Supplementary Information

Small-molecule fluorescent probes for specific RNA targets

Asako Murata,[a] *Shin-ichi Sato,*[a] *Yoshinori Kawazoe,*[b] *and Motonari Uesugi**[a, b]

[a]*Institute for Integrated Cell-Material Sciences.* [b]*Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan. Kyoto University Gokasho, Uji, Kyoto 611-0011, Japan.*

Fax: +81 (0)774-38-3226; Tel: +81 (0)774-38-3225; E-mail: uesugi@scl.kyoto-u.ac.jp

Contents Pages

General information

Reagents. BHQ1 carboxylic acid and BHQ1 amine were purchased from Biosearch Technologies. Alexa594 succinimidyl ester was purchased from Molecular Probes, Cy3 succinimidyl ester from GE Healthcare, and fluorescein isothiocyanate from Sigma Aldrich. EAH sepharose 4B was obtained from GE Healthcare for preparation of BHQ1 affinity resin.

Techniques. ESI-mass spectra were measured using an LCMS-2020 system (Shimadzu). MALDI-TOF mass spectra were measured using an Axima-LNR mass spectrometer (Shimadzu). HPLC was performed using a Resource reversed phase column (6.4 × 30 mm), with a linear gradient of 0.1% trifluoroacetic acid and acetonitrile (0−90%) as an eluent, and a flow rate of 1.0 mL min^{-1}. UV spectra were measured using a HITACHI U-3010 spectrophotometer. Sequencing was carried out using an Applied Biosystems 3130/3130xl Genetic Analyzer (Applied Biosystems). Fluorescence spectra were measured with an LS 55 fluorescence spectrometer (Perkin Elmer). The fluorescence emission data from multiple samples were collected by an MTP-800 multi microplate reader (CORONA). The real-time monitoring of transcription was performed using a 7500/7500Fast Real-Time PCR System (Applied Biosystems).

Synthesis of the fluorophore-BHQ1 conjugates

Amino-PEG6-BHQ1 1. To a solution of N-Fmoc-amido-PEG6-acid (4 mg, 7 μmol) and BHQ-1 amine (1.6 mg, 3.4 μmol) in DCM-DMF were added diisopropylethylamine (0.03 mmol) and 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) (6.5 mg, 17 μmol) at 0 °C. This solution was stirred overnight at room temperature, diluted with $CHCl_3$, and washed with saturated $NaHCO_3$ and brine. The combined organic layer was dried over anhydrous

Na_2SO_4 and concentrated in vacuo. This crude product was used for the next reaction without purification. The Fmoc group was removed with 20% piperidine in DMF, and the solvent was evaporated under reduced pressure. The resulting residues were dissolved in water and washed three times with Et_2O to remove Fmoc residues. The aqueous layer was concentrated under reduced pressure and purified by reversed-phase HPLC to give compound **1** (1.9 μmol, 56% yield determined by UV absorbance at 534 nm). UV-Vis (methanol), λ_{max} = 540 nm; ESI MS $[M+H]^+$ calcd for $C_{40}H_{59}N_8O_{10}$: 811.44, found 811.00.

Conjugate 2 (Alexa594-PEG6-BHQ1). To a solution of compound **1** (540 nmol) in DMF-DMSO were added triethylamine (large excess) and Alexa594 succinimidyl ester (0.8 mg, 980 nmol), and the resulting solution was stirred overnight in dark. The reaction mixture was diluted with water and purified by reversed-phase HPLC to afford conjugate **2** (61.2 nmol, 11% yield determined by UV-Vis absorbance at 590 nm). UV-Vis (10 mM Tris-HCl, pH 7.5, 100 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$), λ_{max} = 543 nm, 588 nm; MALDI-TOF MS $[M+H]^+$ calcd for $C_{75}H_{91}N_{10}O_{20}S_2$: 1516.71, found: 1516.86.

Conjugate 3 (Cy3-PEG6-BHQ1). Cy3 succinimidyl ester (0.7 mg, 913 μmol) was coupled to the amino group of compound **1** in the same manner described in the synthesis of conjugate **2**. The resulting solution was diluted with water and purified by reversed-phase HPLC to afford conjugate **3** (28.6 nmol, 5% yield determined by UV-Vis absorbance at 548 nm). UV-Vis (10 mM Tris-HCl, pH 7.5, 100 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$), λ_{max} = 519 nm; MALDI-TOF MS $[M+H]^+$ calcd for $C_{71}H_{95}N_{10}O_{17}S_2$: 1424.70; found: 1425.03.

Conjugate 4 (fluorescein-PEG6-BHQ1). Fluorescein isothiocyanate (1.1 mg, 2.7 μmol) was coupled to the amino group of compound **1** in the same manner described in the synthesis of conjugate **2**. The resulting solution was diluted with water and purified by reversed-phase HPLC to afford conjugate **4** (13 nmol, 2% yield determined by UV-Vis absorbance at 494 nm). UV-Vis (10

mM Tris-HCl, pH 7.5, 100 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$), λ_{max} = 497 nm; MALDI-TOF MS $[M+H]^+$ calcd for $C_{61}H_{70}N_9O_{15}S$: 1201.32; found: 1200.97.

Preparation of BHQ1 affinity resin

To EAH sepharose 4B resin (0.5 ml, 5 μmol of amino group) were added BHQ-1 carboxylic acid (0.5mg, 1 μmol), 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride: DMT-MM (29.5 mg, 0.1 mmol), and triethylamine (0.1 mmol) in dioxane-H_2O (3:1, v/v). After shaking overnight at room temperature, the resin was drained and washed with dioxane-H_2O. To the resulting resin were added acetic acid (30 mg, 0.5 mmol) and DMT-MM (147.4 mg, 0.5 mmol) to block unreacted amino groups on the resin. After overnight incubation, the resin was drained and washed with aqueous dioxane and aqueous methanol. The resin was stored in 3 volumes of methanol-H_2O (1:3, v/v) at -20 °C.

In vitro selection procedures

The original double-stranded DNA pool was constructed by PCR, using a synthesized oligonucleotide containing 60 random nucleotides [5'- GAA TTC CGC GTG TGC ACA CC - N60 - GTC CGT TGG GAT CCT CAT GG -3'] as a template with Platinum Pfx DNA polymerase (Invitrogen). Primers used for PCR amplification were: forward [5'- GCT AAT ACG ACT CAC TAT AGG GAA TTC CGC GTG TGC ACA CC -3' (T7 promoter sequence is underlined)] and reverse [5'- CCA TGA GGA TCC GAA CGG AC -3']. The original DNA pool was transcribed in vitro with T7 RNA polymerase (TAKARA) to yield the first RNA pools. After transcription, reaction mixtures were treated with RQ1 DNase I (Promega), and precipitated with 2.5 M ammonium acetate/isopropanol. The resulting RNA was subjected to a NAP-5 column (GE Healthcare) to remove unincorporated NTPs, precipitated with 2.5 M

ammonium acetate/isopropanol, pelleted by centrifugation, and used for the selection step. In each selection cycle, RNA was annealed in binding buffer (10 mM Tris-HCl, 100 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$, pH7.5) and incubated on the BHQ1-immobilized affinity resin for 30 minutes at 4 ºC. The resin was then drained and washed 6 times with binding buffer to remove the non-binding species. Bound RNA species were eluted using binding buffer saturated with free BHQ-1 amine, 3 times for 10 min each. The eluted fractions were pooled and precipitated with ethanol. Selected RNAs were reverse transcribed using Prime Script reverse transcriptase (TAKARA), and the resulting cDNA was PCR-amplified with the forward and the reverse primers described above. The DNA templates were transcribed in vitro, and the resulting RNAs were subjected to the next round of selection. After 13 rounds of selection, enriched RNAs were reverse-transcribed, and converted to dsDNA by PCR. The resulting DNAs were ligated into pUC19 at the *Eco*RI and *Bam*HI sites, and transformed to *E. coli* strain DH5α. Twenty-eight clones were isolated and sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems).

Fluorescence measurement and K_d determination

The dsDNA templates for clone A1-A5 were prepared by PCR amplification of individual clones. The dsDNA encoding the mutated version of A1 (A1m) was PCR amplified using the A1 clone as a template, forward primer [5'- AGT GGA TCC CGC GTG TGC ACA CC -3'], and reverse primer [5'- TAG AGA ATT CGA ACG GAC CAg ggt ccg TCC GAA TTT ATC TAG GCC -3' (changes in the A1 sequence are shown in small letters)]. The amplified dsDNA was inserted into the *Bam*HI and *Eco*RI sites in pET28 and cloned into *E.coli* DH5α. The dsDNA template for T7 transcription of A1m was PCR amplified using this vector as template, forward primer forward [5'- GCT AAT ACG ACT CAC TAT AGG GAA TTC CGC GTG

TGC ACA CC -3' (T7 promoter sequence is underlined)], and reverse primer [5'- GTC GAC GGA GCT CGA ATT CGA ACG GAC -3']. The dsDNA templates for the short versions of A1 and A1m (A1sh and A1m-sh) were prepared by annealing of two synthetic ssDNAs: [5'- GCT AAT ACG ACT CAC TAT AGG GCC CAG GCC TAG ATA AAT TCG GAG CCT GGG CCC -3'] and [5'- GGG CCC AGG CTC CGA ATT TAT CTA GGC CTG GGC CCT ATA GTG AGT CGT ATT AGC -3'] for A1-sh; [5'- GCT AAT ACG ACT CAC TAT AGG GCC CAG GCC TAG ATA AAT TCG GAc gga ccc ggg -3'] and [5'- ccc ggg tcc gTC CGA ATT TAT CTA GGC CTG GGC CCT ATA GTG AGT CGT ATT AGC -3'] for A1m-sh. RNAs were transcribed from the dsDNA templates using T7 RNA polymerase and purified as described above. Conjugate **2** was dissolved in the binding buffer in the presence of RNA (0, 0.3, 1, 3, 10, 30 or 100 μM), and fluorescence intensity was measured using an MTP-800 multi microplate reader (CORONA) with an Ex/Em = 550/610 nm filter set. The K_d values of A1-A4 were estimated from the fluorescence titration data, using the equation:

$$F\text{obs} = A((([\text{conjugate } \mathbf{2}]_T + [\text{aptamer}]_T + K_d) - (([\text{conjugate } \mathbf{2}]_T + [\text{aptamer}]_T + K_d)^2 - 4[\text{conjugate } \mathbf{2}]_T[\text{aptamer}]_T)^{1/2}) / 2[\text{conjugate } \mathbf{2}]_T)$$

where A is the increase in fluorescence at saturating A1 concentration (F_{max} - F_{min}), K_d is the dissociation constant, and [conjugate **2**]$_T$ and [aptamer]$_T$ are the total concentrations of conjugate **2** and aptamer A1-A4, respectively.

Resin binding assay

Clones A1-A5 were prepared as described above. RNAs were annealed in binding buffer and UV absorbance was measured at 260 nm using a HITACHI U-3010 spectrophotometer.

After a 30-minute incubation at 4 ºC of RNA solution on the BHQ1-immobilized affinity resin, or Alexa594-immobilized resin as a negative control, the flow-through fraction was collected and UV absorbance was measured at 260 nm. The percent of RNA bound to the resin was calculated as $100(Abs_T - Abs_F)/Abs_T$, where Abs_T is UV absorbance of the RNA solution at 260 nm before binding to the resin, and Abs_F is UV absorbance of the flow-through fraction at 260 nm after binding to the resin.

Fluorescence spectra measurements

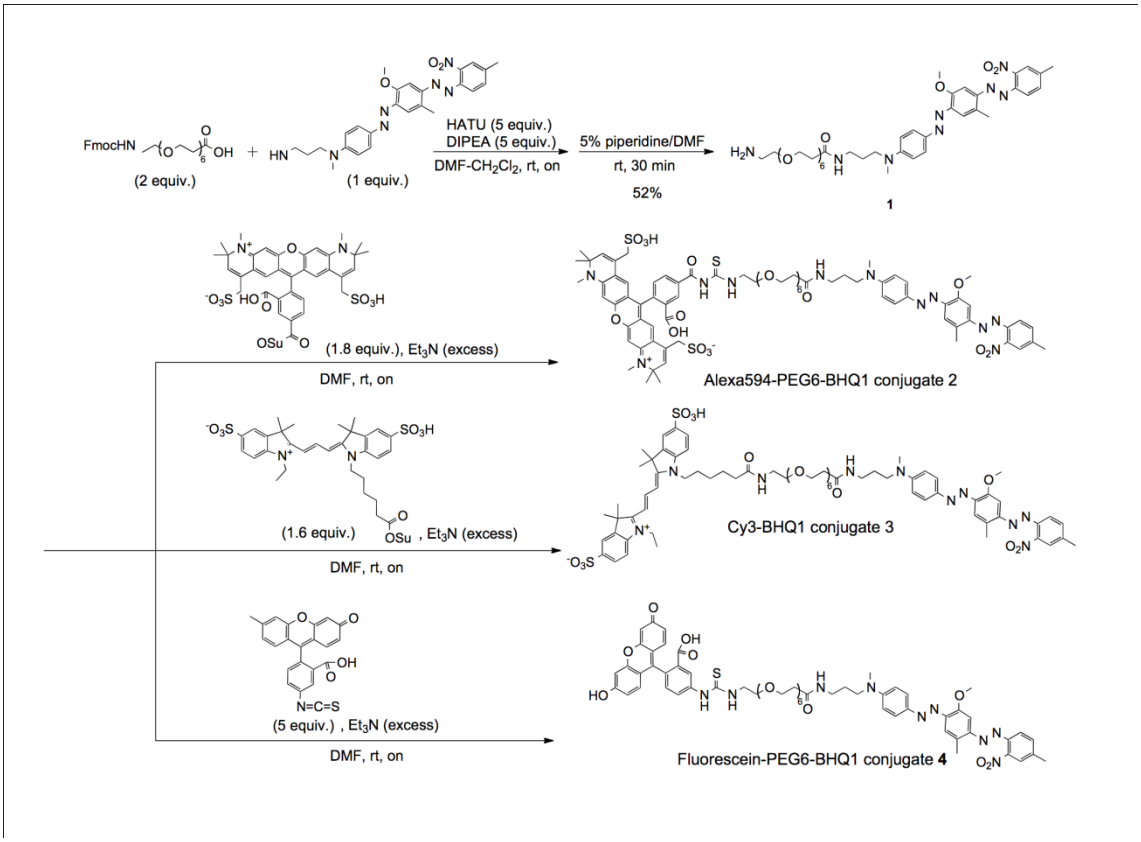
Fluorescence spectra of each fluorophore-BHQ1 conjugate (0.5 μM) were obtained at 4 °C in the binding buffer (10 mM Tris-HCl, 100 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$, pH7.5) in the presence of A1 (0, 0.3, 1, 3, 10, 30 or 100 μM). The excitation wavelengths and fluorescence emissions (Ex/Em) monitored for each conjugates were 570/590-695 nm for conjugate **2**, 520/540-645 nm for conjugate **3**, and 470/490-645 nm for conjugate **4**.

Transcription monitoring

DNA template encoding A1 or A1m sequence were prepared as described above. RNAs were transcribed for 2 hours at 30 °C, from 1 μg of DNA template in 20 μL of transcription buffer (MEGAscript T7 Kit, Ambion) containing 5 μM of conjugate **2** or **3.** Fluorescence emission data were collected every 3 min using a 7500/7500Fast Real-Time PCR System (Applied Biosystems) with the filter for ROX dye or Cy3 dye. The resulting reaction solution was treated with DNase I (Promega) for 30 min at 37 °C, and subjected to native-PAGE analysis. The gel was stained with ethidium bromide, and the band density was quantified using NIH imageJ software.

Real time monitoring of mRNA and protein synthesis of GFP

The dsDNA for A1 and A1m was PCR amplified using A1 and A1m clone as a template, forward primer [5'- CGT TAC TAG TGG ATC CGC GTG TGC ACA CC -3'], and reverse primer [5'- GTT AGC AGC CGG ATC GAA TTC GAA CGG AC -3']. The amplified dsDNA was inserted into the BamH I sites in pIVEX Control Vector GFP (5 Prime Inc.), using an In-Fusion Dry-Down PCR Cloning Kit (Clontech Laboratory Inc.), and cloned into *E. coli* Fusion-Blue. The resulting pIVEX-GFP-A1 or pIVEX-GFP-A1m expression plasmids were used as DNA templates (0.5 μg) in 20 μL of cell-free protein expression mixture (RTS 100 *E.coli* HY kit, 5 Prime Inc.) containing 2.5 μM of conjugate **2**, for 1 h at 30 °C. Fluorescence emission data were collected every 3 min using a 7500/7500Fast Real-Time PCR System (Applied Biosystems) with the filter for FAM and ROX dye. The resulting reaction solutions were subjected to western blot analysis and RT-PCR analyses.



Scheme S1. Synthesis of the fluorophore-BHQ1 conjugates.

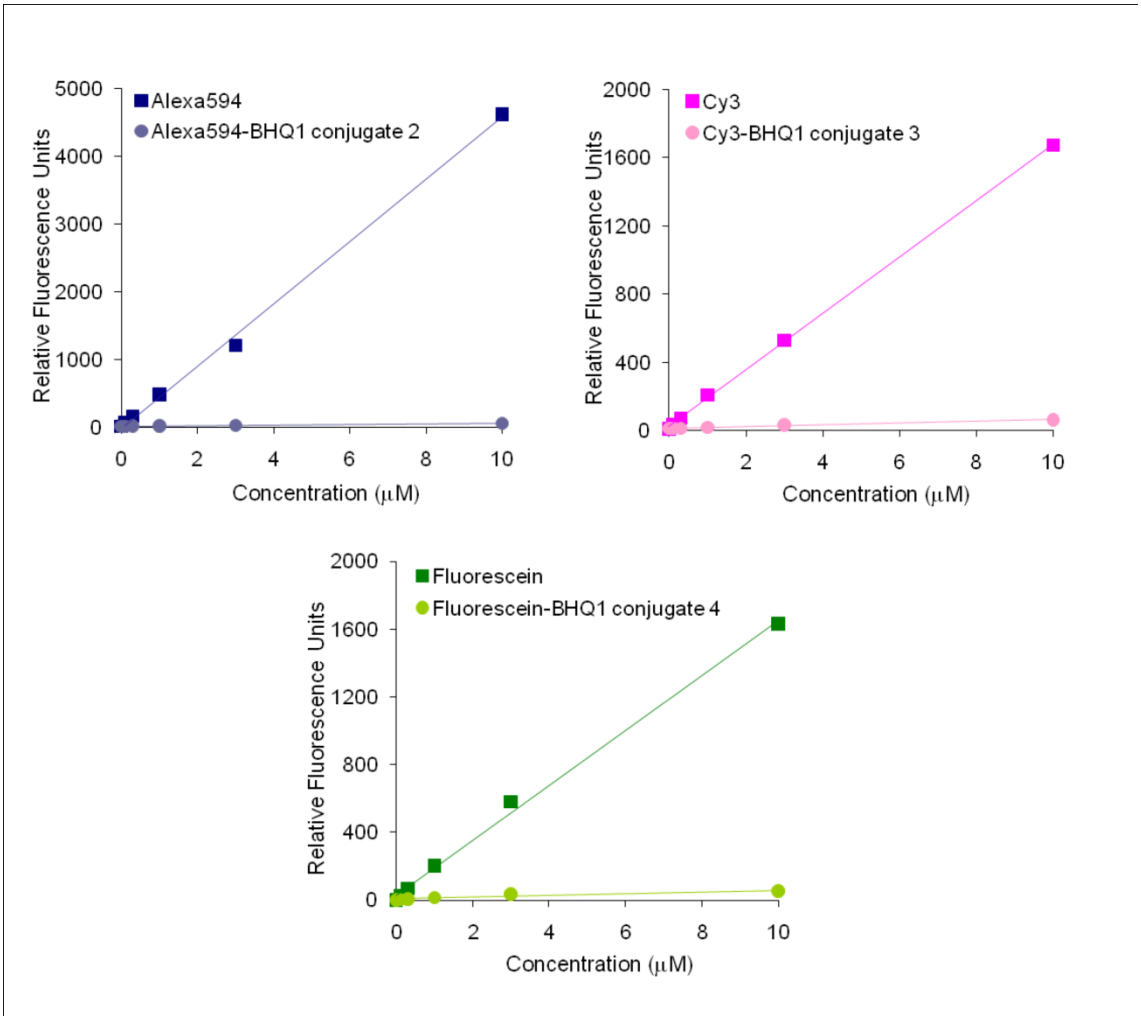


Fig. S1. Determination of quenching efficiency of the fluorophore-BHQ1 conjugates relative to unquenched control fluorophores. The probes were dissolved in 10 mM Tris-HCl buffer (pH7.5)-methanol (7:3, v/v), and fluorescenece intensities were measured using a fluorescence plate reader.

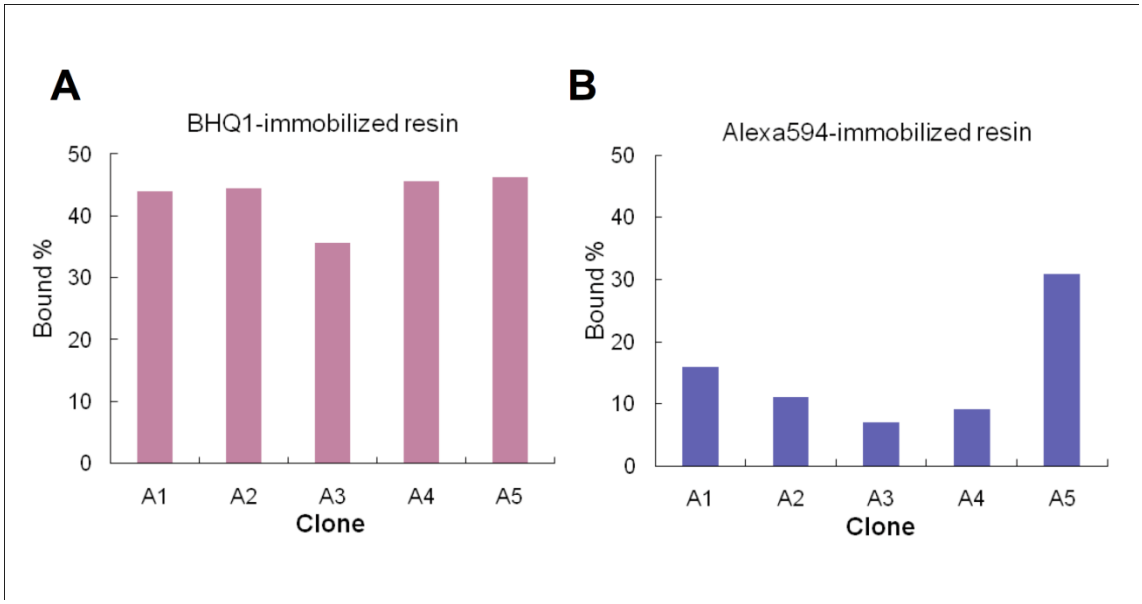


Fig. S2. Percentage of RNA bound to Alexa594-immobilized or BHQ1-immobilized resin.

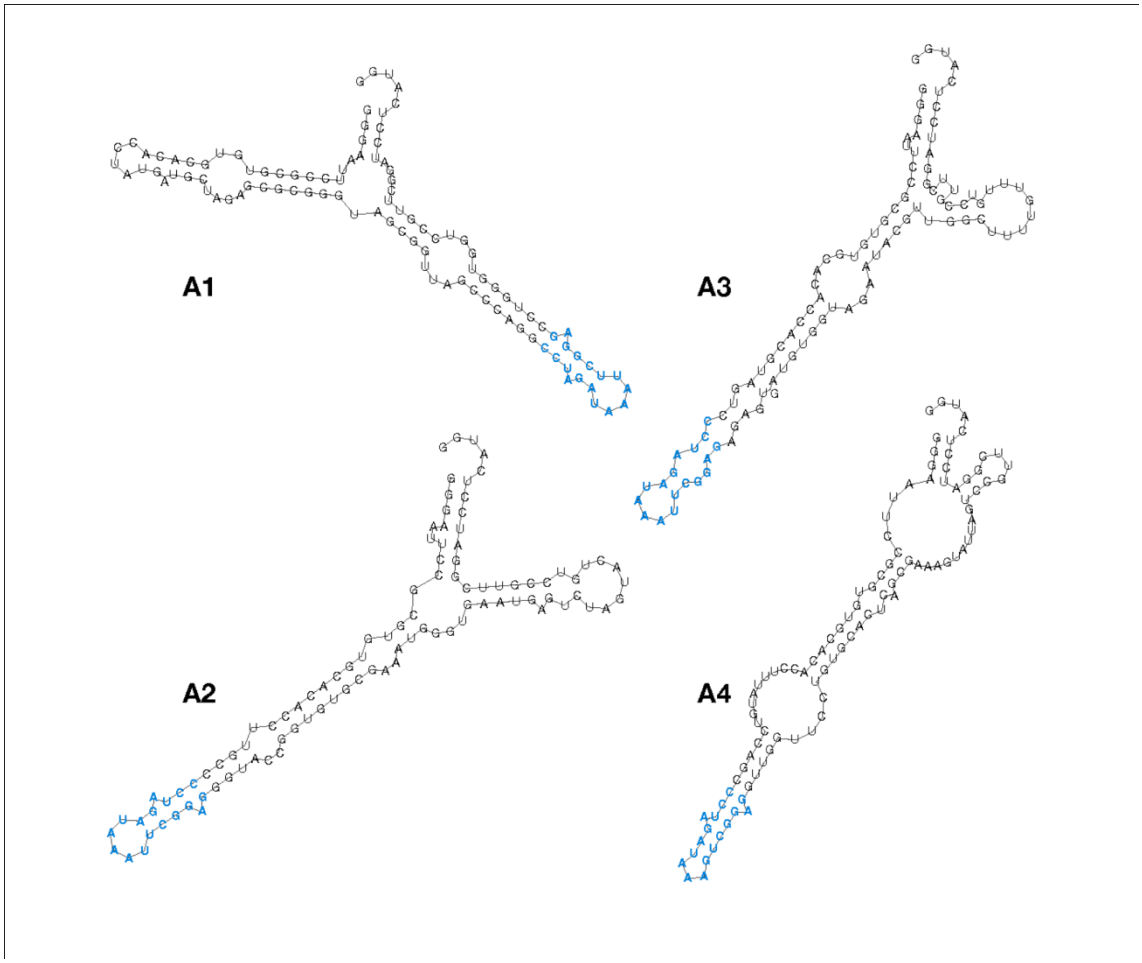


Fig. S3. Secondary structures of the aptamers, A1-A4, predicted by the Centroid fold program (http://www.ncrna.org/software/centroidfold/).

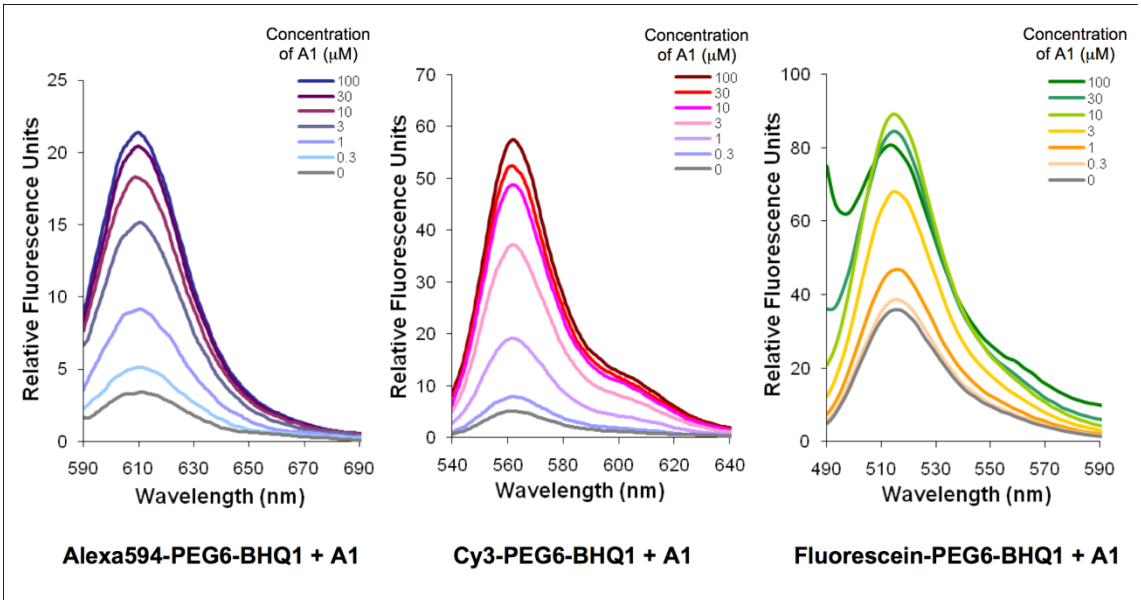


Fig. S4. Fluorescence spectra of fluorophore-BHQ1 conjugates in the presence of A1. Fluorescence was restored by A1 in a concentration-dependent manner, up to 2.5-fold for fluorescein-PEG6-BHQ1, 11.3-fold for Cy3-PEG6-BHQ1, and 6.3-fold for Alexa594-PEG6-BHQ1.

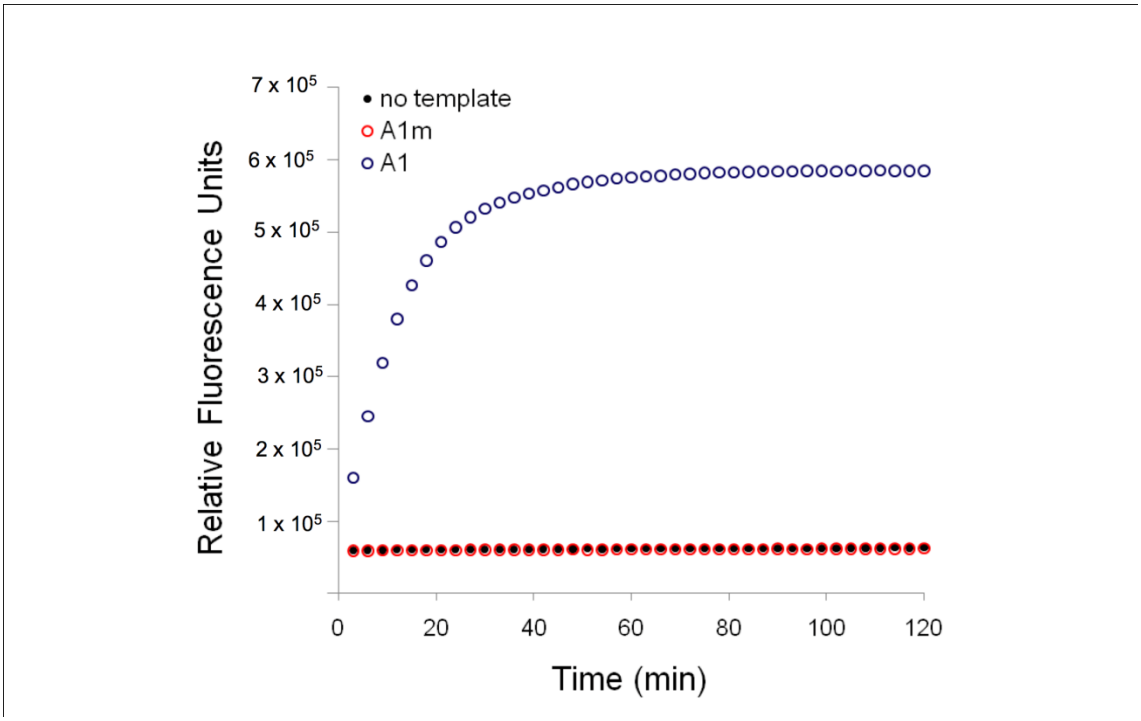


Fig. S5. Real-time monitoring of transcripts corresponding to A1 or A1m, using conjugate **3**.

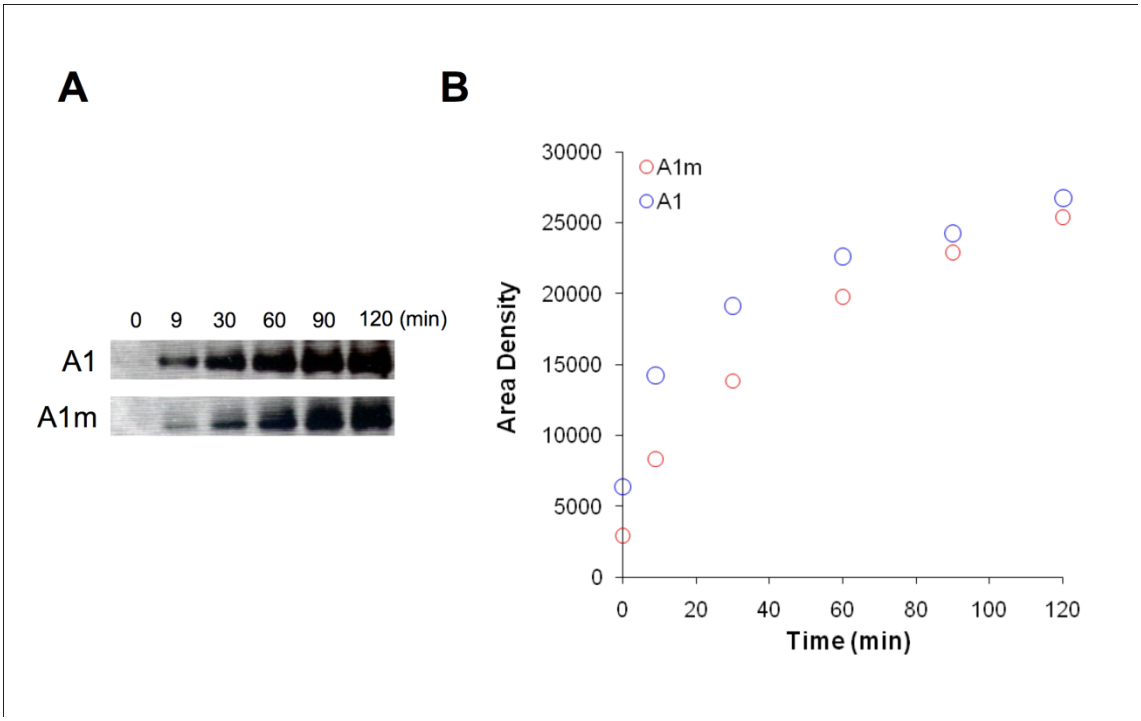


Fig. S6. (A) Time-dependent increase of transcripts synthesized by T7 RNA polymerase from DNA templates encoding A1 or A1m. Gels were stained with ethidium bromide. (B) Densitometric analysis of transcript levels at indicated time points.

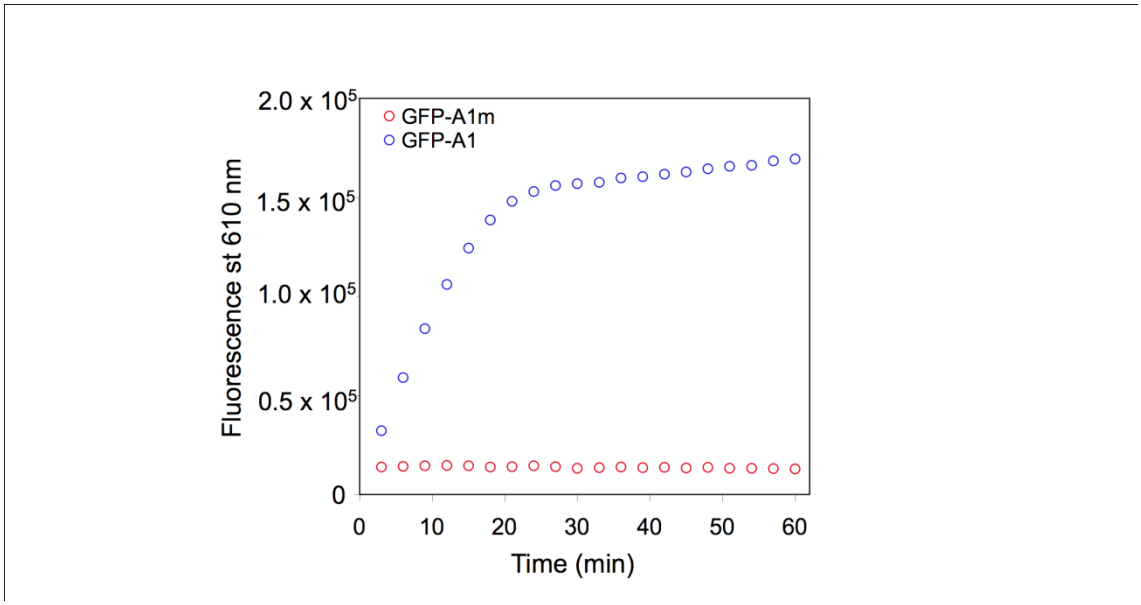


Fig. S7. Real-time monitoring of transcripts from pIVEX-GFP-A1 or pIVEX-GFP-A1m plasmids, using conjugate **2**.

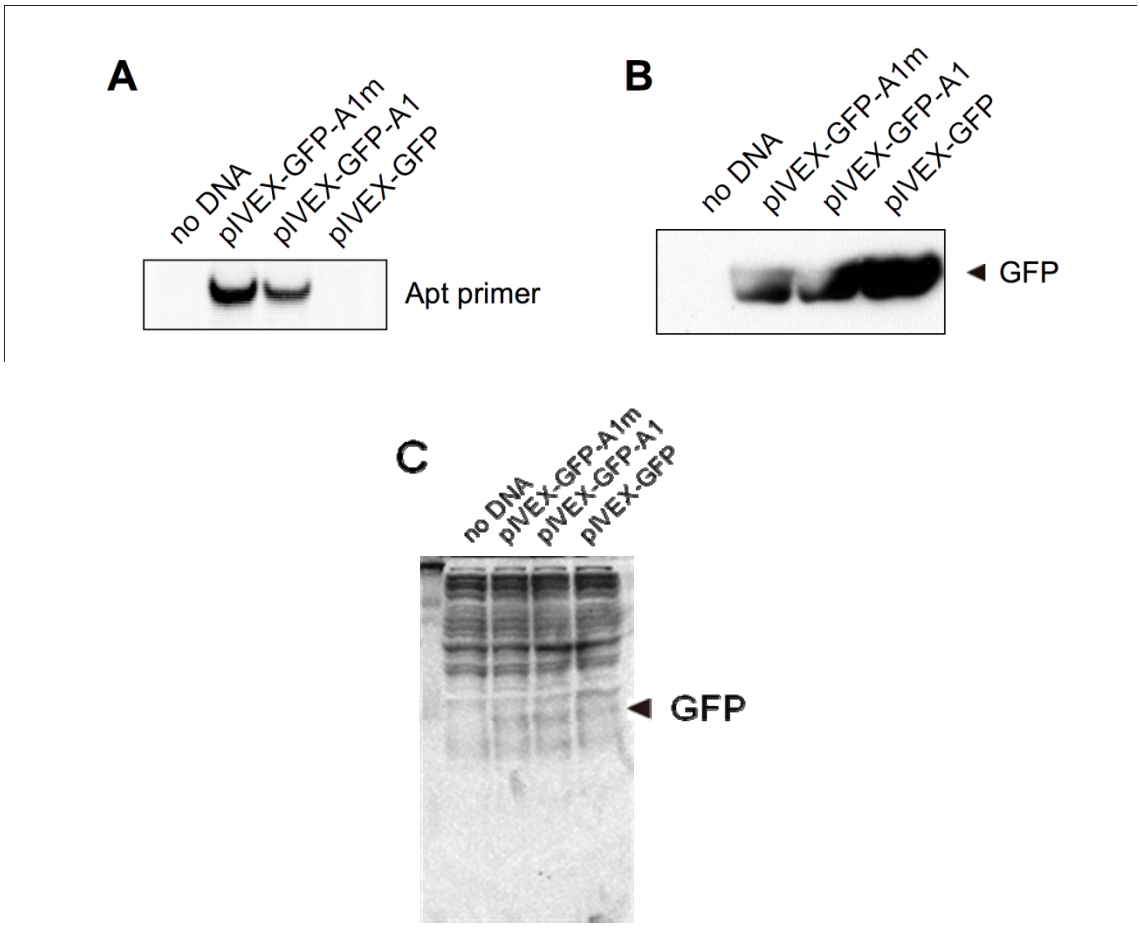


Fig. S8. Verification of RNA and protein synthesis from pIVEX-GFP-A1 and pIVEX-GFP-A1m plasmids. (A) RT-PCR analysis of the transcript from pIVEX-GFP-A1 or pIVEX-GFP-A1m plasmids, using the primer set for aptamer sequence. (B) Western blot analysis of GFP expression, using anti-His antibody. (C) The protein band staining with Coomassie brilliant blue dye (CBB) of the same samples used in (B). The loading amounts of proteins were essentially the same.

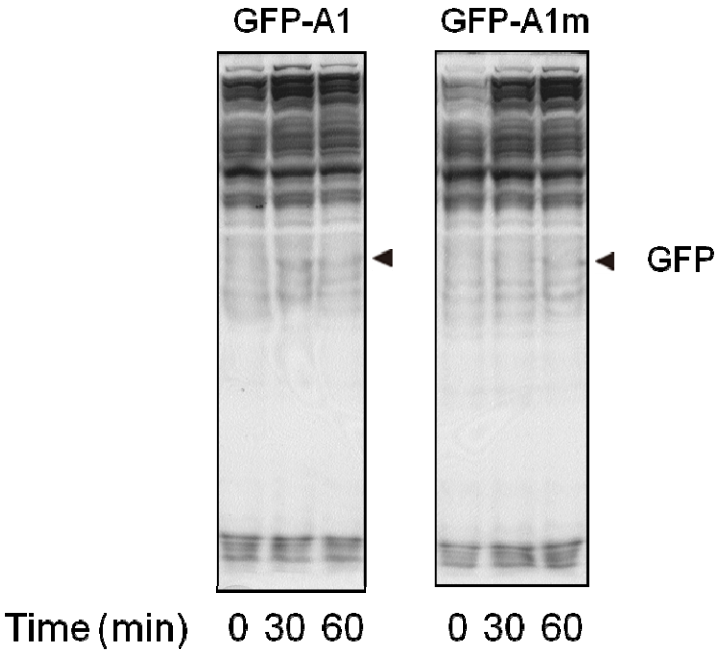


Fig. S9. The protein band staining with Coomassie brilliant blue dye (CBB) of the same samples used in Fig.6C. The loading amounts of proteins were essentially the same.