Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Recognition-then-Reaction Enables Site-Selective Bioconjugation to Proteins on Live-Cell Surfaces

Cheng Cui, Hui Zhang, Ruowen Wang, Sena Cansiz, Xiaoshu Pan, Shuo Wan, Weijia Hou, Long Li, Meiwan Chen, Yuan Liu, Xigao Chen, Qiaoling Liu, and Weihong Tan**

anie_201706285_sm_miscellaneous_information.pdf

SUPPORTING INFORMATION

Table of Contents

Experimental Procedures

Materials

Unless otherwise stated, all reagents used were purchased from Sigma-Aldrich and Thermo Fisher Scientific. DNA synthesis reagents were acquired from Glen Research Corp. and ChemGenes Corp. All oligonucleotides were synthesized and purified in our group, except human IgG Fc aptamer-3', which was purchased from Integrated DNA Technologies. Proteins were purchased from Haematologic Technologies, Inc., Thermo Fisher Scientific, ProSpec-Tany TechnoGene Ltd., R&D Systems, Inc., and Scripps Laboratories. The water used was purified on a Milli-Q Biocell System. SDS-PAGE analysis was performed on precast NuPAGE® Novex 4-12% bis-tris gels according to the manufacturer's protocol and stained with SYBR® Gold and GelCode™ Blue Safe Protein Stain.

DNA synthesis and purification

All oligonucleotides were synthesized based on solid-phase phosphoramidite chemistry at the 1 μmol scale using an ABI3400 DNA/RNA synthesizer (Applied Biosystems). The synthesis and deprotection procedures were performed using the instructions provided by the reagent's manufacturer (Glen Research Corp). Subsequently, the deprotected DNA was precipitated by adding 1/10 volume of 3 M acetate buffer (pH=5.5) and 2.5 times volume of cold ethanol. After placing in a freezer at -20 °C for 30 minutes, the DNA precipitate was collected by centrifugation at 4,000 rpm for 30 minutes. The DNAs were then dissolved in 400 μL of 0.1 M triethylamine acetate (TEAA, Glen Research) and purified by a ProStar HPLC (Varian) with a C18 column (Econosil, 5 μm, 250 × 4.6 mm) using acetonitrile and 0.1 M TEAA aqueous solution as the mobile phase. Finally, the collected sequences were vacuum-dried and quantified at 260 nm using a Cary Bio-300 UV spectrometer (Varian). The detailed sequence information is described in Table S1.

Conjugation of DSS linker to amino-modified DNA to generate RO (3' or 5' NHS ester-modified DNA)

To the solution of amino-modified DNA (21.6 nmol) in water (108 μL) were added disuccinimidyl suberate (DSS, 2 mg, 250 eq.) in DMF (108 μL), acetonitrile (108 μL) and TEA (1.5 μL). After shaking at room temperature for 30 minutes, the DNA was precipitated by adding 3 M acetate buffer (33 μL, pH=5.5), cold ethanol (814 μL) and glycogen (20 mg/mL, 1.5 μL). After placing in a freezer at -80 °C for 20 minutes, the DNA precipitate was collected by centrifugation for 60 minutes (4 °C, 20,000 g). The pellet was then dissolved in 200 μL of 0.1 M TEAA and purified by RP-HPLC (0-30 % acetonitrile in 0.1 M TEAA over 40 minutes, flow rate=1 mL/min). To stabilize the separated NHS ester-modified DNA, an equivalent amount of a 2% (v/v) trifluoroacetic acid (TFA) solution was added to decrease the hydrolysis rate of NHS ester. The DNA was distributed equally into 20 Eppendorf® tubes (approximately 400-800 pmol in each tube; 200-400 pmol for FITC-labeled DNA). Finally, the mixture was vacuum-dried, and the concentration of DNA in each tube was quantified at 260 nm using UV spectrometry, assuming that DSS conjugation did not affect the extinction coefficient at 260 nm.

DNA-protein conjugation reaction procedure

For all protein/DNA conjugation reactions, nearly identical experimental conditions were used. The reactions were performed in LoBind Eppendorf tubes to reduce the surface adhesion of DNA. To better monitor the ATS in protein conjugation, most reactions were analyzed with PAGE and stained for DNA and protein, respectively. Unless stated otherwise, the results of nucleic acid staining with SYBR® Gold and protein staining with GelCode™ Blue were obtained from the same gel.

Preparation of conjugation

RO DNA was stored in TFA to inhibit the hydrolysis of NHS ester residue. Before use, the DNA was neutralized immediately with an aqueous TEA solution (100 mM), 1μL of MES buffer (500 mM, pH=6.0) and water to give pH=6.0 with a final concentration of 5 μM DNA. All proteins were used without any further purification. Protein aliquots were diluted in DPBS buffer (2.67 mM KCl, 138 mM NaCl, 1.47 mM KH_2PO4 and 8.06 mM Na_2HPO4, pH = 7.0–7.3, Thermo Fisher Scientific) to a final concentration of 5μM protein.

Conjugation reaction procedure

Aptamer-guided conjugations were performed on 20 μL to 2 mL scales. The reaction was carried out by mixing protein (0.25 μM), NaCl (400 mM), Tween-20 (0.02%, v/v), aptamer template and RO, with concentrations varying from 0.9 to 2.0 equivalents, to the protein for different targets (detailed numbers given in corresponding figures), in HEPPS buffer (50 mM, pH=7.7) at room temperature

overnight. Then the reaction was quenched by adding 1.0 M glycine solution to reach a final glycine concentration of 50 mM. For reactions with large volume, the mixture was concentrated using Amicon Ultra centrifugal filters (MWCO 3K, spin filtration at 14,000 g for 30 minutes) before gel analysis/purification. RNA aptamer was used; thus, for conjugation reaction on IgG Fc fragment, all materials were autoclaved before reaction to remove ribonuclease.

Template replacement

After reaction, the aptamer template was replaced by adding the complementary-sequenced DNA of the template (2 equivalents) in HEPPS buffer (50 mM, pH=7.7). The replacement was performed for 30 minutes before concentration (spin filtration) and gel analysis.

Trypsin digestion

To demonstrate site selectivity of the reaction, the reaction product was degraded by trypsin (Thermo Fisher Scientific) with equal reaction volume added. The digestion was carried out at 37 °C for 30 minutes, followed by concentration with spin filtration before SDS-PAGE gel analysis.

In gel digestion for mass spectral analysis

After trypsin degradation, gels with target band were digested with sequencing grade trypsin from Promega, using the manufacturer's recommended protocol. Briefly, bands were trimmed as close as possible to minimize background polyacrylamide material. Gel pieces were then washed in nanopure water three times for 5 minutes each. The gel pieces were then washed and/or destained with 1:1 (v/v) methanol:50 mM ammonium bicarbonate twice for 10 minutes. The gel pieces were dehydrated with 1:1 (v/v) acetonitrile: 50 mM ammonium bicarbonate. The gel bands were rehydrated and incubated with dithiothreitol (DTT) solution (25 mM in 100 mM ammonium bicarbonate) for 30 minutes prior to the addition of 55 mM iodoacetamide in 100 mM ammonium bicarbonate solution. Iodoacetamide was incubated with gel bands in the dark for 30 minutes before removal. Gel bands were washed again with two cycles of water and dehydrated with 1:1 (v/v) acetonitrile:50 mM ammonium bicarbonate. The protease was driven into the gel pieces by rehydrating them in 12 ng/mL trypsin in 0.01% ProteaseMAX Surfactant for 5 minutes. The gel pieces were then overlaid with 40 µL of 0.01% ProteaseMAX Surfactant:50 mM ammonium bicarbonate and gently mixed on a shaker for 1 hour. The digestion was stopped by addition of 0.5% TFA. Mass spectral analysis was immediately performed to ensure high-quality tryptic peptides with minimal nonspecific cleavage, or they were frozen at -80 °C until samples could be analyzed.

Experimental conditions, design, and calculations

For aptamer-guided synthesis, key parameters are concentrations of protein and aptamer template. Excessively high concentrations would lead to nonspecific reaction, while insufficient concentrations would cause low binding efficiency between the protein and aptamer. Here, thrombin and aptamer HD22 are used to calculate reaction conditions. The binding reaction is shown as

$$\text{thrombin} + \text{HD22} \rightleftharpoons \text{thrombin} \bullet \text{HD22}$$. Equation (1)

The equilibrium constant of this reaction is

$$\mathrm{K} = \frac{1}{\mathrm{K_d}} = \frac{[\text{thrombin}\bullet\text{HD22}]}{[\text{thrombin}][\text{HD22}]}$$. Equation (2)

For aptamer-guided conjugation, 0.25 µM thrombin and 0.50 µM HD22 template were used. Assuming that the linker would not affect the binding efficiency of aptamer and that x µM thrombin binds to HD22 template, we have (K_d = 0.7 nM):

$$\frac{1}{0.7\times10^{-3}} = \frac{x}{(0.25-x)(0.5-x)}$$, Equation (3)

$$x = 0.249\ \mu\text{M}$$

where x is the equilibrium concentration of the thrombin-aptamer complex. According to this result, more than 99% of thrombin binds to aptamer at this condition. For thrombin and aptamer TBA, x=0.115 µM at the given conditions. This is the result of a high K_d of 450 nM, and the binding ratio between thrombin and aptamer TBA would decrease to 46%. Thus, the theoretical yield could be predicted by K_d of the aptamer and initial concentrations of the protein and aptamer.

SDS-PAGE gel analysis

After conjugation reaction and spin filtration, if necessary, NuPAGE LDS Sample Buffer (4X) was added, and the mixture was heated to 95 °C for 5 minutes to achieve complete denaturing. Electrophoresis was performed on a precast NuPAGE® Novex 4-12% bis-tris gel, according to the manufacturer's protocol, with Fisher BioReagents exACTGene DNA Ladder and PageRuler Plus Prestained Protein Ladder, and stained with SYBR® Gold visualized by Typhoon scanner and/or GelCode™ Blue Safe Protein Stain (Thermo Fisher Scientific).

SUPPORTING INFORMATION

Non-denaturing PAGE gel analysis
It is first necessary to wash the sample with water before loading to decrease the salt concentration. Then, after conjugation and spin filtration, MassRuler DNA Loading Dye (6X) was added to the sample mixture. The sample was loaded on the gel directly, with Fisher BioReagents exACTGene DNA Ladder and NativeMark Unstained Protein Standard, and run at 3 Watts for 80 minutes (5% PAGE).

Non-denaturing 5% gel recipe (to make 20 mL)
Water was added to acrylamide/bis-acrylamide (30%, 3.33 mL), APS (10%, 200 µL) and TBE buffer (10X, 2.0 mL) mixture to make a total volume of 20 mL. Then TEMED (20 µL) was added, and the mixture was gently shaken before use.

Non-denaturing gel extraction
After aptamer template replacement and concentration, the sample was loaded, and electrophoresis was carried out. After visualization with a B-BOX Blue Light LED epi-illuminator (SMOBIO Technology, Inc.), the product band was cut using a scalpel and extracted in DPBS buffer overnight with shaking. The extracted conjugate was analyzed by non-denaturing PAGE, and the extract yield was determined in ImageJ by comparing the SYBR® Gold absorbance of the conjugate band to a series of RO with different concentrations.

LTQ mass spectrometry
Nano-liquid chromatography tandem mass spectrometry (Nano-LC/MS/MS) was performed on a Thermo Scientific LTQ-XL mass spectrometer equipped with an EASY Spray nanospray source (Thermo Scientific) operated in positive ion mode. The LC system was an EASY Nano-LC II from Thermo Scientific. Mobile phase A consisted of water containing 0.1% formic acid, and mobile phase B consisted of acetonitrile with 0.1% formic acid. Five µL of each sample was first injected onto a Thermo Scientific Acclaim PepMap 100 Trap Cartridge (C18 column, 75 µm ID, 2 cm length, 3 µm 100 Å pore size) and then washed with mobile phase A to desalt and concentrate the peptides. The injector port was switched to inject, and the peptides were eluted off the trap onto the column. An EASY Spray PepMAP column from Thermo Scientific was used for chromatographic separations (C18, 75 µm ID, 15 cm length, 3 µm 100 Å pore size). The column temperature was maintained at 35 °C, and peptides were eluted directly off the column into the LTQ system using a gradient of 2-80% B over 45 minutes, with a flow rate of 300 nL/min. The total run time was 60 minutes. MS/MS was acquired according to standard conditions established in the lab. The EASY Spray source operated with a spray voltage of 1.5 KV and a capillary temperature of 200 °C. The scan sequence of the mass spectrometer was based on the TopTen™ method; the analysis was programmed for a full scan recorded between 350 and 2,000 Da. To determine the amino acid sequence, MS/MS scans were collected to generate product ion spectra to determine amino acid sequence in consecutive instrument scans of the ten most abundant peaks in the spectrum. The AGC Target ion number was set at 30,000 ions for full scan and 10,000 ions for MSn mode. Maximum ion injection time was set at 20 ms for full scan and 300 ms for MSn mode. Micro scan number was set at 1 for both full scan and MSn scan. The CID fragmentation energy was set to 35%. Dynamic exclusion was enabled with a repeat count of 1 within 10 seconds, a mass list size of 200, and exclusion duration of 350 seconds. The low mass width was 0.5, and the high mass width was 1.5.
Sequence information from MS/MS data was processed by converting the .raw files into a merged file (.mgf) using msConvert (ProteoWizard) or Mascot Distiller (Matrix Science). The resulting .mgf files were searched using Mascot Daemon by Matrix Science, version 2.4.0, and the database was searched against the full SwissProt database, version 2015_08 (549,008 sequences; 195,692,017 residues). The mass accuracy of the precursor ions was set to 0.8 Da, and the fragment mass accuracy was set to 0.5 Da. Considered variable modifications were methionine oxidation and deamidation NQ. Fixed modification for carbamidomethyl cysteine was considered. One missed cleavage for the enzyme was permitted. A decoy database was searched to determine the false discovery rate (FDR). Peptides were filtered according to the FDR, and the proteins identified required bold red peptides. Protein identifications were checked manually, and proteins with a Mascot significance threshold $p < 0.05$ with a minimum of two unique peptides from one protein having a -b or -y ion sequence tag of five residues, or better, were accepted.

Cell culture
CCRF-CEM (CCL-119 T-cell, human acute lymphoblastic leukemia) and Ramos cells (CRL-1596, B lymphocyte, human Burkitt's lymphoma) were obtained from ATCC (American Type Culture Collection) and were cultured in complete RPMI 1640 medium (ATCC) supplemented with 10% fetal bovine serum (FBS) (heat inactivated, GIBCO) and 100 IU/mL penicillin−streptomycin (Cellgro). The washing buffer contained 4.5 g/L glucose and 5 mM $MgCl_2$ in Dulbecco's PBS (Sigma). Binding buffer used for incubation was prepared by adding yeast tRNA (0.1 mg/mL) (Sigma) and BSA (1 mg/mL) (Fisher) to the washing buffer to reduce background binding.

Cellular conjugation and toehold replacement
After washing with 2 mL washing buffer three times and then removing buffer by centrifugation at 300 g for 3 minutes, 500K CEM/Ramos cells were incubated in 250 µL binding buffer with 50 mM NaCl, 250 nM Sgc8-3' template and 250 nM FITC-modified 3RO for 90 minutes at room temperature. Then 25 µL cDNA of 3RO (25 µM) were added to the mixture and incubated for 30 minutes. Finally, the treated cells were washed three times with 2 mL of washing buffer and analyzed by flow cytometry.

Flow cytometry analysis

SUPPORTING INFORMATION

To monitor the conjugation reaction between 3RO and cell membrane protein, FITC-labeled 3RO was used. Cells were washed and suspended in 0.3 mL washing buffer. The fluorescence of FITC was determined with a BD Accuri C6 cytometer (Becton, Dickinson and Company) by counting 10,000 events. Negative control groups were prepared, as shown in Fig. S18.

Results and Discussion

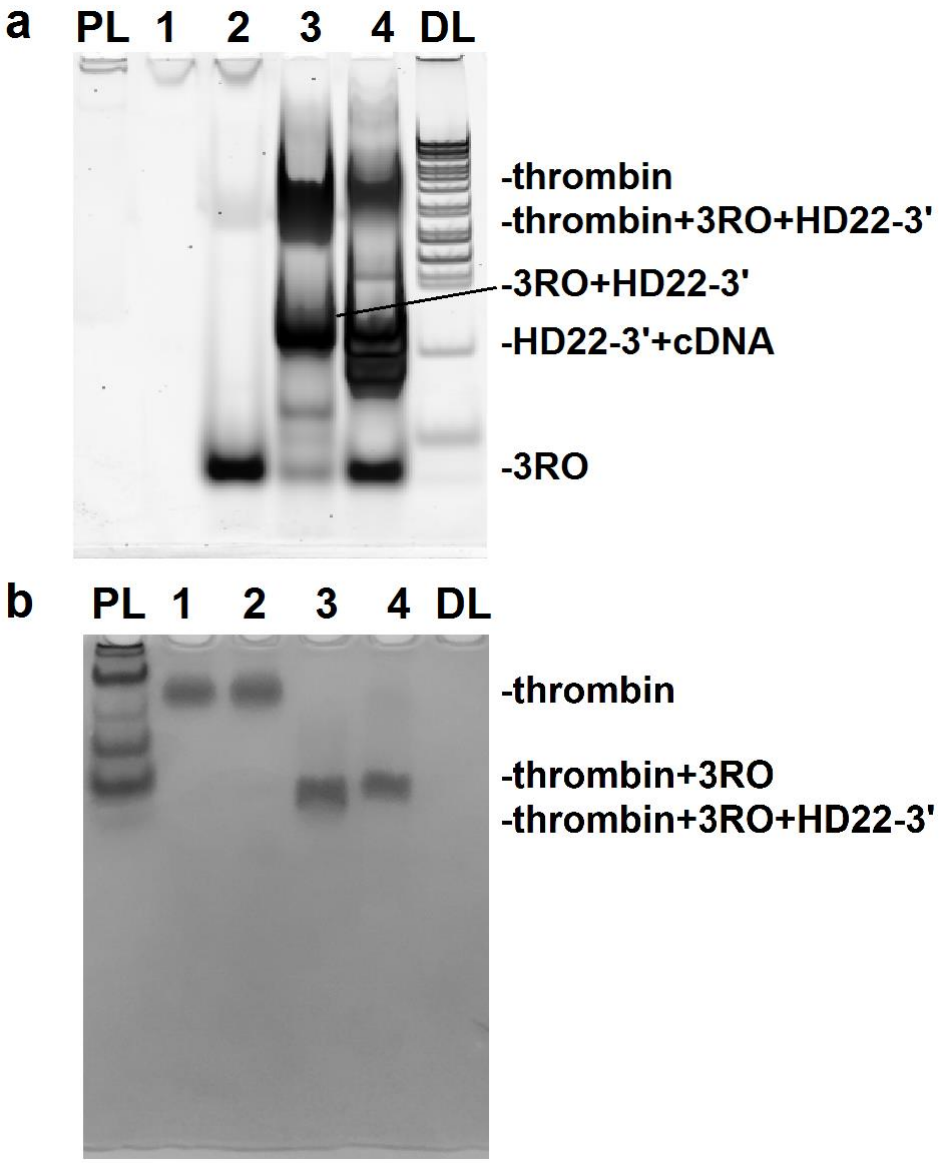


Figure S1. Non-denaturing PAGE (5%) showing aptamer-templated conjugation on thrombin. a, Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin; lane 2, thrombin and 3RO (2 eq.); lane 3, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); lane 4, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), with cDNA of HD22-3' replacement (4 eq.) after conjugation. PL, NativeMark™ Unstained Protein Standard (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, The gel is stained with GelCode™ Blue for protein with same lane 1-4.

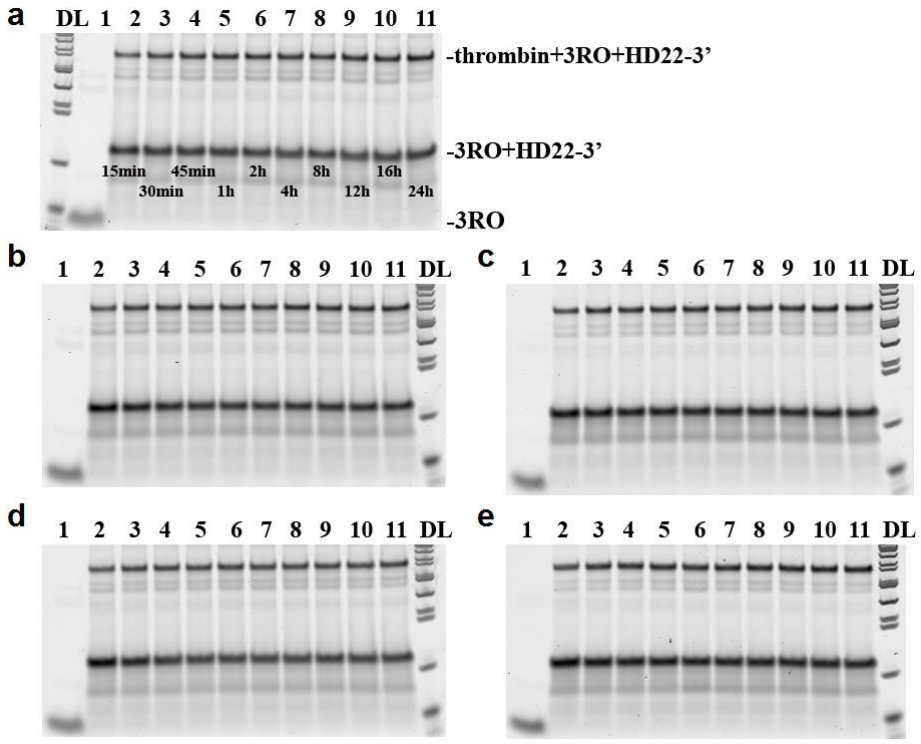


Figure S2. SDS PAGE (4-12%) showing the kinetics of site-selective conjugate formation. To investigate the rate of conjugate formation, ATS on thrombin with template HD22-3' was performed using different reaction times and analyzed by SDS- PAGE, as shown in Figure S2. Analysis indicates that 1 hour reaction time would be sufficient for optimal conversion (analysis results shown in Figure 1c). ATS of thrombin and HD22-3' with different reaction times was repeated five times and generated reproducible results. Lane 1, thrombin and 3RO (2 eq.), reaction time: 24 hours; lane 2-11, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); reaction times were 15 minutes, 30 minutes, 45 minutes, 1 hour, 2 hours, 4 hours, 8 hours, 12 hours, 16 hours and 24 hours, respectively. DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). All gels are stained with SYBR® Gold for DNA.

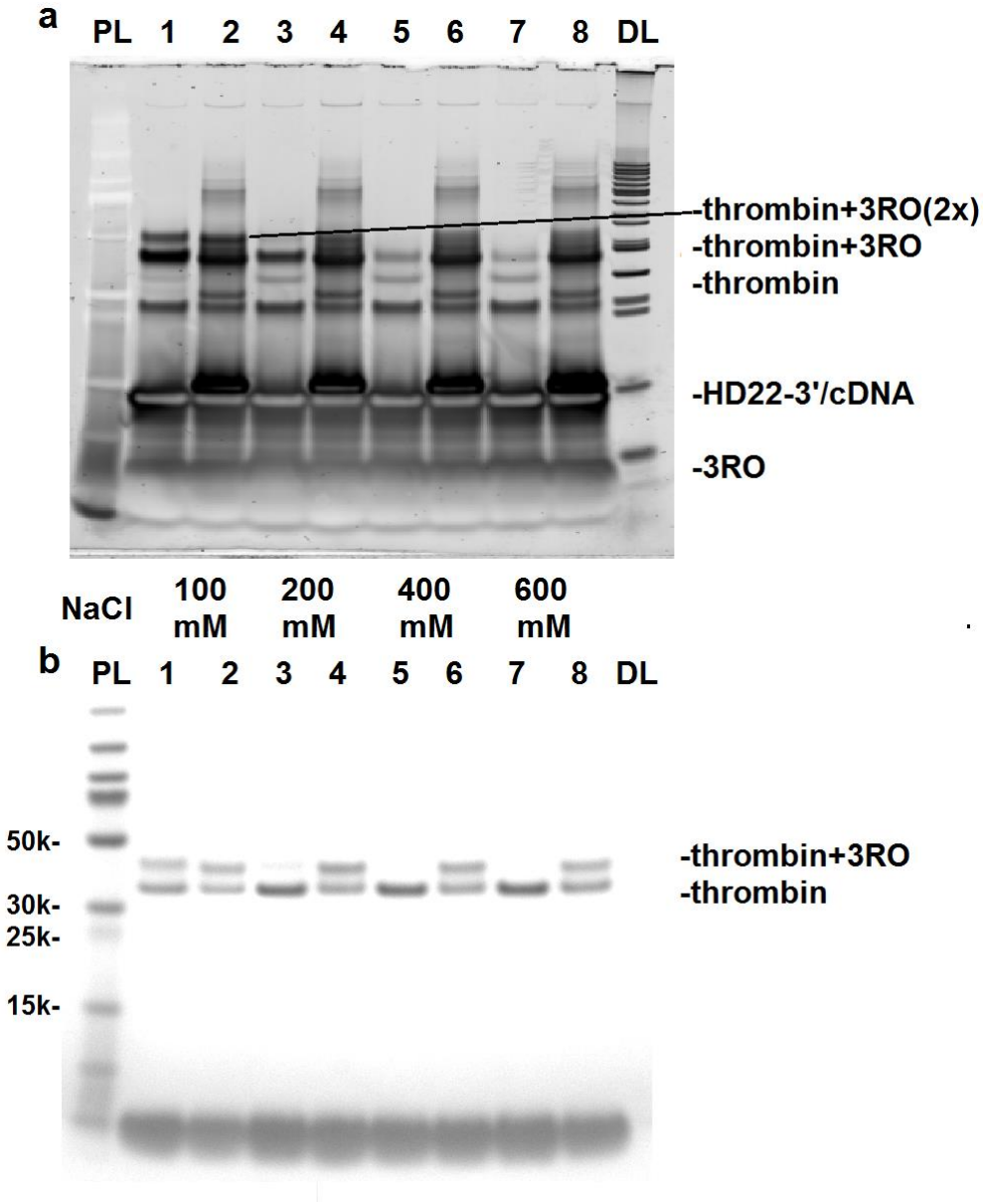


Figure S3. Salt dependency of ATS on thrombin. From the SDS-PAGE (4-12%) results, only negligible nonspecific reaction occurs between 3RO and thrombin at 400mM or higher NaCl concentration. **a**, Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin and 3RO (2 eq.) in 100 mM NaCl; lane 2, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.) in 100 mM NaCl; lane 3, thrombin and 3RO (2 eq.) in 200 mM NaCl; lane 4, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.) in 200 mM NaCl; lane 5, thrombin and 3RO (2 eq.) in 400 mM NaCl; lane 6, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.) in 400 mM NaCl; lane 7, thrombin and 3RO (2 eq.) in 600 mM NaCl; lane 8, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.) in 600 mM NaCl. PL, NativeMark™ Unstained Protein Standard (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, Same gel is stained with GelCode™ Blue for protein.

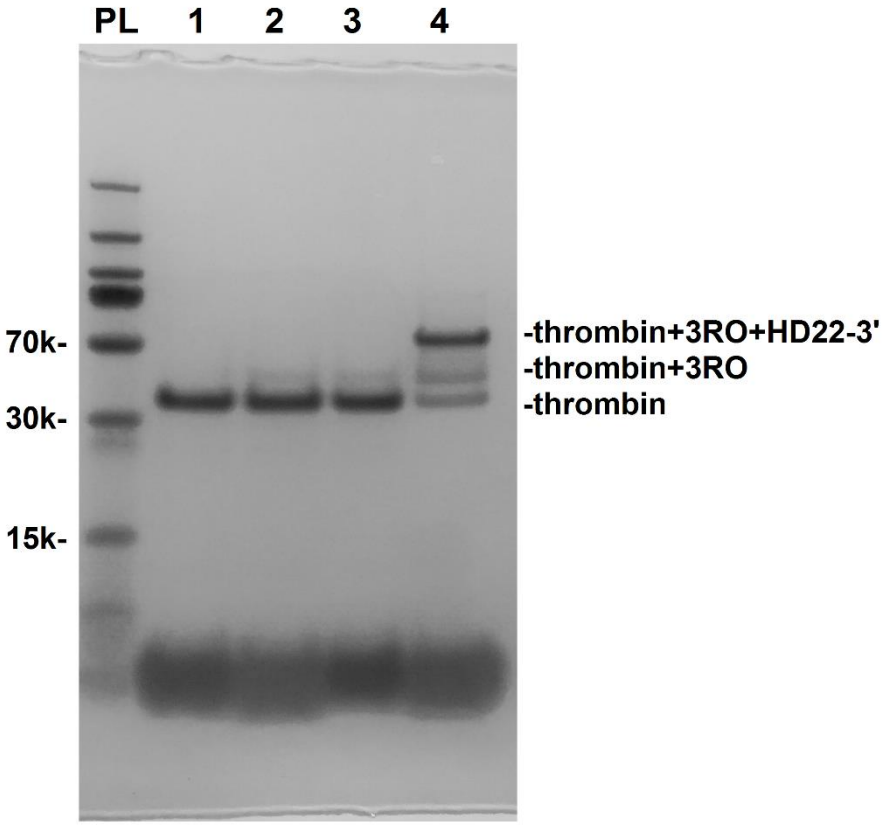


Figure S4. SDS PAGE (4-12%) with protein staining showing the necessity of using aptamer template in thrombin conjugation. Lane 1, thrombin; lane 2, thrombin, 3RO (2 eq.) and template containing HD22 aptamer and random sequence in linker (2 eq.); lane 3, thrombin, 3RO (2 eq.) and template containing a linker and random sequence in aptamer part (same length as HD-22, 2 eq.); lane 4, thrombin, 3RO (2 eq.), and HD22-3' (2 eq.). PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). Gel is stained with GelCode™ Blue for protein.

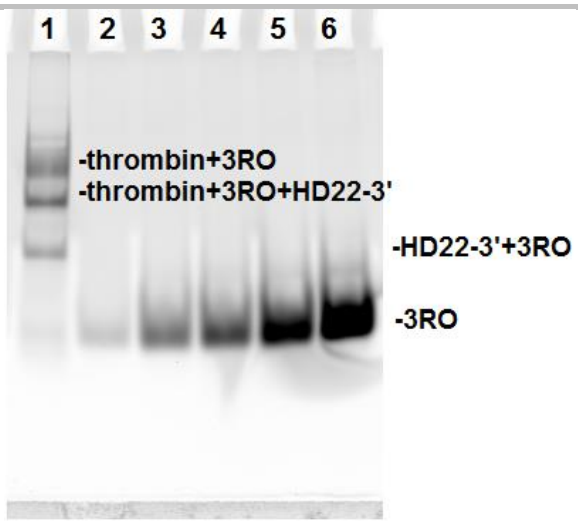


Figure S5. Non-denaturing PAGE (5%) for analysis of conjugate extract yield from gel. Lane 1, thrombin/3RO-FITC conjugate extracted from another non-denaturing PAGE gel at reaction scale of 600 µL; lane 2, 3RO-FITC (1 pmol); lane 3, 3RO-FITC (5 pmol); lane 4, 3RO-FITC (10 pmol); lane 5, 3RO-FITC (20 pmol); lane 6, 3RO-FITC (50 pmol). The band signals are from the FITC labeled on 3RO. The intensity of the 3RO-FITC was analyzed with ImageJ. By comparing to a calibration plot of absorbance *vs*. free 3RO concentration, the extract yield of conjugate was determined to be 55%.

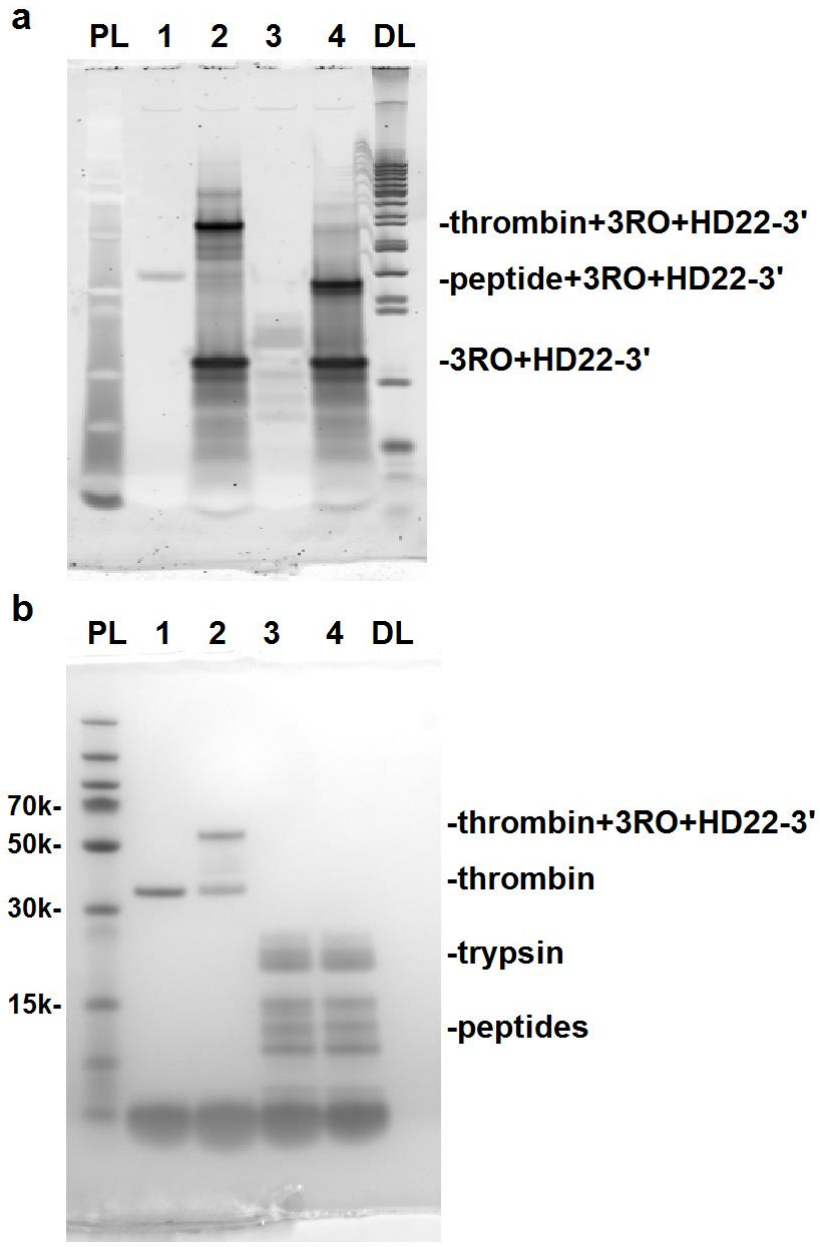


Figure S6. SDS-PAGE (4-12%) for analysis of trypsin digestion products. a, Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin; lane 2, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); lane 3, thrombin after trypsin digestion for 30 minutes; lane 4, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.) after trypsin digestion for 30 minutes after conjugation. PL, NativeMark™ Unstained Protein Standard (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, Same gel is stained with GelCode™ Blue for protein. One clear band can be observed in **a** after digestion (lane 4) showing the unique conjugation position of ATS.

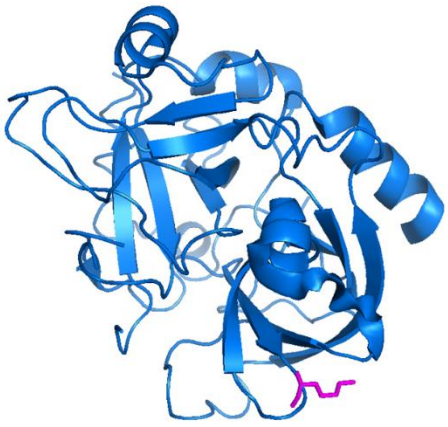
a

Query	Start - End	Observed	Mr(expt)	Mr(calc)	Delta	M	Score	Expect	Rank	U	Peptide
1416	364 - 383	755.4200	2263.2382	2264.0970	-0.8588	0	82	6.5e-006	1	U	R.IVEGSDAEIGMSPWQVMLFR.K
1417	364 - 383	755.5200	2263.5382	2264.0970	-0.5588	0	74	1.3e-005	1	U	R.IVEGSDAEIGMSPWQVMLFR.K
1982	364 - 383	1133.3900	2264.7654	2264.0970	0.6685	0	105	1e-008	1	U	R.IVEGSDAEIGMSPWQVMLFR.K
1419	364 - 383	755.9300	2264.7682	2264.0970	0.6712	0	63	0.00038	1	U	R.IVEGSDAEIGMSPWQVMLFR.K
1983	364 - 383	1133.4700	2264.9254	2264.0970	0.8285	0	97	7.3e-008	1	U	R.IVEGSDAEIGMSPWQVMLFR.K
1420	364 - 383	755.9900	2264.9482	2265.0810	-0.1328	0	42	0.039	1	U	R.IVEGSDAEIGMSPWQVMLFR.K + Deamidated (NQ)
1984	364 - 383	1133.9400	2265.8654	2265.0810	0.7845	0	82	7.2e-007	1	U	R.IVEGSDAEIGMSPWQVMLFR.K + Deamidated (NQ)
1421	364 - 383	756.3600	2266.0582	2265.0810	0.9772	0	77	2e-005	1	U	R.IVEGSDAEIGMSPWQVMLFR.K + Deamidated (NQ)
1422	364 - 383	756.4300	2266.2682	2265.0810	1.1872	0	81	2.8e-006	1	U	R.IVEGSDAEIGMSPWQVMLFR.K + Deamidated (NQ)
1520	364 - 384	798.3100	2391.9082	2392.1919	-0.2837	1	61	0.00076	1	U	R.IVEGSDAEIGMSPWQVMLFRK.S
1523	364 - 384	798.8100	2393.4082	2393.1759	0.2323	1	67	0.00012	1	U	R.IVEGSDAEIGMSPWQVMLFRK.S + Deamidated (NQ)
1524	364 - 384	799.1100	2394.3082	2393.1759	1.1323	1	51	0.0087	1	U	R.IVEGSDAEIGMSPWQVMLFRK.S + Deamidated (NQ)
1545	364 - 384	804.2500	2409.7282	2409.1708	0.5573	1	44	0.008	1	U	R.IVEGSDAEIGMSPWQVMLFRK.S + Deamidated (NQ); Oxidation (M)
1072	384 - 399	592.0400	1773.0982	1772.9091	0.1891	1	61	0.007	1	U	R.KSPQELLCGASLISDR.W
1073	384 - 399	592.4900	1774.4482	1773.8931	0.5551	1	43	0.045	1	U	R.KSPQELLCGASLISDR.W + Deamidated (NQ)
954	385 - 399	549.5000	1645.4782	1644.8141	0.6641	0	54	0.0012	1	U	K.SPQELLCGASLISDR.W
955	385 - 399	549.5100	1645.5082	1644.8141	0.6941	0	49	0.0074	1	U	K.SPQELLCGASLISDR.W
1570	385 - 399	823.8000	1645.5854	1644.8141	0.7713	0	86	5.4e-006	1	U	K.SPQELLCGASLISDR.W
1571	385 - 399	823.9300	1645.8454	1644.8141	1.0313	0	98	1.2e-007	1	U	K.SPQELLCGASLISDR.W
1292	612 - 622	682.0700	1362.1254	1361.6980	0.4275	1	54	0.00033	1	U	K.WIQKVIDQFGE.-
1295	612 - 622	682.2400	1362.4654	1361.6980	0.7675	1	52	0.0022	1	U	K.WIQKVIDQFGE.-
1299	612 - 622	682.8500	1363.6854	1362.6820	1.0035	1	43	0.032	1	U	K.WIQKVIDQFGE.- + Deamidated (NQ)

b

```
        10         20         30         40         50         60
TFGSGEADCG LRPLFEKKSL EDKTERELLE SYIDGRIVEG SDAEIGMSPW QVMLFRKSPQ

        70         80         90        100        110        120
ELLCGASLIS DRWVLTAAHC LLYPPWDKNF TENDLLVRIG KHSRTRYERN IEKISMLEKI

       130        140        150        160        170        180
YIHPRYNWRE NLDRDIALMK LKKPVAFSDY IHPVCLPDRE TAASLLQAGY KGRVTGWGNL

       190        200        210        220        230        240
KETWTANVGK GQPSVLQVVN LPIVERPVCK DSTRIRITDN MFCAGYKPDE GKRGDACEGD

       250        260        270        280        290
SGGPFVMKSP FNNRWYQMGI VSWGEGCDRD GKYGFYTHVF RLKKWIQKVI DQFGE
```

c



d

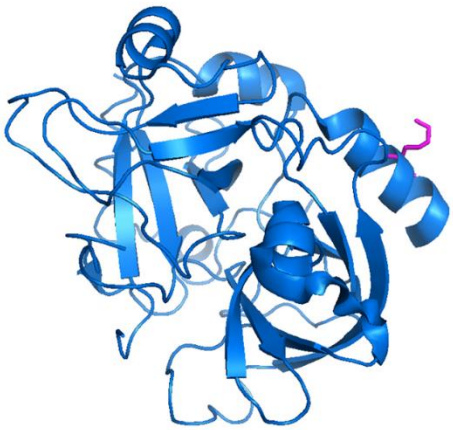


Figure S7. Position determination by mass spectrometry and labeling of the reaction points on thrombin. a, Mass spectral results for extracted band analysis. **b**, Complete sequence of human alpha thrombin. The observed peptide sequences were compared with complete sequence of thrombin. The observed strands are highlighted in Figure S7b, and the reacted lysines are labeled in red. It was determined that the ATS of thrombin occurs on Lysine 57 and Lysine 288. **c**,**d**, The positions of Lysine 57 and Lysine 288 are shown on the structure of thrombin. The two lysine residues were labeled as magenta sticks.

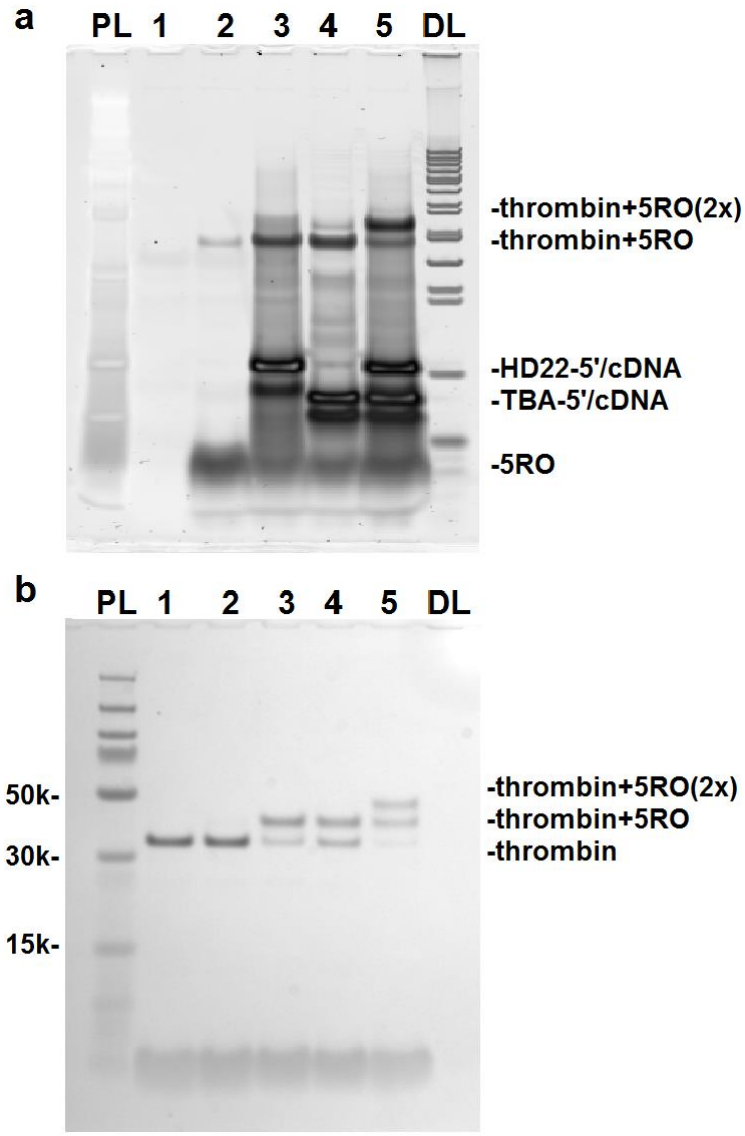


Figure S8. SDS-PAGE (4-12%) showing the formation of double conjugation product with two aptamer templates. a, Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin; lane 2, thrombin and 5RO (4 eq.); lane 3, thrombin, 5RO (2 eq.) and HD22-5' (2 eq.), with cDNA of HD22-5' replacement (4 eq.) after conjugation; lane 4, thrombin, 5RO (2 eq.) and TBA-5' (2 eq.), with cDNA of TBA-5' replacement (4 eq.) after conjugation; lane 5, thrombin, 5RO (4 eq.), HD22-5' (2 eq.) and TBA-5' (2 eq.), with cDNA of HD22-5' (4 eq.) and cDNA of TBA-5' (4 eq.) replacement after conjugation. PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, Same gel is stained with GelCode™ Blue for protein.

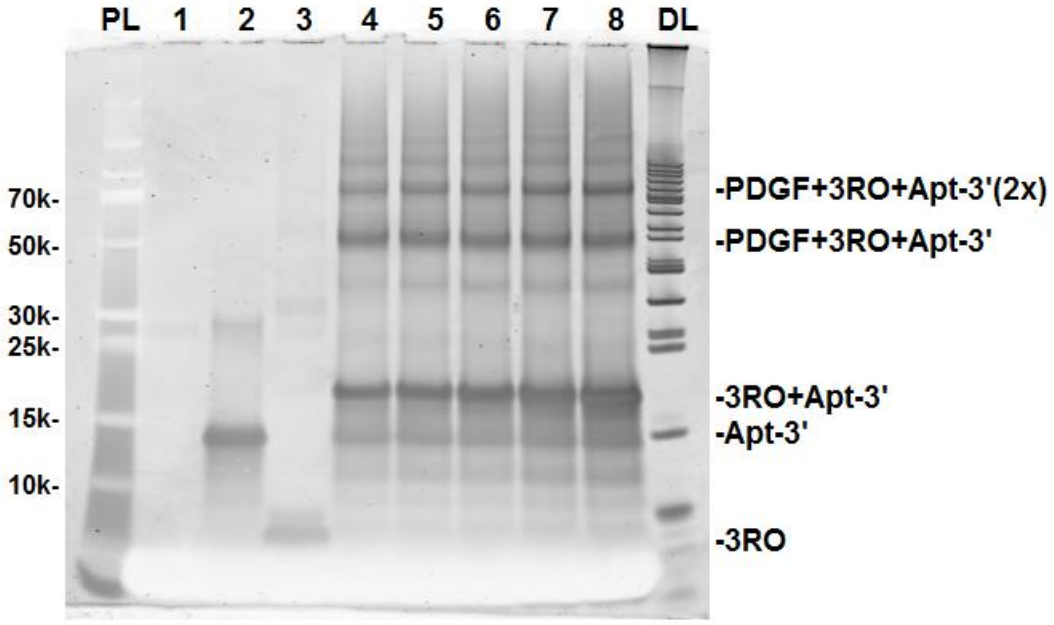


Figure S9. SDS-PAGE (4-12%) showing ATS on PDGF-BB with different amounts of DNA. Gel is stained with SYBR® Gold for DNA. More double-conjugate product can be seen with increased concentration of DNA. Lane 1, PDGF-BB; lane 2, PDGF-BB and PD-apt-3' (1 eq.); lane 3, PDGF-BB and 3RO (1 eq.); lane 4, PDGF-BB, 3RO (1 eq.) and PD-apt-3' (1 eq.); lane 5, PDGF-BB, 3RO (1.5 eq.) and PD-apt-3' (1.5 eq.); lane 6, PDGF-BB, 3RO (2 eq.) and PD-apt-3' (2 eq.); lane 7, PDGF-BB, 3RO (3 eq.) and PD-apt-3' (3 eq.); lane 8, PDGF-BB, 3RO (4 eq.) and PD-apt-3' (4 eq.). PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific).

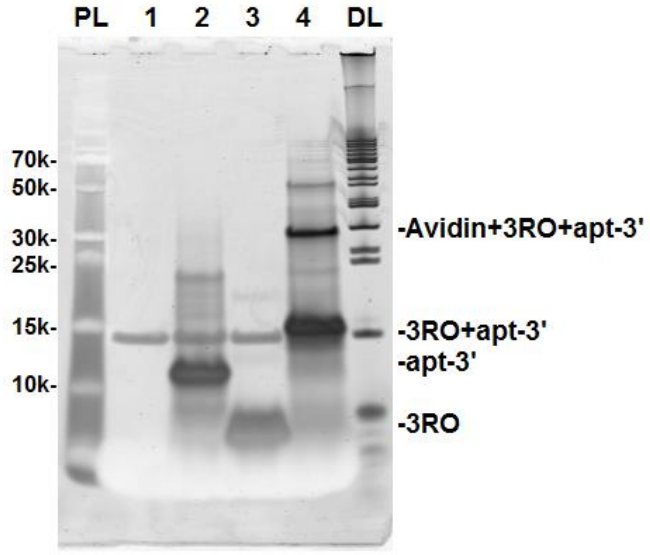


Figure S10. SDS-PAGE (4-12%) showing ATS on Streptavidin. We started the reaction with Streptavidin (homotetramer of Avidin). The protein decomposed during the preparation of SDS-PAGE and gave the Avidin band at ~13 kDa and product at ~36 kDa. Lane 1, Avidin; lane 2, Avidin and apt-3' (1 eq.); lane 3, Avidin and 3RO (1 eq.); lane 4, Avidin, 3RO (1 eq.) and apt-3' (1 eq.). PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). Gel is stained with SYBR® Gold for DNA.

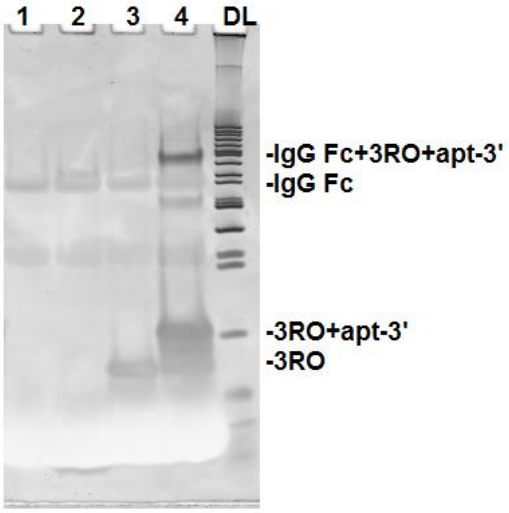


Figure S11. SDS-PAGE (4-12%) showing ATS on Human IgG Fc domain. RNA aptamer with 2'-fluoro modification is used. Lane 1, human IgG Fc; lane 2, human IgG Fc and aptamer-3' (2 eq.); lane 3, human IgG Fc and 3RO (2 eq.); lane 4, human IgG Fc, 3RO (2 eq.) and aptamer-3' (2 eq.). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). Gel is stained with SYBR® Gold for DNA. Lane 2 does not stain well since it is an RNA aptamer.

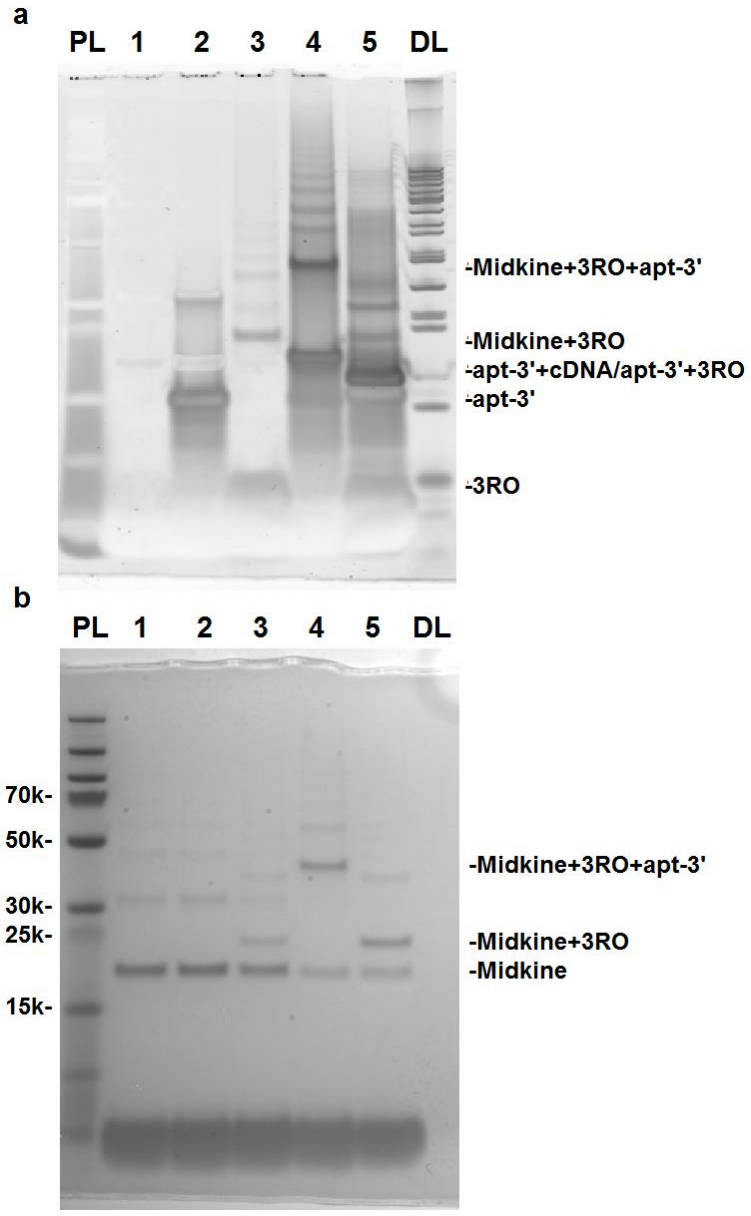


Figure S12. SDS-PAGE (4-12%) showing ATS on Midkine with N' His tag and His tag aptamer-3' template. a, Gel is stained with SYBR® Gold for DNA. Lane 1, Midkine (with N' His tag); lane 2, Midkine and His-apt-3' (0.9 eq.); lane 3, Midkine and 3RO (0.9 eq.); lane 4, Midkine, 3RO (0.9 eq.) and His-apt-3' (0.9 eq.); lane 5, Midkine, 3RO (0.9 eq.) and His-apt-3' (0.9 eq.), with cDNA of His-apt-3' (1.8 eq.) replacement after conjugation. PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, Same gel is stained with GelCode™ Blue for protein. The conversion is 62% based on ImageJ analysis.

For the protein stain, the band with high molar mass is due to the impurity of the protein.

The multiple conjugation is due to the strucuture of Midkine. The sequence of the Midkine is shown as follow:

MKHHHHHHHM KKKDKVKKGG PGSECAEWAW GPCTPSSKDC GVGFREGTCG AQTQRIRCRV PCNWKKEFGA DCKYKFENWG ACDGGTGTKV RQGTLKKARY NAQCQETIRV TKPCTPKTKA KAKAKKGKGK D.

There are multiple lysines very close to the His-tag, thus it would be difficult to achieve only single-modification of the protein.

Similar situation for Figure S13.

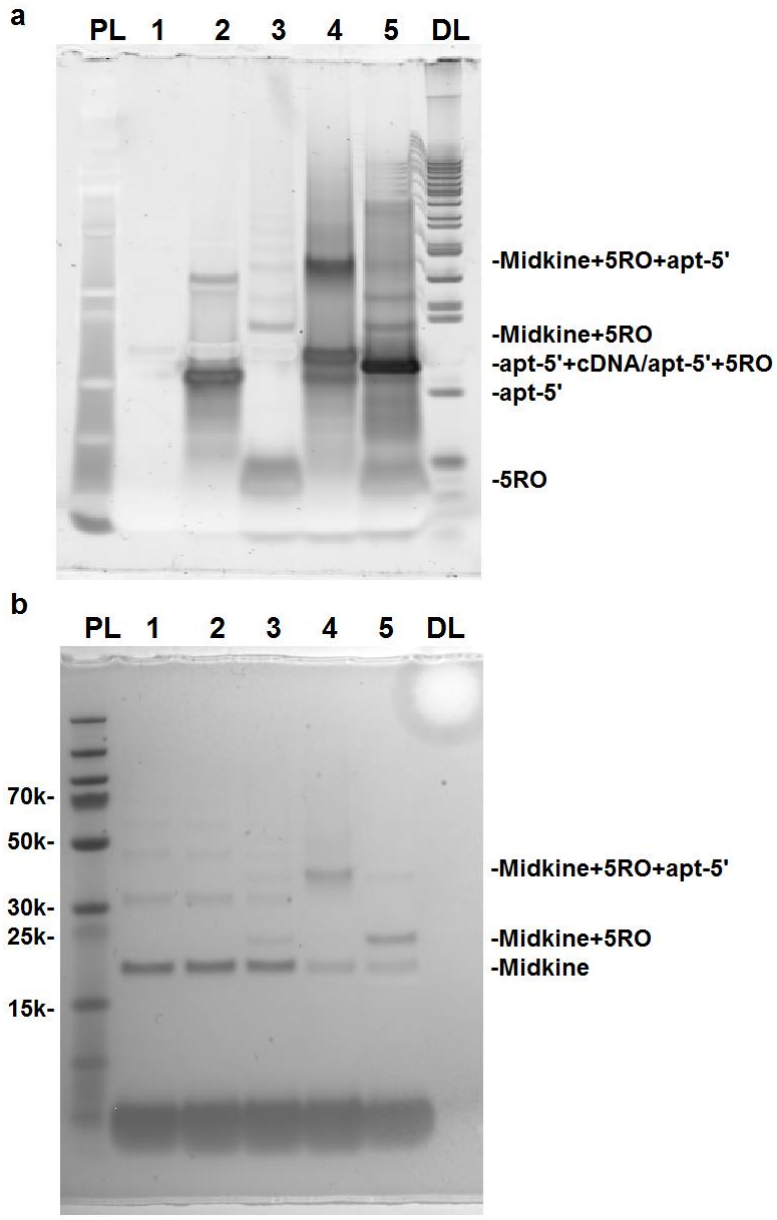


Figure S13. SDS-PAGE (4-12%) showing ATS on Midkine with N' His tag with His tag aptamer-5' template. a, Gel is stained with SYBR® Gold for DNA. Lane 1, Midkine (with N' His tag); lane 2, Midkine and His-apt-5' (0.9 eq.); lane 3, Midkine and 5RO (0.9 eq.); lane 4, Midkine, 5RO (0.9 eq.) and His-apt-5' (0.9 eq.); lane 5, Midkine, 5RO (0.9 eq.) and His-apt-5' (0.9 eq.), with cDNA of His-apt-5' (1.8 eq.) replacement after conjugation. PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, Same gel is stained with GelCode™ Blue for protein. The conversion is 69% based on ImageJ analysis.

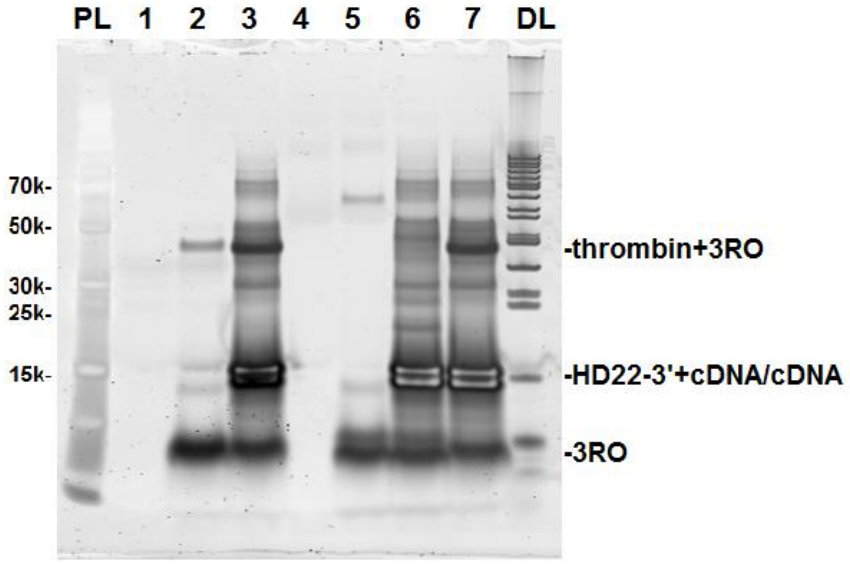


Figure S14. SDS-PAGE (4-12%) showing selective conjugation on thrombin protein using HD22-3' template. Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin; lane 2, thrombin and 3RO (2 eq.); lane 3, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); lane 4, BSA; lane 5, BSA and 3RO (2 eq.); lane 6, BSA, 3RO (2 eq.) and HD22-3' (2 eq.); lane 7, thrombin, BSA, 3RO (2 eq.), and HD22-3' (2 eq.). For lane 3, 6 and 7, cDNA of HD22-3' (4 eq.) was added for replacement after conjugation. PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific).

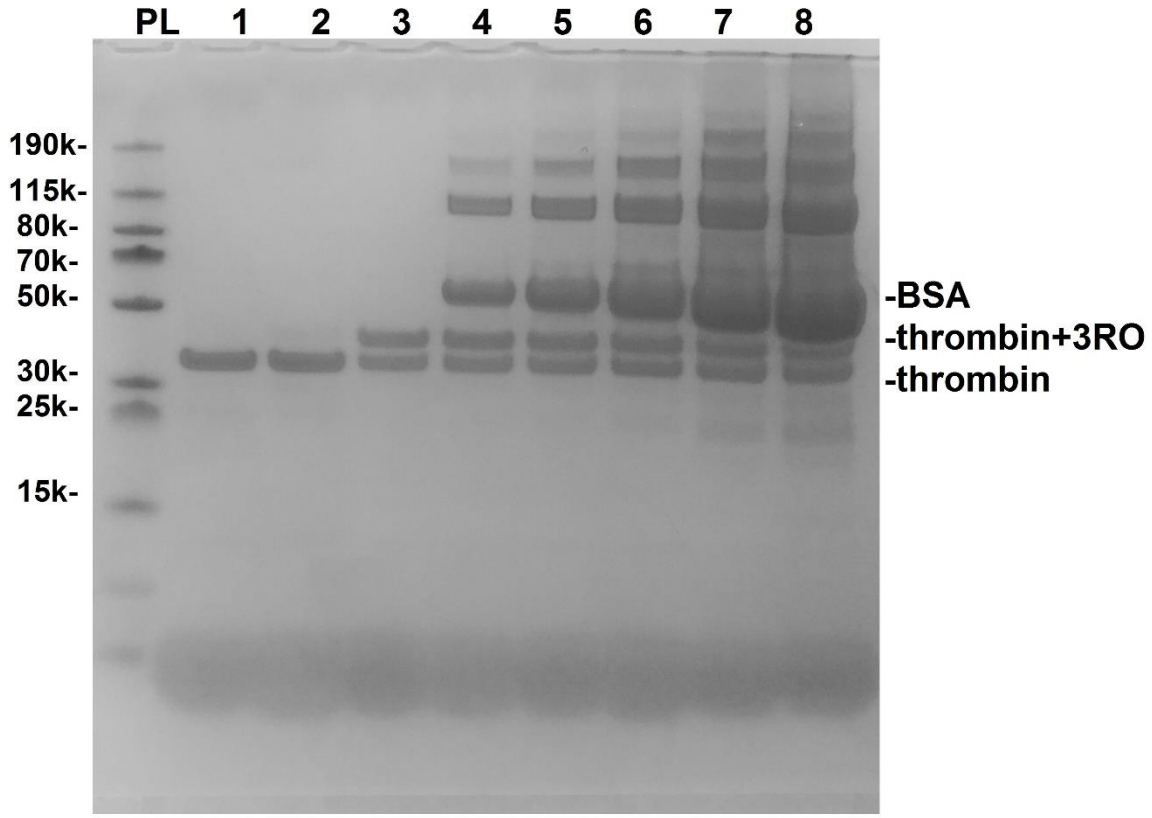


Figure S15. SDS-PAGE (4-12%) showing high concentration of negative protein BSA (20 times) would not affect the selective conjugation on thrombin protein using HD22-3' template. Gel is stained with GelCode™ Blue for protein. Lane 1, thrombin; lane 2, thrombin and 3RO (2 eq.); lane 3, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); lane 4, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), BSA (1 eq.); lane 5, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), BSA (2 eq.); lane 6, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), BSA (5 eq.); lane 7, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), BSA (10 eq.); lane 8, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.), BSA (20 eq.). For lane 3-8, cDNA of HD22-3' (4 eq.) was added for replacement after conjugation. PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific).

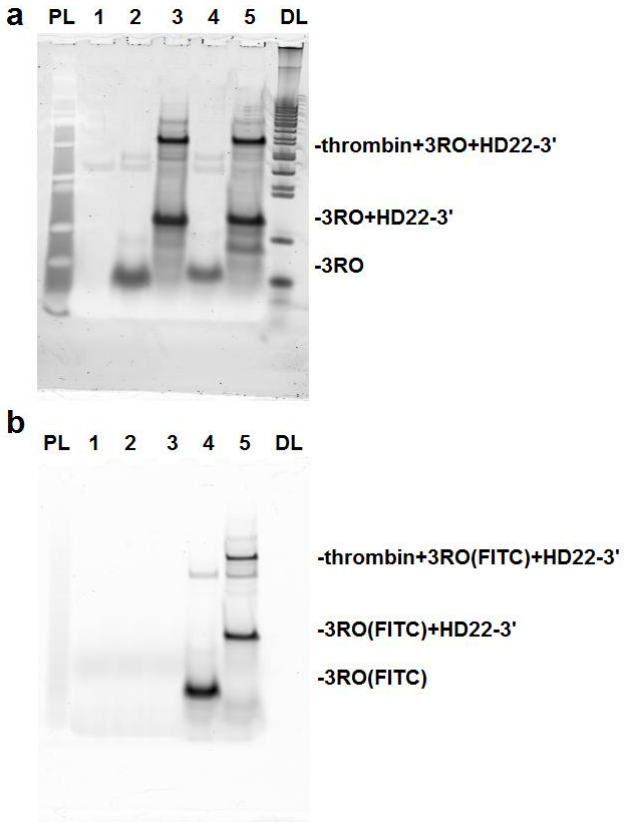


Figure S16. SDS-PAGE (4-12%) showing the feasibility of FITC-labeled RO in ATS. a, Gel is stained with SYBR® Gold for DNA. Lane 1, thrombin; lane 2, thrombin and 3RO (2 eq.); lane 3, thrombin, 3RO (2 eq.) and HD22-3' (2 eq.); lane 4, thrombin and FITC-3RO (2 eq.); lane 5, thrombin, FITC-3RO (2 eq.) and HD22-3' (2 eq.). PL, PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). DL, Fisher BioReagents™ exACTGene™ DNA Ladders (Fisher Scientific). **b**, FITC fluorescence scanned for the same gel.

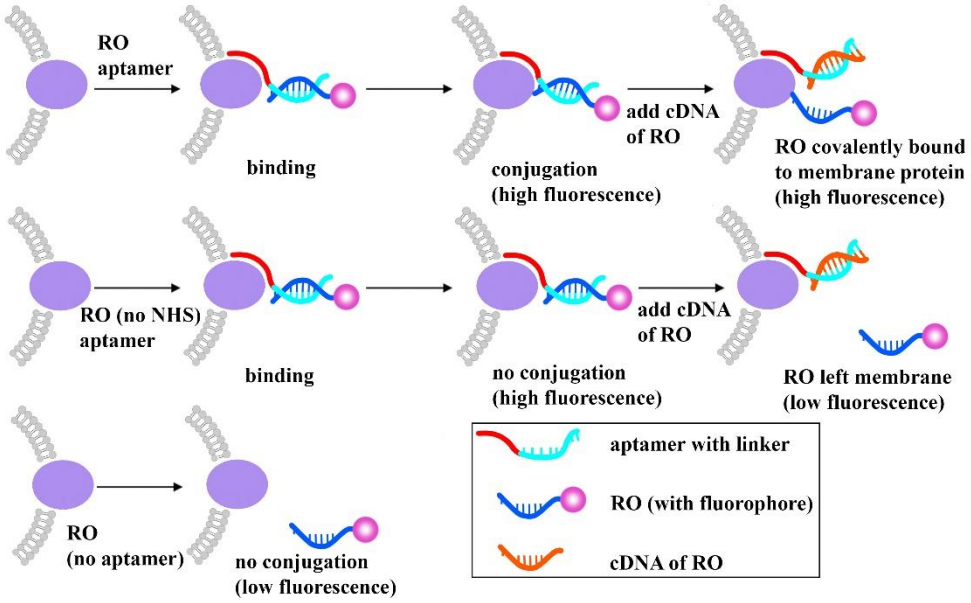


Figure S17. Cellular ATS experimental design with negative control groups. Among three groups, cells react with 1) FITC-RO and aptamer template, 2) FITC-RO, but not with NHS ester (no reacting residue) and template, and 3) FITC-RO and no template. After binding and reaction, RO is on the cell for the first and second groups, but not the third group. Then the cDNA of RO is added. If no covalent bond is formed, as in the second group, RO leaves the cell membrane as a negative control.

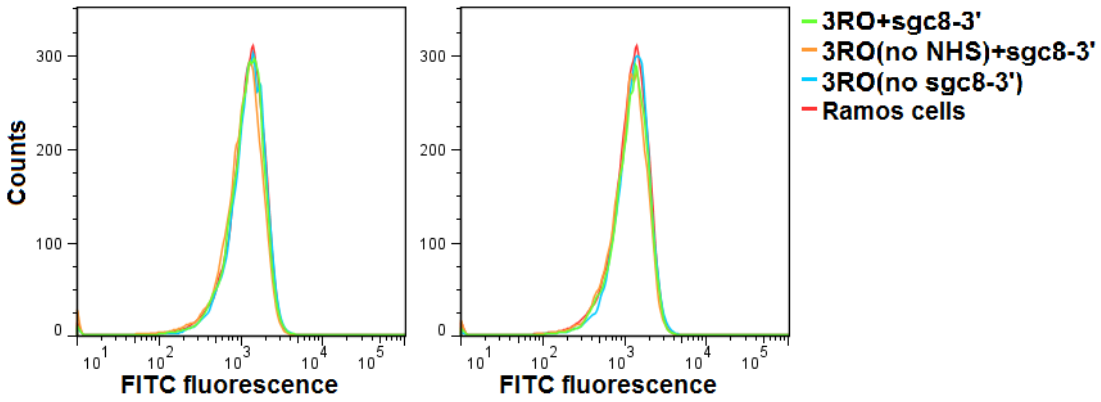


Figure S18. Flow cytometry of Ramos cells showing the selectivity of ATS. Sgc8 aptamer template, which does not bind to Ramos cells, was used. The conditions were set the same as those for CEM ATS. Left, Ramos cells before adding cDNA of 3RO; right, Ramos cells after adding cDNA of 3RO.

SUPPORTING INFORMATION

Table S1

DNA sequences. All linkers are shown in red. In the IgG Fc aptamer, rA and rG stand for RNA A and G, while i2FU and i2FC stand for int 2'-Fluoro RNA U and C.

3RO (3' NHS ester-modified DNA)	5' TAC ATA CAG CCT CGC ATG AGC CC-Amino 3'
5RO (5' NHS ester-modified DNA)	5' Amino-CGT AGA CCG TAG ATG GCA GAT CT 3'
HD22-3'	5' AGT CCG TGG TAG GGC AGG TTG GGG TGA CTG GGC TCA TGC GAG GCT TAC GAA C 3'
HD22-5'	5' TAG AAT CGC ATC TAC GGT CTA CGA GTC CGT GGT AGG GCA GGT TGG GGT GAC T 3'
HD22-3'5'	5' TAG AAT CGC ATC TAC GGT CTA CGA GTC CGT GGT AGG GCA GGT TGG GGT GAC TGG GCT CAT GCG AGG CTT ACG AAC 3'
cDNA of HD22-3'	5' GTT CGT AAG CCT CGC ATG AGC CCA GTC ACC CCA ACC TGC CCT ACC ACG GAC T 3'
cDNA of HD22-5'	5' AGT CAC CCC AAC CTG CCC TAC CAC GGA CTC GTA GAC CGT AGA TGC GAT TCT A 3'
cDNA of HD22-3'5'	5' GTT CGT AAG CCT CGC ATG AGC CCA GTC ACC CCA ACC TGC CCT ACC ACG GAC TCG TAG ACC GTA GAT GCG ATT CTA 3'
TBA-3'	5' GGT TGG TGT GGT TGG GGG CTC ATG CGA GGC TTA CGA AC 3'
TBA-5'	5' TAG AAT CGC ATC TAC GGT CTA CGG GTT GGT GTG GTT GG 3'
cDNA of TBA-3'	5' GTT CGT AAG CCT CGC ATG AGC CCC CAA CCA CAC CAA CC 3'
cDNA of TBA-5'	5' CCA ACC ACA CCA ACC CGT AGA CCG TAG ATG CGA TTC TA 3'
PDGF aptamer-3'	5' CAC AGG CTA CGG CAC GTA GAG CAT CAC CAT GAT CCT GTG GGG CTC ATG CGA GGC TTA CGA AC 3'
StreptAvidin aptamer-3'	5' GGG AAC GCA CCG ATC GCA GGT TTC CCG GGC TCA TGC GAG GCT TAC GAA C 3'
Human IgG Fc aptamer-3'	5' rGrGrA rGrG/i2FU/ rG/i2FC//i2FU/ /i2FC//i2FC/rG rArArA rGrGrA rA/i2FC//i2FU/ /i2FC//i2FC/rA GGG CTC ATG CGA GGC TTA CGA AC 3'
His-tag aptamer-3'	5' GCT ATG GGT GGT CTG GTT GGG ATT GGC CCC GGG AGC TGG CGG GCT CAT GCG AGG CTT ACG AAC 3'
His-tag aptamer-5'	5' TAG AAT CGC ATC TAC GGT CTA CGG CTA TGG GTG GTC TGG TTG GGA TTG GCC CCG GGA GCT GGC 3'
cDNA of His-tag aptamer-3'	5' GTT CGT AAG CCT CGC ATG AGC CCG CCA GCT CCC GGG GCC AAT CCC AAC CAG ACC ACC CAT AGC 3'
cDNA of His-tag aptamer-5'	5' GCC AGC TCC CGG GGC CAA TCC CAA CCA GAC CAC CCA TAG CCG TAG ACC GTA GAT GCG ATT CTA 3'
sgc8-3'	5' ATC TAA CTG CTG CGC CGC CGG GAA AAT ACT GTA CGG TTA GAG GGC TCA TGC GAG GCT TAC GAA C 3'
cDNA of 3RO	5' GGG CTC ATG CGA GGC TGT ATG TA 3'
mismatch (random aptamer) template	5' TAG AAT CGC ATC TAC GGT CTA CGG CTA GTC TCA TGC GAG GCT TAC GAA C 3'
mismatch (random linker) template	5' AGT CCG TGG TAG GGC AGG TTG GGG TGA CCT CGG ATA CCG ATT TAC GAA C 3'

Author Contributions

C.C. and W.T. conceived the idea and designed the project. C.C., H.Z., R.W., Q.L. and S.C. performed most of the experiments. S.W., W.H., X.P., L.L., M.C., Y.L., R.W., X.C. and Q.L. provided advice on performing experiments and performed some experiments. C.C., X.P. Q.L. and W.T. analyzed the data and wrote the manuscript. W.T. supervised the project.