: 12317.38 Da. Taken by Bruker MALDI-TOF mass spectrometer in negative ion mode with FLEX Control software. The raw mass spectra data was analyzed with X-TOF. The matrix was 45.4 g/L 3-Hydroxypicolinic acid (3-HPA) and 10 g/L diammonium hydrogencitrate dibasic in 50% ACN.

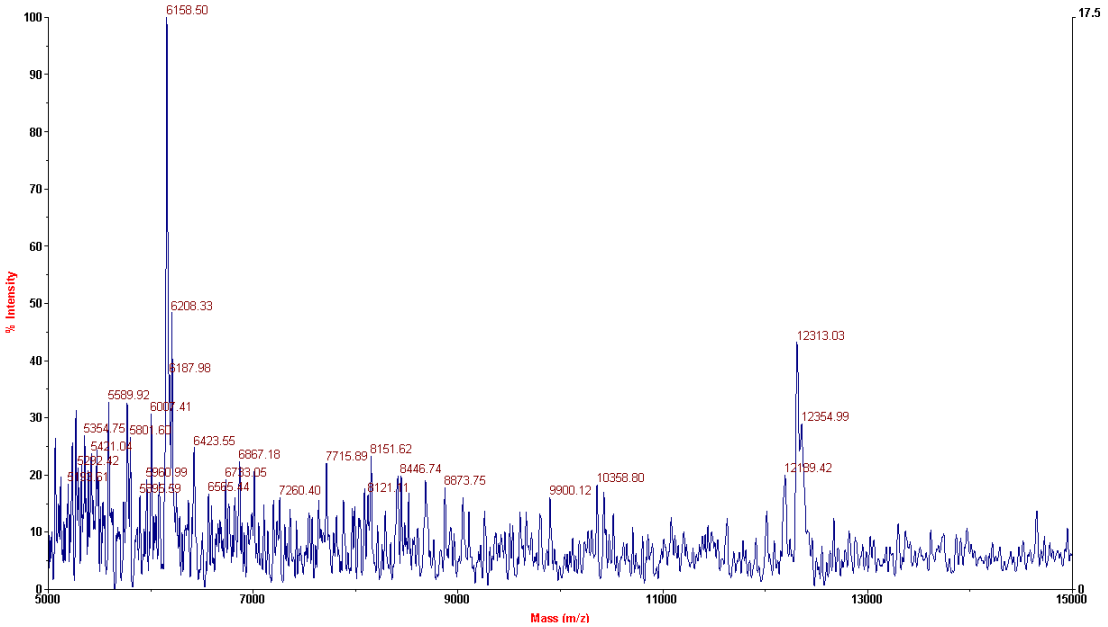

Figure S13. MALDI-TOF mass spectrum of **E3**. Expected molar mass: 12317.38 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager^{TM} Biospectrometry^{TM} Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

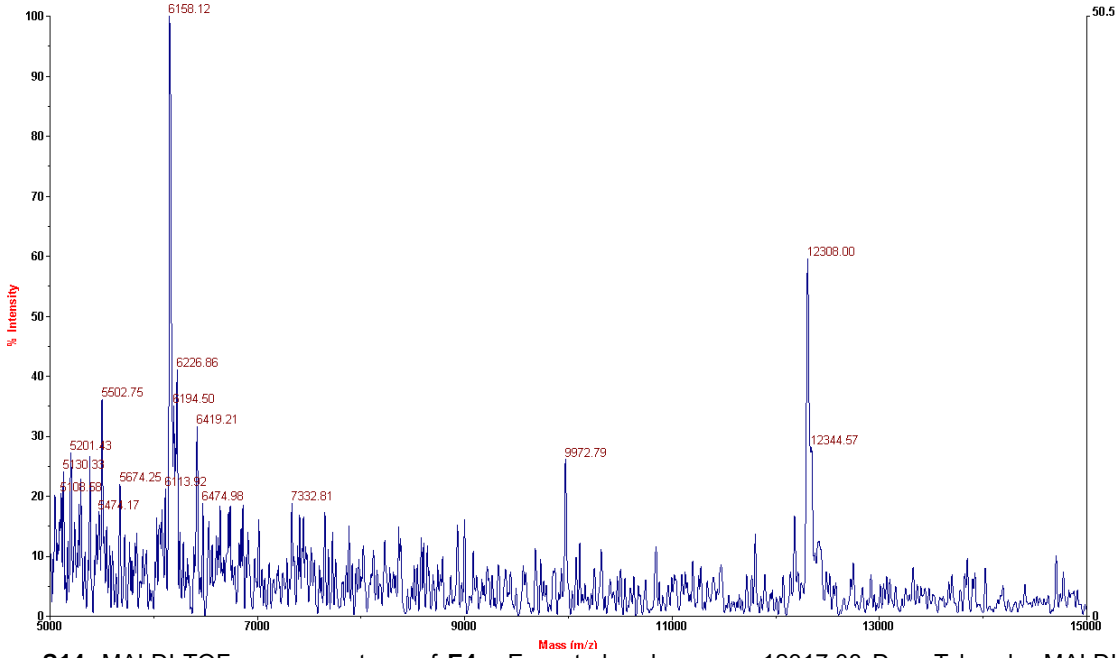

Figure S14. MALDI-TOF mass spectrum of **E4**. Expected molar mass: 12317.38 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager^{TM} Biospectrometry^{TM} Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

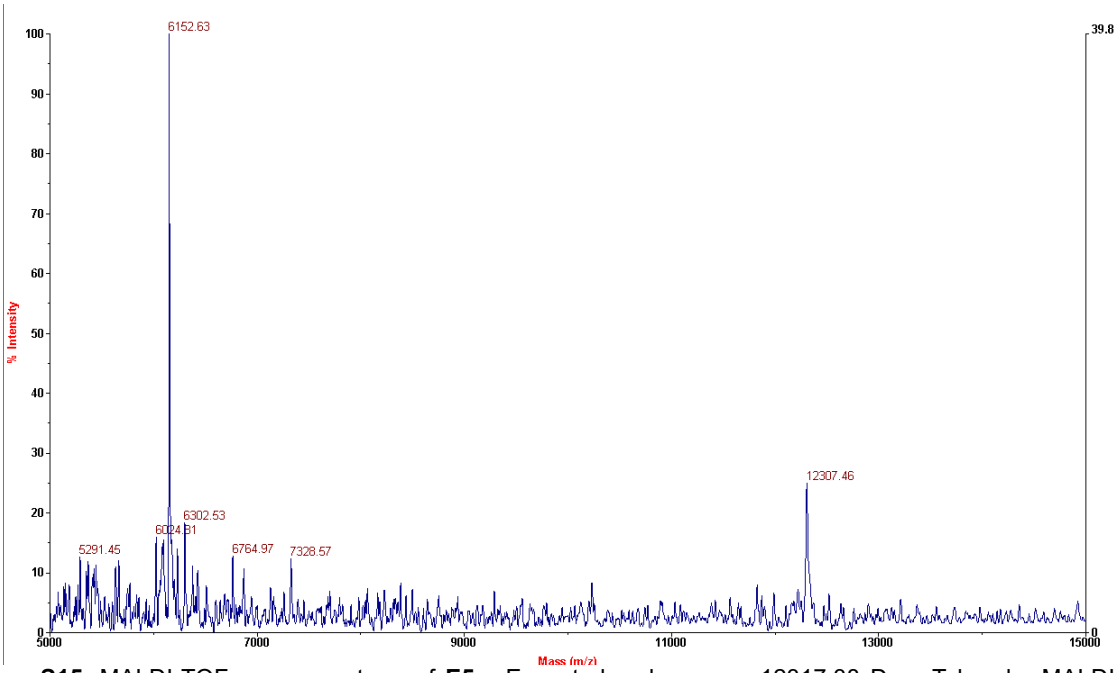

Figure S15. MALDI-TOF mass spectrum of **E5**. Expected molar mass: 12317.38 Da. Taken by MALDI-TOF mass spectrometer in negative ion mode with Voyager™ Biospectrometry™ Workstation software. The raw mass spectra data was analyzed with Data Explorer version 4.0.0.0. The matrix was 20 g/L 2,4,6-trihydroxyacetophenone (THAP) and 4 g/L diammonium hydrogencitrate dibasic in 50% ACN.

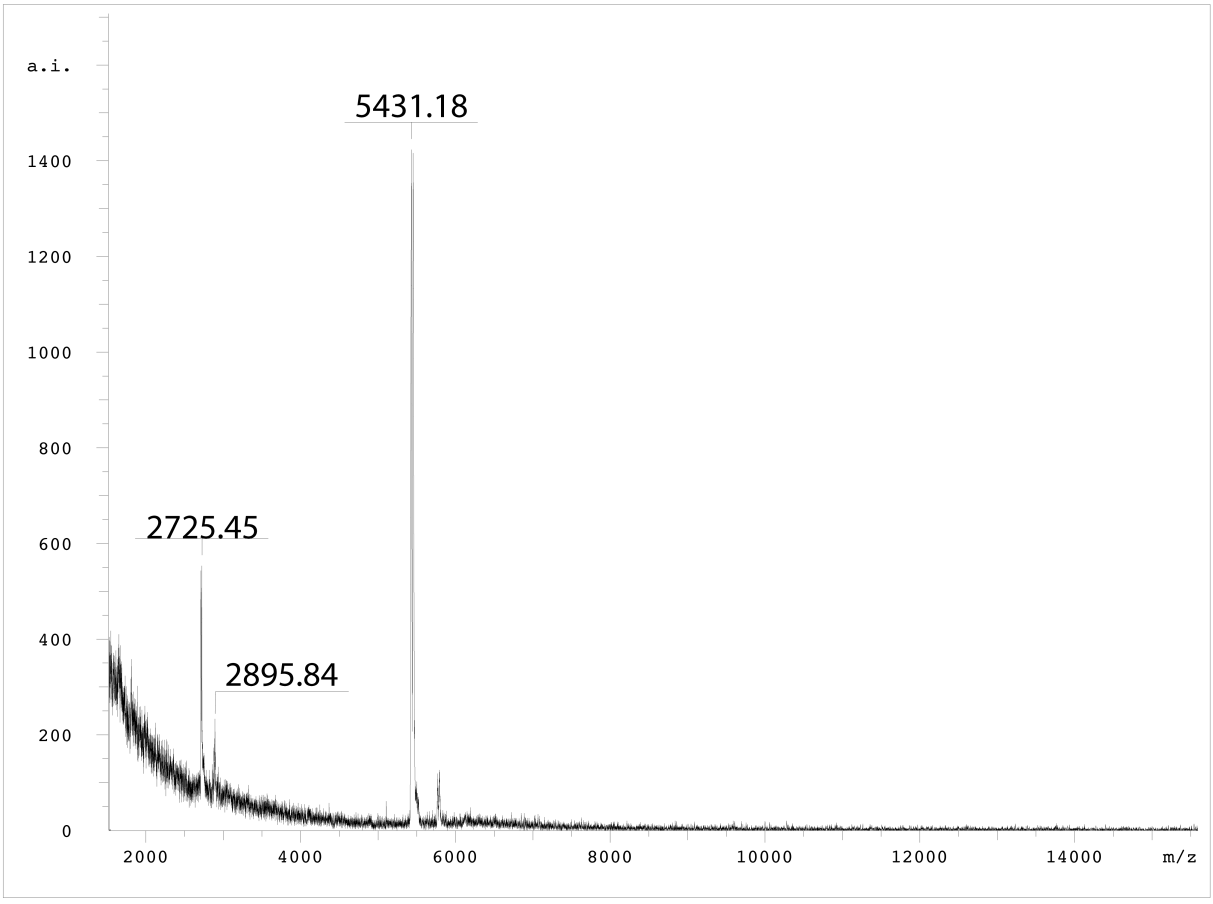

Figure S16. MALDI-TOF mass spectrum of **S3**. Expected molar mass: 5433.28 Da. Taken on a Bruker MALDI-TOF mass spectrometer in negative ion mode with FLEX Control software. The raw mass spectra data was analyzed with X-TOF. The matrix was 45.4 g/L 3-Hydroxypicolinic acid (3-HPA) and 10 g/L diammonium hydrogencitrate dibasic in 50% ACN.

6.3. ^{32}P-Postlabeling/polyacrylamide gel

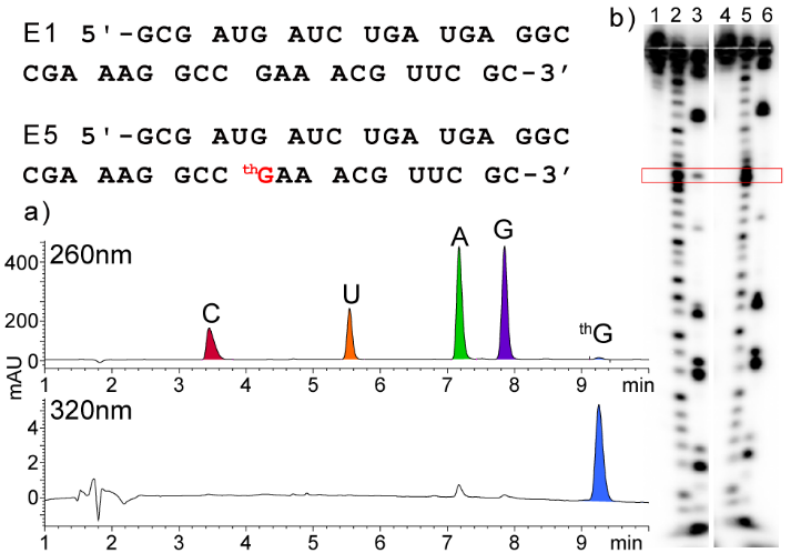

Figure S17. Characterization of **E5**. a) HPLC traces of **S1** digestion of **E5**. b) T1 sequencing of **E1** and **E5**. Lane 1, non-reacted E1. Lane 2, alkaline hydrolysis of **E1**. Lane 3, T1 digestion of **E1**. Lane 4, non-reacted **E5**. Lane 5, alkaline hydrolysis of **E5**. Lane 6, T1 digestion of **E1**. Position 12 is red boxed.

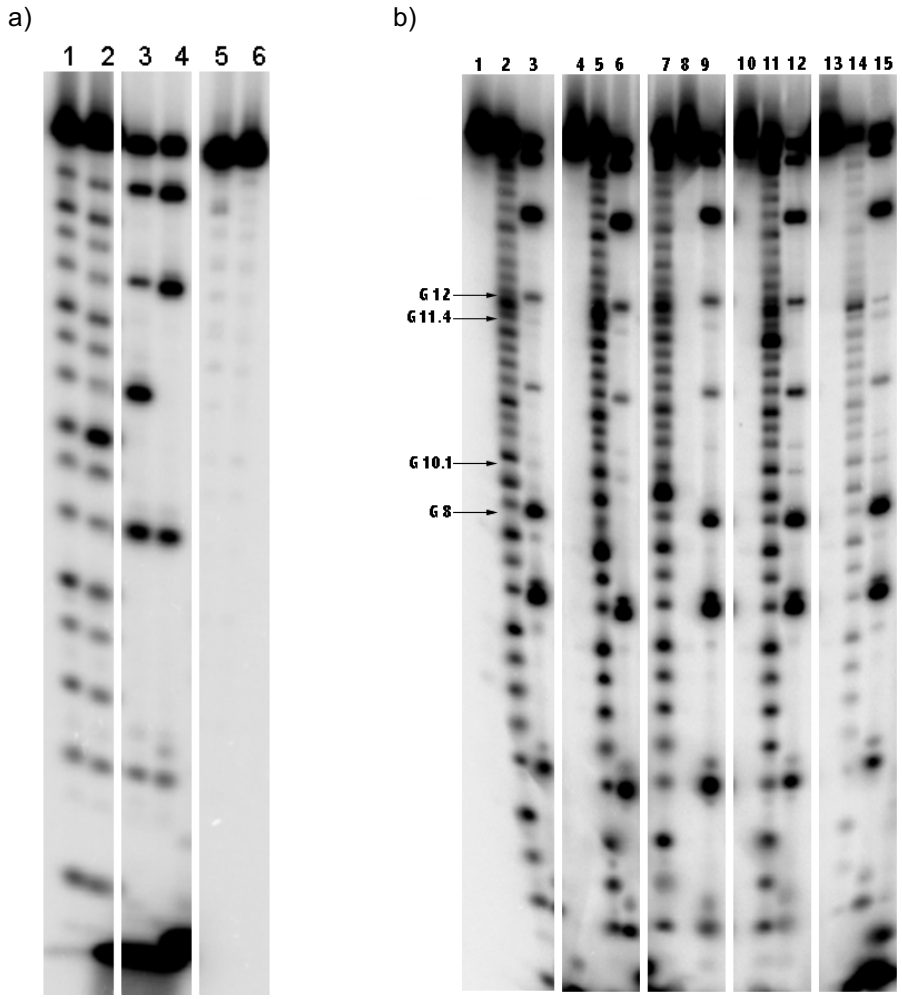

Figure S18. Alkaline hydrolysis and T1 nuclease digestion of thG-modified RNA oligonucleotides. a) Lane 1 and 2, alkaline hydrolysis of substrates **S2** and **S3**; lane 3 and 4, T1 nuclease digestion of **S2** and **S3**; lane 5 and 6, non-treated **S2** and **S3**. b) Lane 1, non-treated enzyme strand **E1**; lane 2, alkaline hydrolysis of **E1**; lane 3, T1 nuclease digestion of **E1**; **E4**, non-treated enzyme strand **E2**; lane 5, alkaline hydrolysis of **E2**; lane 6, T1 nuclease digestion of **E2**; lane 7, alkaline hydrolysis of enzyme strand **E3**; lane 8, non-treated **E3**; lane 9, T1 nuclease digestion of **E3**; lane 10, non-treated enzyme strand **E4**; lane 11, alkaline hydrolysis of **E4**; lane 12, T1 nuclease digestion of **E4**; lane 13, non-treated enzyme strand **E5**; lane 14, alkaline hydrolysis of **E5**; lane 15, T1 nuclease digestion of **E5**. Position of G8, G10.1, G11.4 and G12 were indicated with arrows. T1 digestion at guanosine residues of helix II of the enzyme strands were not observed even for the native enzyme strand **E1**, which was most likely due to the formation of the duplex structure of helix II.

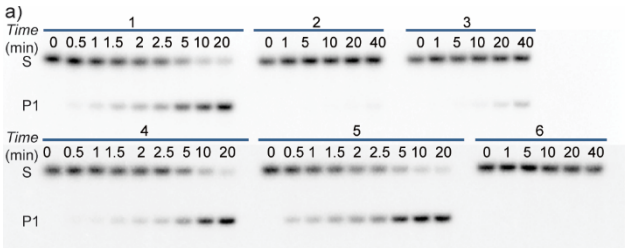

Figure S19. Cleavage of ^{32}P-labeled S1 by of HH enzyme strands with replacement of G for thG at different positions. E5 (1), E2 (2), E3 (3), E1 (4), E4 (5) and fully thG-modified enzyme, thG-E (6) with native substrate (S1). S1 and P1 indicate the substrate and the product strands respectively.

6.4 Fluorescence spectra

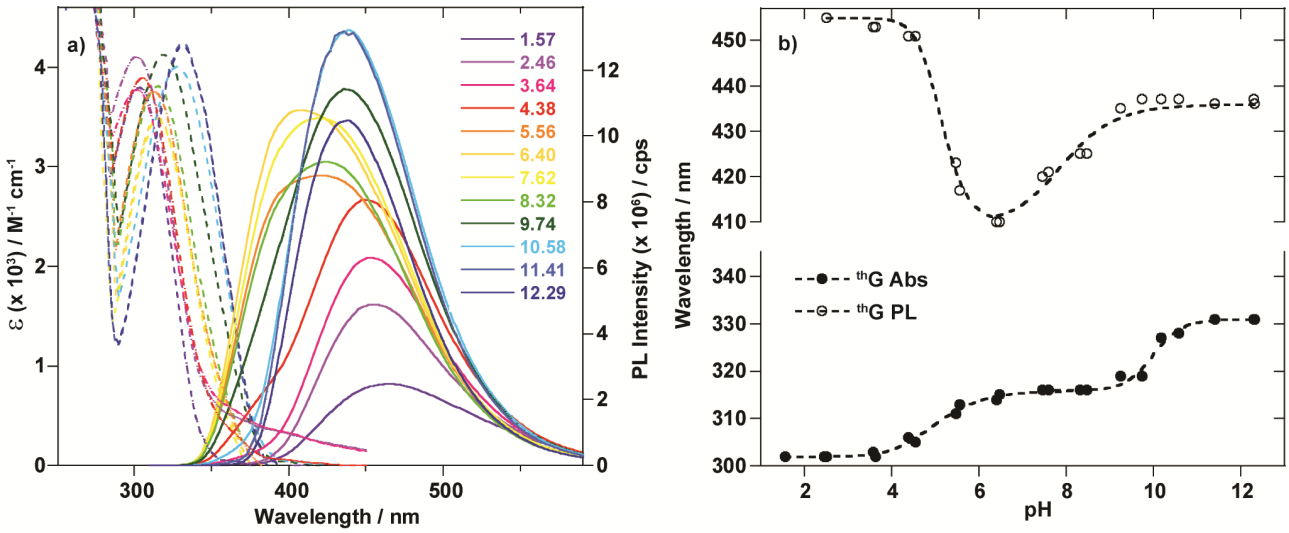

Figure S20. a) Absorption (dashed lines) and emission (solid lines) traces in buffer solutions at different pH for th**G**. The emission spectra were normalized to 0.1 intensity at the excitation wavelength. b) Absorption (black) and emission (white) maxima variation versus pH for th**G**.

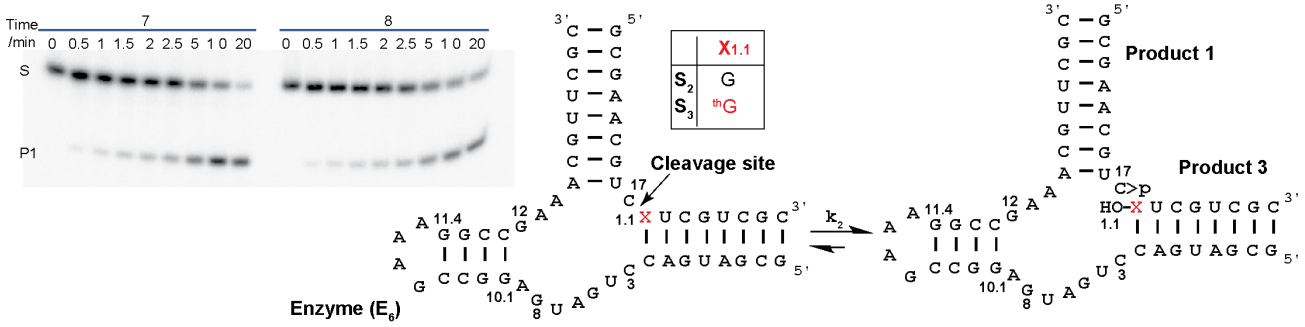

Figure S21. Hammerhead ribozymes and cleavage reactions of G1.1-S (**S2**), thG1.1-S (**S3**) and corresponding enzyme (**E6**). Hammerhead ribozyme cleavage reactions, **S2-E6** (7) and **S3-E6** (8), were followed by ^{32}P radioactive labeling of substrate strands. **S** and **P1** indicate the substrates and the product strands respectively.

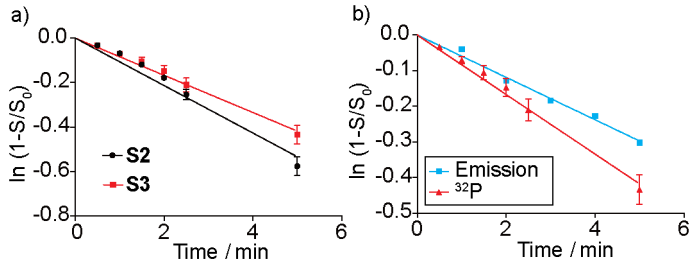

Figure S22. a) Initial kinetics of E6 with the native S2 (black) and the thG-modified substrate (red). The pseudo-first-order rate constants (k_2) of the cleavage reactions are determined as the slope of semi-logarithmic plot of the fraction cleaved as function of time. b) Initial kinetics of **E6** with **S3** monitored by radio-active assay (red) and fluorescence spectroscopy (cyan).

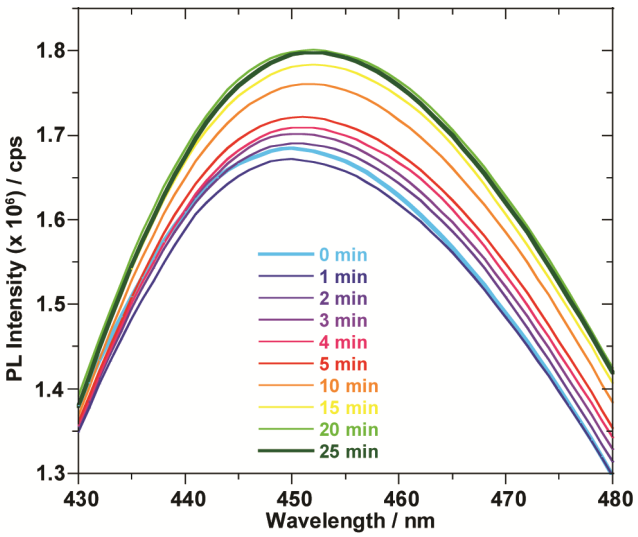

Figure S23. Emission spectra over time for the enzymatic cleavage of substrate S3 and enzyme **E6**.

6.5 Schematic mechanism

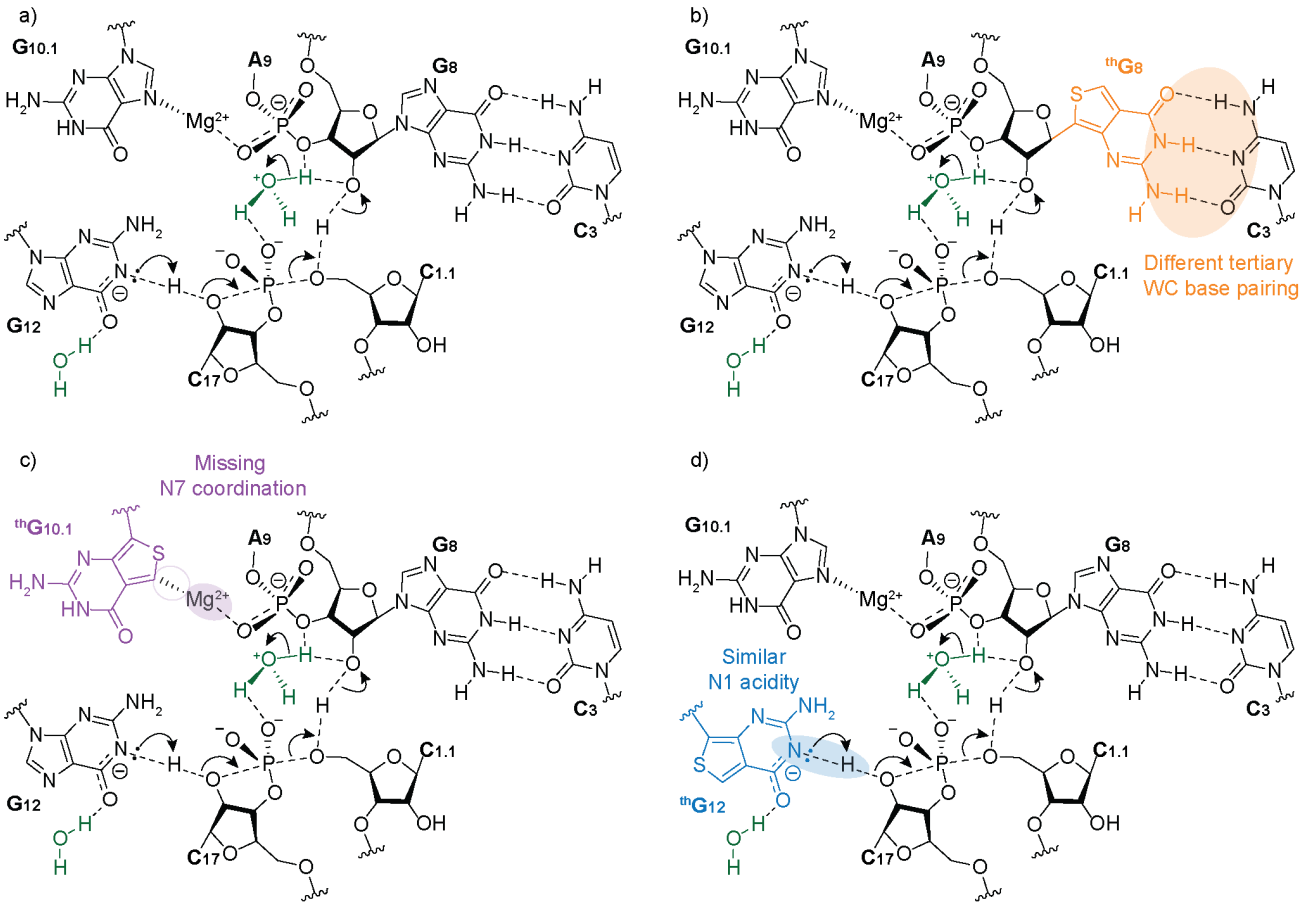

Figure S24. Schematic representation of the enzymatic cycle for the native **E1** (a), **E2** (b), **E3** (c) and **E4** (d) HH enzymes.

7. Supplementary Tables

Table S1. thG pK_a values

	pK_a	
Absorbance	5.06 ± 0.06 (N3)	10.08 ± 0.08 (N1)
Emission	5.17 ±0.05 (N3)	7.95 ±0.09 (N1)

Table S2. Initiator Nucleotides and Nucleosides

Description	Initiator Molecule	Modification site of molecule	Oligo modification	Follow up reaction
AMP and AMP analogue	5'- or N6- modified AMP derivatives [S2]	5' phosphate or N6	5'	
	cyanine-AMP conjugates [S3]	5' phosphate	5'	
	thio-AMP dimer [S4]	5' phosphate	5'	Biotinylation of 5' HS-RNA
GMP and GMP analogue	GMP [S5]	5'	5'	
	Norbornene modified GMP [S6]	5' phosphate	5'	Diels-Alder reaction
	Levulinic Amide-GMP conjugate [S7]	5' phosphate	5'	Aldol reaction catalyzed by ribozyme 11DA$_{7\text{-}194}$
	Ternary conjugates of GMP [S8,S9]	5' phosphate	5'	
	5′-HS-PEG$_n$-GMP [S10]	5' phosphate	5'	Conjugation with thiol-reactive agents
	5′-deoxy-5′-thioguanosine-5′-monophosphorothioate (GSMP) [S10, S11]	5'	5'	Dephosphorylation, yielding a terminal sulfhydryl group
	acetal-protected aldehyde–GMP [S12]	5' phosphate	5'	Conjugation with hydrazine-containing moieties after deprotection
Dinucleotide	octadiynyl-dUpG (OdUpG) [S13]	OdU conjugated to GMP	5'	Splinted ligation
	Adenosine–guanosine conjugates [S14]	Biotinylated adenosine moieties conjugated to GMP	5'	
	Adenosine–guanosine conjugates with 2'-O-methylation [S15]	Methylated adenosine conjugated to methylated GMP	5'	
Guanosine analogue	5'-amino-5'-deoxyguanosine-5'-N-phosphoramidate (GNHP) [S16]	5'	5'	Conjugation with activated esters
	5-deoxy-5-hydrazinylguanosine [S17]	5'	5'	Conjugation with Fluorescein Isothiocyanate (FITC)
	5'-amino and 5'-thiol-hexaethylene glycol guanosine [S18]	5'	5'	Conjugation with electrophiles

TABLE S3. CLEAVAGE RATE OF E1-S1					
DATE	09/24/14	08/02/15	08/03/15	**Ave.**	**Std.**
RATE	-0.14	-0.12	-0.14	**-0.13**	**0.01**
TIME/MIN			ln(1-S/S0)		
0	0.00	0.00	0.00	0.00	0.00
0.5	-0.07	-0.03	-0.04	-0.05	0.02
1	-0.14	-0.07	-0.09	-0.10	0.04
1.5	-0.22	-0.15	-0.18	-0.18	0.03
2	-0.27	-0.23	-0.27	-0.26	0.03
2.5	-0.35	-0.34	-0.37	-0.35	0.02

TABLE S4. CLEAVAGE RATE OF E4-S1					
DATE	12/07/14	08/02/15	08/03/15	**Ave.**	**Std.**
RATE	-0.44	-0.48	-0.48	**-0.46**	**0.02**
TIME/MIN			ln(1-S/S0)		
0	0.00	0.00	0.00	0.00	0.00
0.5	-0.27	-0.28	-0.24	-0.26	0.02
1	-0.42	-0.46	-0.48	-0.46	0.03
1.5	-0.71	-0.69	-0.80	-0.74	0.06
2	-0.88	-1.01	-0.96	-0.95	0.07
2.5	-1.06	-1.16	-1.14	-1.12	0.05

TABLE S5. CLEAVAGE RATE OF E5-S1					
DATE	09/29/14	10/06/14	08/05/15	**Ave.**	**Std.**
RATE	-0.14	-0.10	-0.15	**-0.13**	**0.02**
TIME/MIN			ln(1-S/S0)		
0	0.00	0.00	0.00	0.00	0.00
0.5	-0.07	-0.05	-0.06	-0.06	0.01
1	-0.16	-0.09	-0.11	-0.12	0.04
1.5	-0.21	-0.12	-0.20	-0.18	0.05
2	-0.26	-0.19	-0.29	-0.25	0.05
2.5	-0.34	-0.28	-0.40	-0.34	0.06

TABLE S6. CLEAVAGE RATE OF E1-S2					
DATE	08/01/15	08/04/15	08/06/15	**Ave.**	**Std.**
RATE	-0.10	-0.11	-0.11	**-0.11**	**0.01**
TIME/MIN			ln(1-S/S0)		
0	0.00	0.00	0.00	0.00	0.00
0.5	-0.03	-0.03	-0.04	-0.03	0.01
1	-0.07	-0.06	-0.08	-0.07	0.01
1.5	-0.12	-0.11	-0.12	-0.12	0.01
2	-0.17	-0.17	-0.19	-0.18	0.01
2.5	-0.25	-0.23	-0.28	-0.25	0.02
5	-0.53	-0.58	-0.61	-0.58	0.04

TABLE S7. CLEAVAGE RATE OF E1-S3					
DATE	80115	8/4/15	8/6/15	**Ave.**	**Std.**
RATE	-0.09	-0.08	-0.07	**-0.08**	**0.01**
TIME/MIN			ln(1-S/S0)		
0	0.00	0.00	0.00	0.00	0.00
0.5	-0.04	-0.04	-0.03	-0.03	0.00
1	-0.08	-0.08	-0.06	-0.07	0.01
1.5	-0.13	-0.11	-0.08	-0.11	0.02
2	-0.17	-0.15	-0.12	-0.15	0.03
2.5	-0.24	-0.21	-0.18	-0.21	0.03
5	-0.48	-0.42	-0.40	-0.43	0.04

8. Supplementary references

[S1] D. Shin, R. W. Sinkeldam, Y. Tor, *J. Am. Chem. Soc.* **2011**, *133*, 14912-14915.

[S2] F. Huang, G. Wang, T. Coleman, N. Li, *RNA* **2003**, *9*, 1562–1570.

[S3] N. Li, C. Yu, F. Huang, *Nucleic Acids Research* **2005**, *33*.

[S4] F. Huang, Y. Shi, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6254–6257.

[S5] J. R. Sampson, O. C. Uhlenbeck, *Proc. Natl. Acad. Sci.USA* **1988**,*85,* 1033–1037.

[S6] J. Schoch, S. Ameta, A. Jaschke, *Chem. Commun.,* **2011**, *47*, 12536–12537.

[S7] S. Fusz *et al*., *J. Org. Chem.* **2008**, *73*, 5069–5077.

[S8] R. Fiammengo, K. Musilek, A. Jaschke, *J. AM. CHEM. SOC*. **2005**, *127*, 9271-9276.

[S9] B. Seelig, A. Jaschke, *Bioconjugate Chem*. **1999**, *10*, 371−378.

[S10] L. Zhang *et al., Bioconjugate Chem.* **2001**, *12,* 939−948.

[S11] I. H. Kim, S. Shin, Y. J. Jeong, S. S. Hah, *Tetrahedron Letters* **2010**, *50*, 3446–3448.

[S12] S. Pfander *et al., Nucleic Acids Research*, **2007**, *35*

[S13] A. Samanta, A. Krause, A. Jaschke, *Chem. Commun.,* **2014**, *50*, 1313–1316.

[S14] J. Wolf, V. Dombos, B. Appel, S. Muller, *Org. Biomol. Chem.* **2008***, 6*, 899–907.

[S15] C. Pitulle, R. G. Kleineidam, B. Sproat, G. Krupp, *Gene* **1992**, *112,* 101–105.

[S16] D. Williamson, M. J. Cann, D. R. W. Hodgson, *Chem. Commun.* **2007**, 5096–5098.

[S17] M. Skipsey *et al., Nucleosides, Nucleotides and Nucleic Acids* **2013**, *32*, 670–681.

[S18] J. C. Schlatterer, A. Jaschke, *Biochem. Biophys. Res. Commun*. **2006**, *344,* 887–892.