Electronic Supplementary Information

Modified dinucleotide for site-specific RNA-labelling by transcription priming and click chemistry

Ayan Samanta[#], André Krause[#] and Andres Jäschke*

Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, Heidelberg 69120, Germany. Phone: +49 6221 54 4851, Fax: +49 6221 54 6430.

E-mail: jaeschke@uni-hd.de

[#] Both authors contributed equally to this work

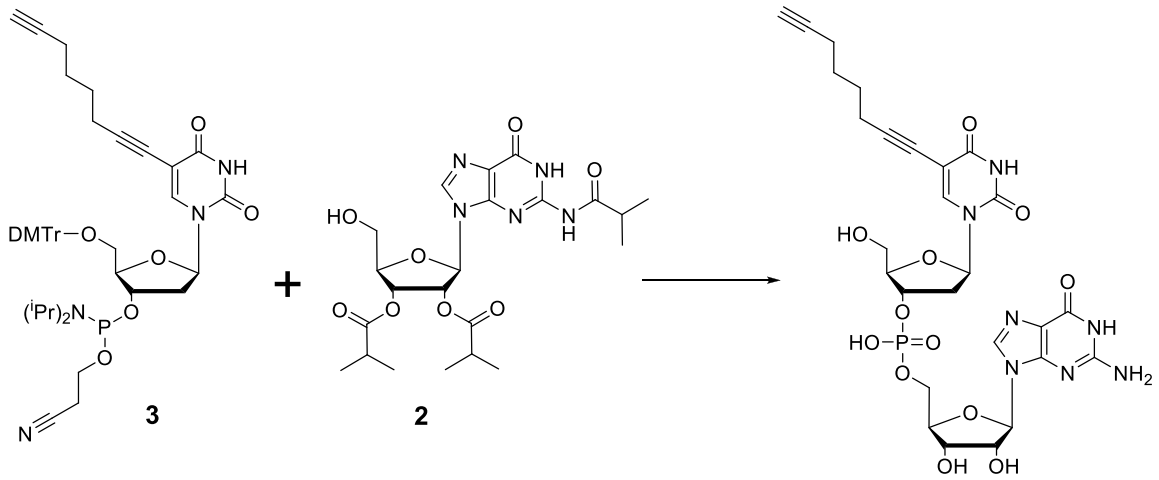
List of abbreviations:

DMTr-Cl = 4, 4´-dimethoxytrityl chloride, LC-ESI-HRMS = liquid-chromatography electron spray ionisation high-resolution mass spectrometry, TBHP = *tert*-butyl hydroperoxide.

General materials and methods for synthesis:

All chemicals, required for synthesis, were obtained from commercial providers and used without further purification. If not otherwise stated, reactions were performed under argon atmosphere in anhydrous solvents, previously stored over molecular sieves. Reactions were monitored by thin layer chromatography (TLC) on silica gel sheets Polygram® Sil G/UV254 (0.2 mm, 40 x 80 mm) and visualised under UV-light or by staining with blue-shift reagent [$Ce(SO_4)_2$; molybdophosphoric acid; sulfuric acid]. Flash column chromatography was performed using self-packed columns of silica gel 60 M (0.04-0.063 mm/230-240 mesh ASTM) or on pre-packed cartridges (puriFlash™ Silica High Capacity 50 µm, Interchim or TELOS™ Silica Flash Chromatography Columns) on an IntelliFlash™ 310 chromatography workstation (Varian). Preparative reversed-phase chromatography for final purification of dinucleotide initiator was carried out on pre-packed C18 cartridges (SuperFlash™ C18, 18% carbon C18, bonded and end-capped on 50µm silica support, Analogix or TELOS™ C18 Flash chromatography columns). Purified initiator dinucleotide was analysed by HPLC on an Agilent 1100 system equipped with a diode-array detector on an analytical Phenomenex Luna C18 column (5 µm, 5.0 × 250 mm) at a flow rate of 1 ml/min, eluting in a gradient of 100 mM triethylammonium acetate pH 7.0 (buffer A) and 100 mM triethylammonium acetate in 80 % acetonitrile (buffer B) (Fig. S1). LC-ESI-HRMS analyses of initiated transcripts were carried out on a Bruker microTOFQ-II ESI mass spectrometer combined with an Agilent 1200 Series HPLC system equipped with a multi-wavelength detector. Samples were resolved on a Phenomenex Kinetex C18 column (2.6 µm, 2.1 × 100 mm) in a gradient of 100 mM hexafluoroisopropanol/ 8.6 mM triethylamine pH 8.3 and methanol (LC-MS grade) at a flow rate of 0.2 ml/min. LC-MS data were analysed using Hyphenation Star PP (Version 3.2.44.0) and DataAnalysis (Version 4.0, SP 4) software (Bruker Daltonics). MS-spectra were deconvoluted using Maximum Entropy deconvolution. High-resolution mass spectra were obtained by internal calibration (enhanced quadratic mode) using ESI Tunemix (Fluka) as a calibrant. Calculated molecular weights refer to the m/z values given by the DataAnalysis software. NMR spectra were recorded on a Varian Mercury Plus 500 MHz NMR spectrometer (^{1}H: 500 MHz, ^{13}C: 126 MHz, ^{31}P: 202 MHz) and calibrated using the residual undeuterated solvent signal as internal standard.

Synthesis of OdUpG initiator dinucleotide (1):



OdUpG initiator dinucleotide (**1**)

Triisobutyryl guanosine (**2**) was synthesised according to published procedure.[1] A flame-dried 1 ml Schlenk flask, equipped with a stirring bar, was charged with 33 mg (0.067 mmol) of **2** and a solution of 55 mg (0.066 mmol) of **3** (Baseclick, Germany) in 1.2 ml of anhydrous acetonitrile under argon. 1.47 ml (0.66 mmol) of a 0.45 M solution of tetrazole in anhydrous acetonitrile was added and the mixture was stirred for 2 h at room temperature. 95 µl (0.95 mmol) of TBHP was added. After stirring for an additional 10 min the solvent was evaporated under reduced pressure and the formation of the intermediate fully protected dinucleotide was verified by MALDI-MS. Deprotection of ^{i}Bu- and cyanoethyl-protected functionalities was achieved by dissolving the residue in 5 ml of 28-30 % aqueous ammonia and heating for 8 h at 55 °C in a sealed vessel. The solution was evaporated under reduced pressure and the residue was redissolved in 20 ml of water. Lyophilisation yielded 80 mg of a white powder which was detritylated in 5 ml of 2 % dichloroacetic acid in dichloromethane for 2 min. The mixture was neutralised by the addition of 5 ml of saturated aqueous NH_4HCO_3 solution and the organic layer was extracted 4X with 5 ml of saturated NH_4HCO_3 solution. The combined aqueous fractions were lyophilised. Reversed phase chromatography on an Analytix-C18 column (flow-rate 12 ml/min, water: acetonitrile 1:0 to 7:3, 2 %/min) yielded 12 mg (27 %) of OdUpG initiator dinucleotide (**1**). ESI-HRMS: cal. $[M-H]^-$ $C_{27}H_{31}N_7O_{12}P$= 676.1774, found 676.1848.

General materials and methods for molecular biology:

All oligonucleotides were purchased from IBA (Göttingen, Germany). Radioactive gels were exposed to storage phosphor screens (GE Healthcare) and scanned (ex: 633 nm, em: 390 BP) with a Typhoon 9400 variable mode imager (GE Healthcare). For all quantifications, radioactive bands were background-corrected to the object average of a region in the gel not containing any radioactive bands and quantified using ImageQuant software (Molecular Dynamics; version 5.2). All azide-bearing compounds used for Fig. S2 were obtained from commercial providers except biotin azides (compounds **5** and **7**) and sugar azides (compounds **11** and **12**) which were synthesised according to published procedures.[2, 3] The copper-stabilising ligand, tris-(3-hydroxypropyltriazolylmethyl)-amine (THPTA) was synthesised following a published protocol.[4]

Optimisation of transcription priming by three different polymerases:

The corresponding sense and antisense strands (DNA1 and DNA2 for T7; DNA3 and DNA4 for T3; DNA5 and DNA6 for SP6 RNAP, Table S2) were mixed together in hybridisation buffer (10 mM Tris-HCl pH 7.5, 80 mM NaCl) at a concentration of 8 µM each followed by heating at 70 ^{0}C for 2 min and then cooling down to room temperature over 15 min to ensure annealing of the two strands. The double-stranded transcription template was added to the transcription mixture [40 mM Tris-HCl (pH 8.1, at 25 ^{0}C), 1 mM spermidine, 22 mM $MgCl_2$, 0.01 % Triton-X-100, 10 mM dithiothreitol, 40 µg/ml bovine serum albumin, 1 µM [α-^{32}P]-CTP (10 µCi/µl, 800 Ci/mmol; Hartmann Analytic), 4 mM initiator nucleotide (OdUpG or GMP as reference)] at a concentration of 0.2 µM containing GTP and the other three NTPs (as described in individual experiments), followed by an additional annealing step as described above. The respective RNAPs [T7, SP6 RNAPs (Fermentas) and T3 RNAP (Promega)] were added at a concentration of 2 U/µl and the reaction mixture was incubated for 2 h at 37 ^{0}C. The resulting mixture was subjected to aq. phenol-diethyl ether extraction, followed by ethanol-precipitation of the transcript from NH_4OAc buffer (0.5 M, pH 5.6) using glycogen (RNA grade; Fermentas) as carrier. The RNA pellet was washed with 80 % ethanol and redissolved in water.

Biotinylation of OdUpG-primed transcripts by CuAAC and streptavidin electrophoretic mobility shift assay (Strep-EMSA):

The CuAAC reaction was performed as described previously.[5] RNA concentration was maintained between 0.1 – 2.6 µM in the CuAAC reaction mixture. Briefly, reactions were performed in sodium

phosphate buffer (50 mM, pH 7.0) containing 0.5 mM $CuSO_4$, 5 mM sodium ascorbate and 2.5 mM THPTA ligand. Biotin azide (compound **5**) concentration was maintained at 0.25 mM for octadiynyl-modified transcripts. The reaction mixture was incubated at 37 ^{0}C for 2 h after which the RNA was isolated by ethanol precipitation from NH_4OAc buffer (0.5 M, pH 5.6) using glycogen as carrier. The resulting RNA pellet was washed with 80 % ethanol and redissolved in water followed by incubation with streptavidin (1 mg/ml stock in water, New England Biolabs, at least 15-fold molar excess over the RNA), and analysed on an 18 % 1 mm denaturing polyacrylamide gel (1 h at 25W with a pre-run of 15 min at 25W). 6X DNA loading dye (Fermentas) was used for loading the samples onto a gel. Labelling efficiencies were calculated in percentage from ratiometric comparison of labelled and unlabelled RNA-bands (Fig. 2). Relative transcript (modified + unmodified) yields were calculated in percentage by normalising the transcript amounts of the corresponding GMP-primed reactions in Fig. 2 to 100 % (Fig. 3).

Syntheses of transcripts conjugated with various organic moieties:

A template with two 2´-OMe substitutions at the 5´-end of the antisense strand was used for these experiments.[6] DNA7 and DNA8 (Table S2) were used for preparing the template. The transcription was carried out as described earlier with 4 mM of OdUpG, 0.8 mM of GTP, 1 mM of the other three NTPs and 2 U/µl of T7 RNAP. The concentration of the azide component in the CuAAC reaction was always maintained at 0.25 mM. All other conditions were kept identical as before. The azide component was omitted for the negative controls. Samples were analysed on 18 % 1 mm denaturing polyacrylamide gels [for Azide-Fluor 488 (**8**), ex: 488 nm, em: 520 BP 40; Alexa Fluor® 594 Azide (**9**), ex: 532 nm, em: 610 BP 30; Cy5-Azide (**10**), ex: 633 nm, em: 670 BP 30]. The fluorescence scans were shown as a single overlaid image of all three channels (Fig. S2).

Labelling long RNA with a fluorophore:

Bacillus subtilis genomic DNA was isolated by GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich). Genomic DNA (~150 ng) was added to 50 µl of the PCR mixture [1X Hybrid buffer, 0.2 mM dNTPs, 4 % DMSO, 0.5 µM of each primer (DNA10 and DNA11, Table S2) and 1 U of Polymerase X - Hybrid DNA Polymerase (Roboklon)] and subjected to the following PCR conditions: 98 ^{0}C 30 s, [98 ^{0}C 10 s, 67 ^{0}C 30 s, 72 ^{0}C 30 s] 25 cycles, 72 ^{0}C 7 min. The PCR product was purified using a QIAquick PCR Purification Kit (Qiagen).

The 252 nucleotide double-stranded DNA pool, bearing constant 5′- and 3′- ends [5′-TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGC-3′ with the promoter sequence underlined (5′-constant region) and 5′-GGCACCACGGTCGGATCCAC-3′ (3′-constant region)] and a randomised region in between, was used for transcription-priming with OdUpG by T7 RNAP. The resulting transcript has a length of 233 nucleotides.

For transcription, the aforesaid PCR products were used as transcription templates in combination with 4 mM of OdUpG, 0.8 mM of GTP, 1 mM of the other three NTPs and 2 U/µl of T7 RNAP. For click labelling, Cy5-azide (**10**) was used at a concentration of 0.25 mM. All other conditions were kept identical to previous experiments. Samples were analysed on 8 % 1 mm denaturing polyacrylamide gels. A GMP-primed transcript was used as a negative control in these experiments (Fig. S3).

Ligation of OdUpG-primed transcript:

RNA1 (Table S2) was prepared by transcription-priming with either OdUpG or UpG (IBA). The UpG-primed transcript served as a control in the following experiments.

For enzymatic phosphorylation of the 5′-end of transcripts, ~14 pmol of gel-purified transcripts were incubated in kinase reaction mixture [50 mM Tris-HCl (pH 7.6 at 25 ^{0}C), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 2 mM ATP, 15 U of T4 polynucleotide kinase (PNK) (Fermentas); final volume: 20 µl] at 37 ^{0}C for 1 h after which the RNAs were purified by ethanol precipitation from NH_4OAc buffer (0.5 M, pH 5.6) using glycogen as carrier.

For splinted ligation, a mixture of either OdUpG- or UpG-primed transcript (~0.35 µM), RNA2 (2.5 µM) and DNA9 (2.5 µM) [Table S2] in ligation buffer [50 mM Tris-HCl (pH 7.4 at 25 ^{0}C), 10 mM $MgCl_2$, 50 µM ATP, 5 mM DTT; final volume: 20 µl] was heated at 90 ^{0}C for 1 min followed by cooling down to room temperature over 15 min to allow the ligation fragments to anneal on the complementary DNA splint after which ligases were added [T4 RNA ligase 2 (500 nmol/µl) (home-prep), T4 RNA ligase 1 (0.2 U/µl) (Fermentas) and T4 DNA ligase (0.6 U/µl) (Fermentas)] and the resulting mixture was incubated at 37 ^{0}C for 1 h. The DNA splint was removed by the addition of DNase I (0.5 U/µl) (Fermentas) and incubating the reaction mixture at 37 ^{0}C for 15 min. The reaction was stopped by the addition of formamide gel loading buffer and samples were analysed on an 18 % 1 mm denaturing polyacrylamide gel (Fig. S4). Since the gel was scanned in phosphorimaging mode, only the donor RNA (primed transcript, 26 nucleotides considering the alkyne-bearing residue) and the ligated product (66 nucleotides) are visible.

Supplementary figures:

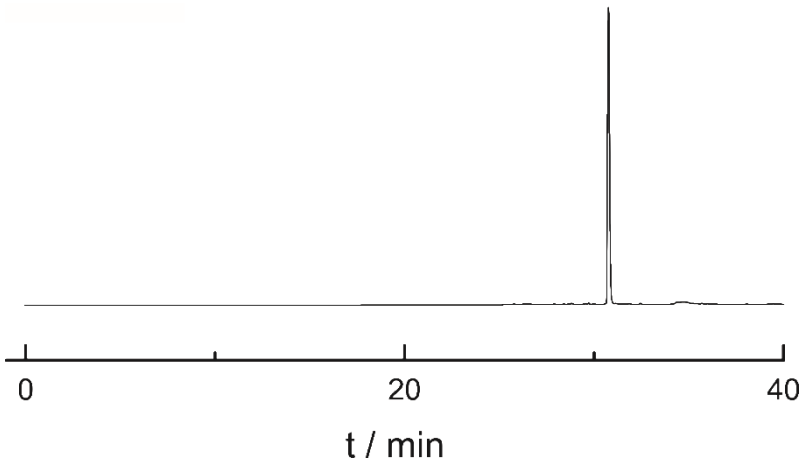


Fig. S1 HPLC analysis of purified OdUpG initiator dinucleotide (**1**) (30.8 min) based on UV-absorption at 254 nm.

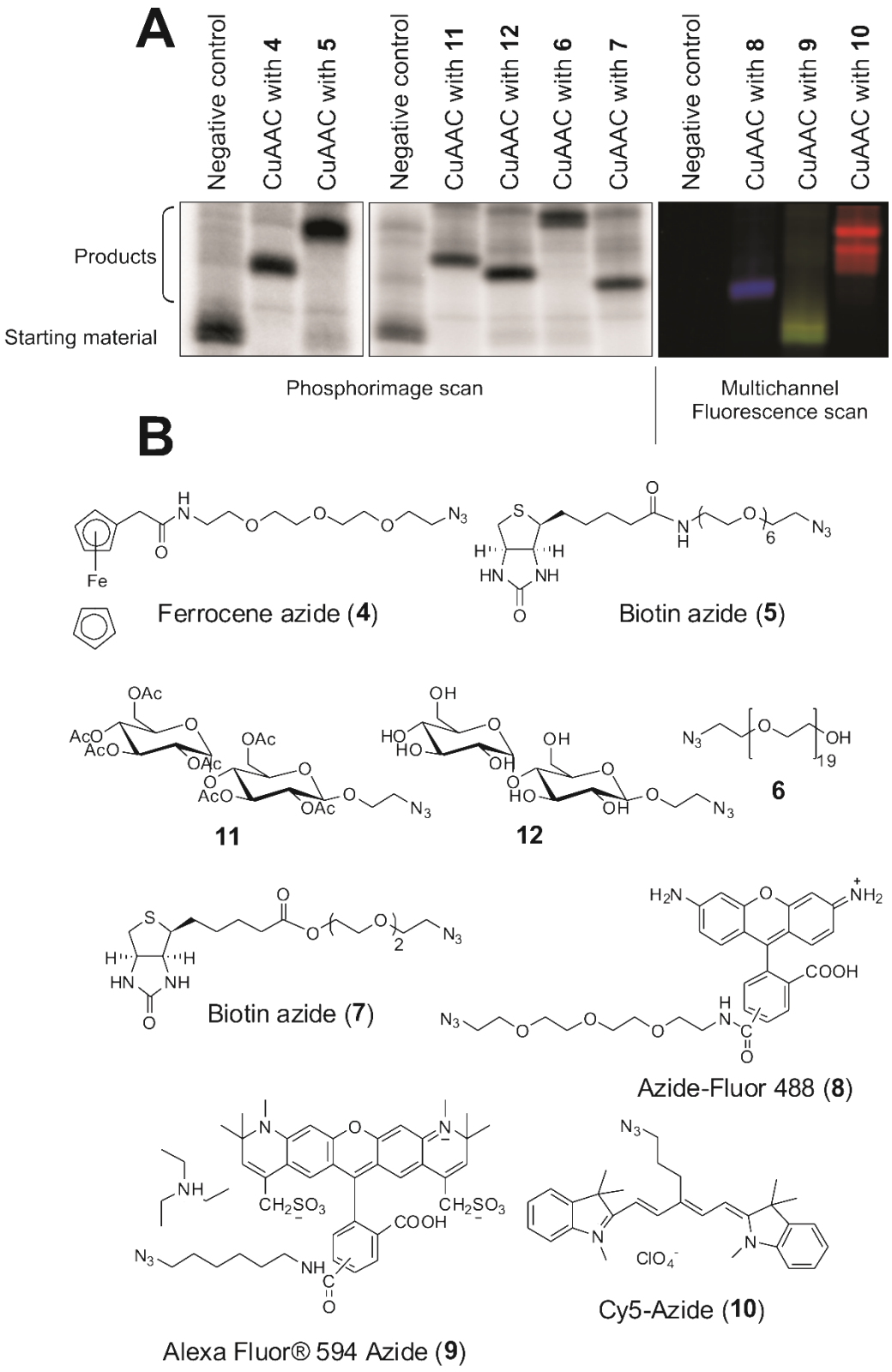


Fig. S2 Syntheses of transcripts conjugated with various organic moieties. **A**) Gel analysis. **B**) Chemical structures.

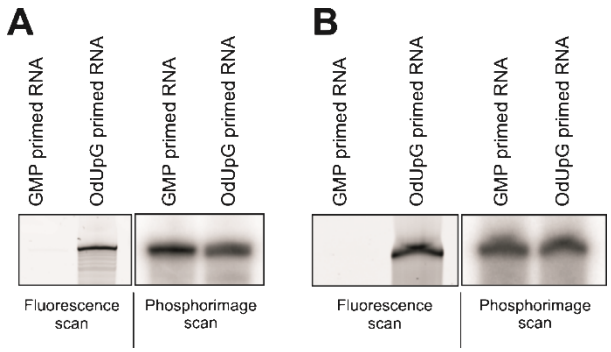


Fig. S3 Labelling long RNA with a fluorophore. **A**) Glycine riboswitch and **B**) random RNA pool were modified with OdUpG by T7 RNAP transcription priming. Radioactive transcripts were purified and subjected to CuAAC with Cy5-azide (**10**) in both cases.

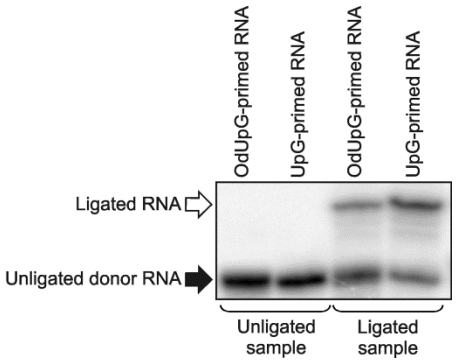


Fig. S4 Splinted ligation of OdUpG- and UpG-primed transcripts.

Supplementary tables:

Table S1 ESI-HRMS data of initiated transcripts

Transcription primed with	Sequence of the primed transcript	Calc. $[M-H]^-$	Observed
GMP	5′-**pG**GAGCUCAGCCUACGAGCCUGAGCC-3′	8090.0885	8090.1058
OdUpG (**1**)	5′-**OdUpG**GAGCUCAGCCUACGAGCCUGAGCC-3′	8404.2370	8404.2949

Table S2 Oligonucleotide sequences used in this study

Name	Sequence[a]	Remarks
DNA1	5′-TCTAATACGACTCACTATAGGAGCTCAGCCTACGAGCCTGAGCC-3′	T7 template sense strand
DNA2	5′-GGCTCAGGCTCGTAGGCTGAGCTCCTATAGTGAGTCGTATTAGA-3′	T7 template antisense strand
RNA1[b]	5′-GGAGCUCAGCCUACGAGCCUGAGCC-3′	Transcript encoded from a duplex of DNA1+DNA2
DNA3	5′-TCAATTAACCCTCACTAAAGGGAGACAGCCTACGAGCCTGAGCC-3′	T3 template sense strand
DNA4	5′-GGCTCAGGCTCGTAGGCTGTCTCCCTTTAGTGAGGGTTAATTGA-3′	T3 template antisense strand
DNA5	5′-TCATTTAGGTGACACTATAGAAGAGCAGCCTACGAGCCTGAGCC-3′	SP6 template sense strand
DNA6	5′-GGCTCAGGCTCGTAGGCTGCTCTTCTATAGTGTCACCTAAATGA-3′	SP6 template antisense strand
DNA7	5′-TCTAATACGACTCACTATA-3′	T7 template sense strand
DNA8	5′-**GG**AGTGAAGGCTGATGACCTATAGTGAGTCGTATTAGA-3′	T7 template antisense strand (methoxy)[c]
RNA2	5′-GGGUUGGGAAGAAACUGUGGCACUUCGGUGCCAGCAACCC-3′	For splinted-ligation to RNA1
DNA9	5′-GGCTCAGGCTCGTAGGCTGAGCTCCGGGTTGCTGGCACCGAAGTGCCACAGTTTCTTCCCAACCC-3′	DNA-splint for ligation between RNA1 and RNA2
DNA10	5′-TAATACGACTCACTATAGGGATATGAGCGAATGACAGCAAGGG-3′	For. primer[d]
DNA11	5′-GGTTCTCTGTCCTGGCACCTGAAAGTTTACTTTGC-3′	Rev. primer[d]

[a] Promoter sequences are underlined.
[b] Modified with respective initiator nucleotides at the 5′-end by transcription-priming as indicated in individual experiments.
[c] Two bold G′s are carrying 2′-methoxy modification.
[d] Primer pair required for amplifying the transcription template of glycine riboswitch from *B. subtilis* genomic DNA by PCR, carrying a T7 RNAP promoter.

Supplementary references:

1. B. Seelig and A. Jaschke, *Bioconjug. Chem.*, 1999, **10**, 371-378.
2. N. Vundyala, C. Sun, F. Sidime, W. Shi, W. L'Amoreaux, K. Raja and R. M. Peetz, *Tetrahedron Lett.*, 2008, **49**, 6386-6389.
3. W. Wang, C. Jin, L. Guo, Y. Liu, Y. Wan, X. Wang, L. Li, W. Zhao and P. G. Wang, *Chem. Commun.*, 2011, **47**, 11240-11242.
4. V. Hong, S. I. Presolski, C. Ma and M. G. Finn, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 9879-9883.
5. M. L. Winz, A. Samanta, D. Benzinger and A. Jaschke, *Nucleic Acids Res.*, 2012, **40**, e78.
6. C. Kao, S. Rudisser and M. Zheng, *Methods*, 2001, **23**, 201-205.