Supporting Information

Quencher-Free Molecular Beacons as Probes for Oligonucleotides Containing CAG Repeat Sequences

Ki Tae Kim,[a] Rakesh N. Veedu,[b] Young Jun Seo[c] and Byeang Hyean Kim*[a]

[a] *Department of Chemistry, Pohang University of Science and Technology, Pohang 790-784, South Korea*

[b] *School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD 4072, Australia*

[c] *Department of Chemistry, Chonbuk National University, Jeonju 561-756, South Korea*

General experimental details

Synthesis of ^{Py}U and corresponding phosphoramidites

The ^{Py}**U** and its phosphoramidites were synthesized according to previous synthetic methods.[1] Each reaction was performed under anhydrous conditions under Ar or N_2. All reagents, including 1-ethynylpyrne and 5-iodo-2´-deoxyuridine, were purchased from Sigma-Aldrich, Fluka, TCI, Alfa, and Proligo and used without additional purification. We obtained the MC, THF, DMF from distill tower. For confirmation of synthesized compound, we used 1H NMR spectra recorded by FT-300MHz Bruker Aspect 3000 spectrometer and high-resolution fast atom bombardment (HRMS-FAB) mass spectra recorded by a Jeol JMS700 HR mass spectrometer at the Korea Basic Science Center, Daegu, Korea.

Synthesis and identification of oligonucleotides

ON-A, **ON-T**, $(CAG)_{20}$, $(CAG)_{30}$, and all mRNA samples were purchased from Integrated DNA Technologies (IDT). All reagents for synthesis of oligonucleotides were purchased from Proligo, Sigma-Aldrich, and Glen research. LNA-T phosphoramidities were provided by Exiqon. DNA synthesis was performed using a Polygen 12 synthesizer. Synthesized oligonucleotides were removed from the controlled pore glass (CPG) by treatment with 28 ~ 30 % aqueous ammonia (13h. at 55 °C). After filtration of CPG, the filtrate was dried by a freeze dryer. DMT-on and off oligonucleotides were purified by HPLC. An Agilent high-performance liquid chromatography system (1100 Series) was used to purify the synthesized ODNs; Agilent, ZORBA X Eclipse XDB-C18, 4.6 150 mm; gradient elution: 0 min, A:B = 90:10; 10 min, A:B = 90:10; 20 min, A:B = 0:100; 25 min, A:B = 0:100; 30 min, A:B = 90:10; solution A, 0.1 M triethylammonium acetate (TEAA) buffer (pH 7.2)/MeCN = 95:5; solution B, 0.1 M TEAA buffer (pH 7.2)/MeCN = 1:1); flow rate: 2.5 mL/min; UV detection: 254 and 380 nm. MALDI-TOF mass spectra were recorded at Bioneer, Taejon, Korea for confirmation of synthesized oligonucleotides; 10 μL of 0.5 M 3-Hydroxypicolinic acid (99 %, Fluka) was added to oligonucleotide samples. After voltexing and spindown, we loaded 1~2 μL of the mixed sample in MALDI-TOF plate and waited for formation of crsyal. These samples were analyzed by Kratos AXIMA-LNR MALDI-TOF mass (voltage : 20 kV, mode : linear mode, laser wavelength : 337 nm, pulse width : 4 ns, pulse energy : 300 μJ at 10 Hz).

Quantification of oligonucleotides

UV-visible spectrophotometer (Cary 100) was used to determine the ODN concentrations. 10 μL of stock sample was diluted to 1 mL for quantification. The absorbance at 260 nm of the sample was measured. For quantification of oligonucleotides, 15400, 7400, 11500, 8700, and 4200 M^{-1} was used as extinction coefficient at 260 nm for A, C, G, T, and ^{Py}**U**, respectively. We used next equation for calculation.

$$\text{Concentration } (\mu M) = A_{260} \times \frac{1000}{15.4N_A + 7.4N_C + 11.5N_G + 8.7N_T}$$

DNA sample preparation

For CD, UV, and fluorescence measurement, 1.5 nmol of each Qf-MB (P1-5 and P3L) and 125 μL of pH 7.2 200 mM Trizma buffer containing 400 mM NaCl and 80 mM $MgCl_2$ were mixed in 1.5 mL microtube. After addition of 1.5 nmol (1 equiv.) of each target sequences and nuclease-free water, total 1 mL solution was mixed by vortex for 1 min. For annealed sample, we kept the sample at 95 °C for 3 min. and cooling down at ambient condition for 3 h. For time-dependent fluorescence enhancement test, we carried out same procedure except annealing processes.

For testing mRNA sequence, 0.1 nmol of mRNA samples (HTT REP17, SCA6 CAG13, SCA3 CAG8, and DRPLA CAG15) were annealed in 50 μL of pH 7.2 20mM Trizma buffer, 80 mM NaCl, 20mM $MgCl_2$ by heating samples at 80 °C for one minute and slow cooling to 37 °C. Prepared samples were added to 0.1 nmol of P3 sequence in 950 μL of same buffer solution. The fluorescence enhancement was recorded as a function of time at 37 °C.

CD spectra of oligonucleotides

CD spectra of oligonucleotides were recorded using JASCO J-810 (JASCO); sensitivity: standard (100 mdeg), scanning range : 320 ~ 220 nm, data pitch : 1 nm, scanning mode : continuous, scanning speed : 200 nm/min, response: 1 sec, band width : 1 nm, accumulation : 5 times, cell length : 1 cm, temperature : 20°C.

UV and fluorescence spectra

UV and fluorescence spectra were recorded using Cary 100 and Eclipse spectrometers (Varian), respectively. Samples for UV/fluorescence spectroscopy were prepared in a quartz cell (path length: 1 cm). All samples were measured after baseline correction for UV spectra. Parameters for UV spectra; scanning range : 500~200 nm, data interval : 1.0 nm, scan rate : 600 nm/min, SBW: 2.0 nm, beam mode : double, temperature : 25 °C. Parameters for fluorescence spectra; exc. wavelength: 377 nm, scanning range: 382~600nm, excitation and emission slit : 5 nm, data interval : 1.0 nm, scan rate : 600 nm/min, PMT Detector voltage : 600 V, Temperature : 25 °C.

Thermal stability measurement of oligonucleotides

For T_m determination, we utilized Cary 100 UV-visible spectrophotometer and Cary temperature controller to record the hyperchromicity; wavelength: 260nm, SBW: 1.0 nm, start: 10 °C, end: 98 °C, data interval: 1 °C, rate: 2 °C/min. All samples were recorded three times. The T_m curves were analyzed by general method for T_m determination (temperature at which half of the sample is folded, and half is unfolded).[2] T_m determination was done by both linear fitting of upper and lower baseline and first derivative of the absorbance signal. For the T_m curve of P3L exhibiting no upper baseline, only first derivative method was used.

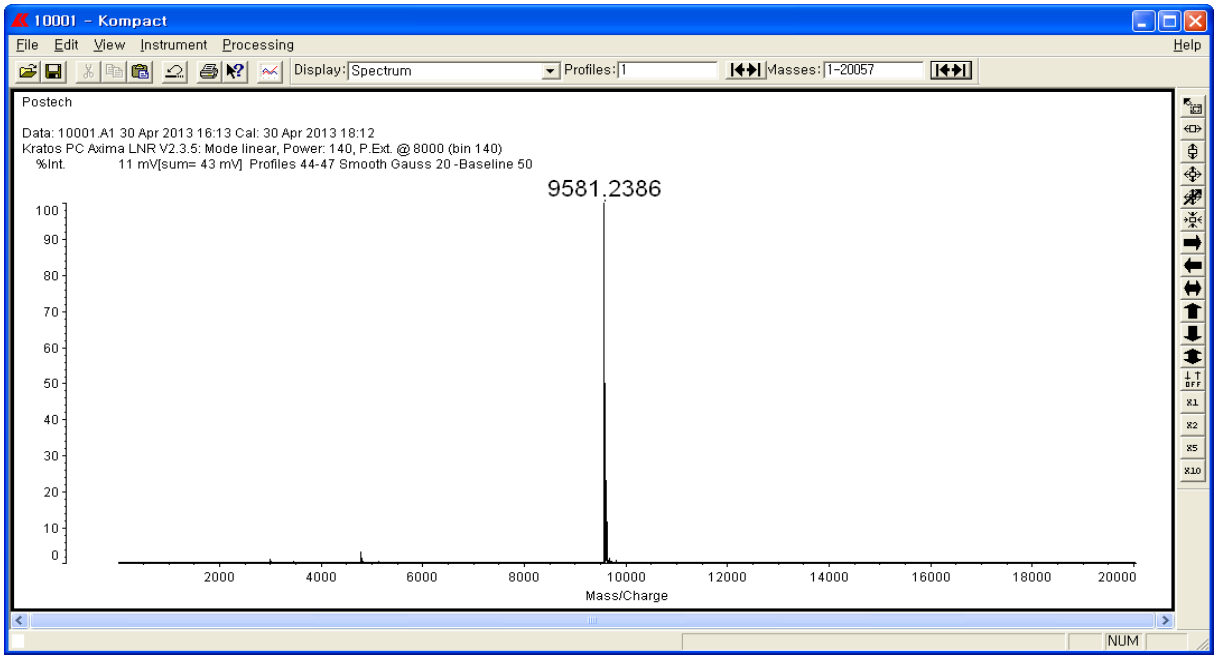
Polyacrylamide Gel electrophoresis (PAGE)

5 mL of 40% acrylamide, 2 mL of 5X TBE buffer, and 3 mL of distilled water were mixed for 20 % non-denaturing gel. 12 mg of ammonium persulfate was added to the mixed solution. For initiation of gel formation, 10 µL of N, N**,** N',N'-Tetramethylethylenediamine (TEMED) was added. 100 pmol of samples (hairpin or duplex) were used for PAGE. Dried samples were dissolved in 6 µL of water/formamide mixture (v/v = 1: 1) for sample loading. We used 8 µL of O'RangeRuler 5 bp DNA Ladder, ready-to-use (Fermentas) for size marker. PAGE was carried out at 90V, 260mA, 22W, 25°C for 3 h. After running, gels were mixed with stains all (Sigma-Aldrich) in formamide for 30 min. Gels were dried and exposed on light for visualization of DNA bands.

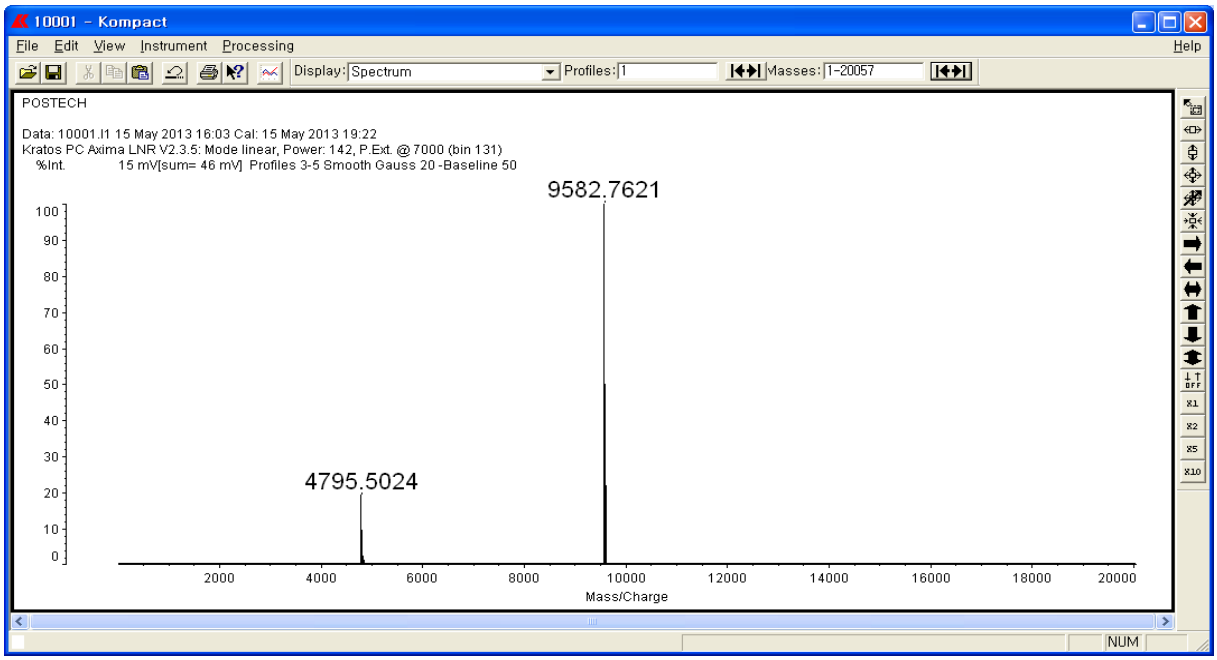
Calculation of photophysical properties

For calculation of quantum yield, we used quinine sulfate in 0.5 M H_2SO_4 (1~ 10 µM, 0.546, excitation at 310 nm) as reference material.[3] The quantum yields (QYs) of oligonucleotides containing the ^{Py}**U** were calculated according to a previous report.[4] We checked absorbance at 377 nm and used it for excitation of the ^{Py}**U**s. Refractive indices were 1.333. For calculation of extinction coefficients of oligonucleotides, we used Beer–Lambert law (A = ε l c, l: 1 cm, c: 1.5 µM). The fluorescence enhancement (FE) was calculated by dividing fluorescence intensity at 430 nm of the initial sample by that of the final sample. The Brightness was expressed by QY × ε_{377}.

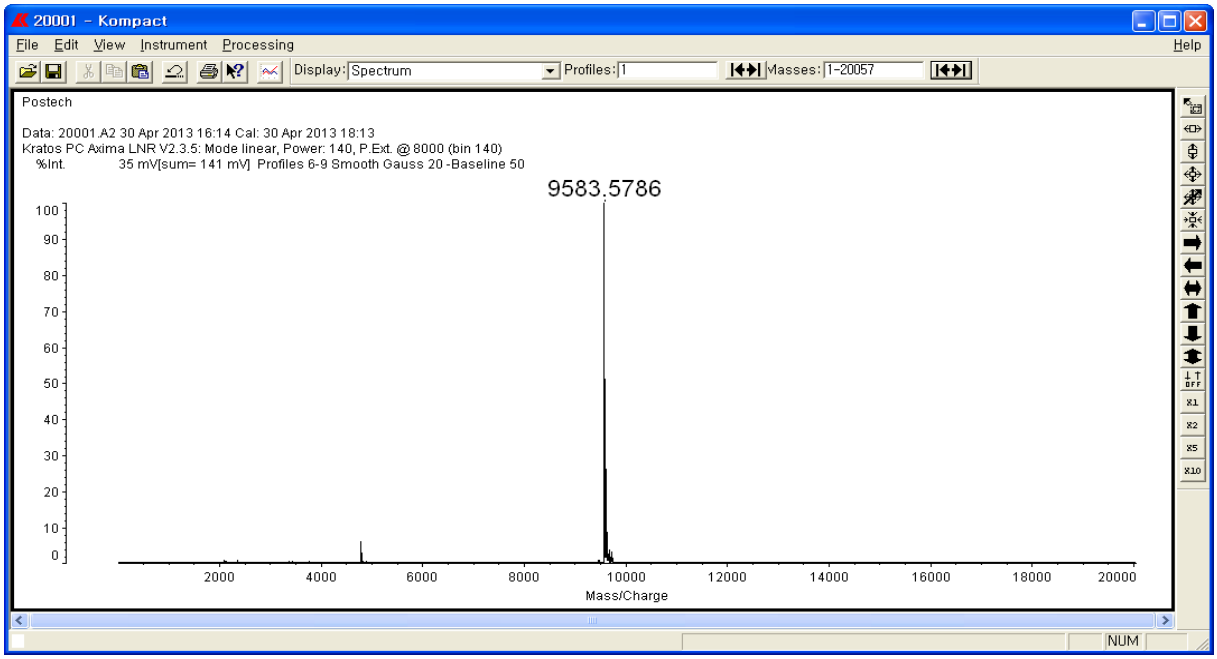
Fig. S1 MALDI-TOF mass spectra of the tested ODN sequences.



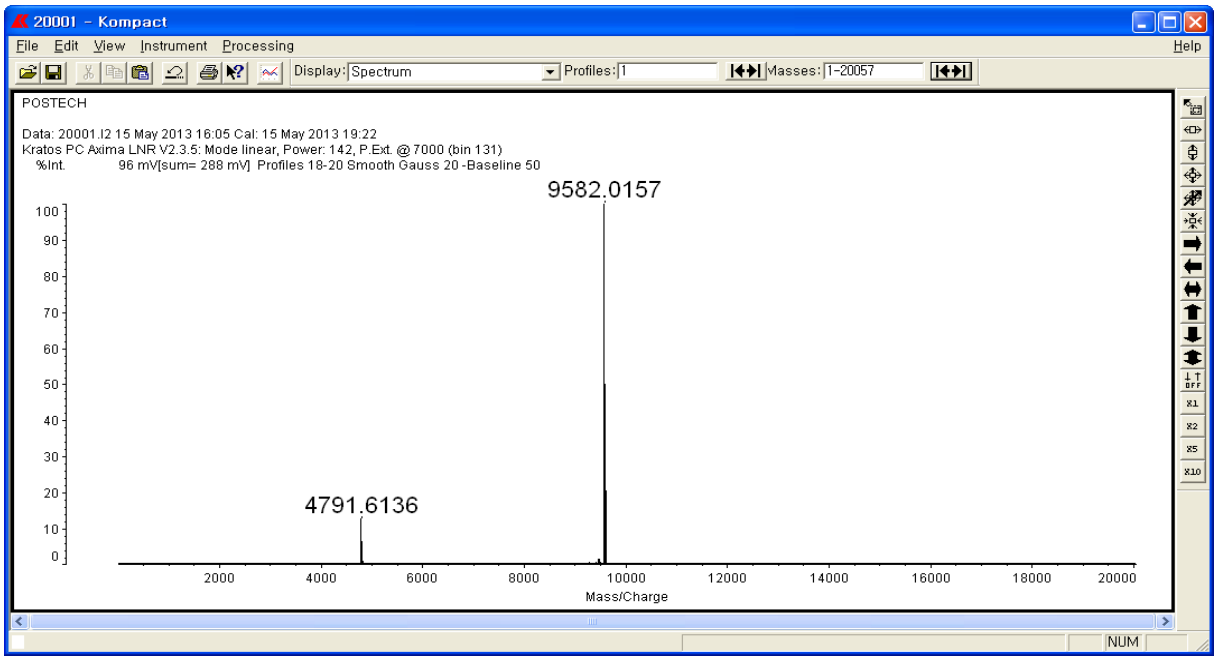
(a) P1; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



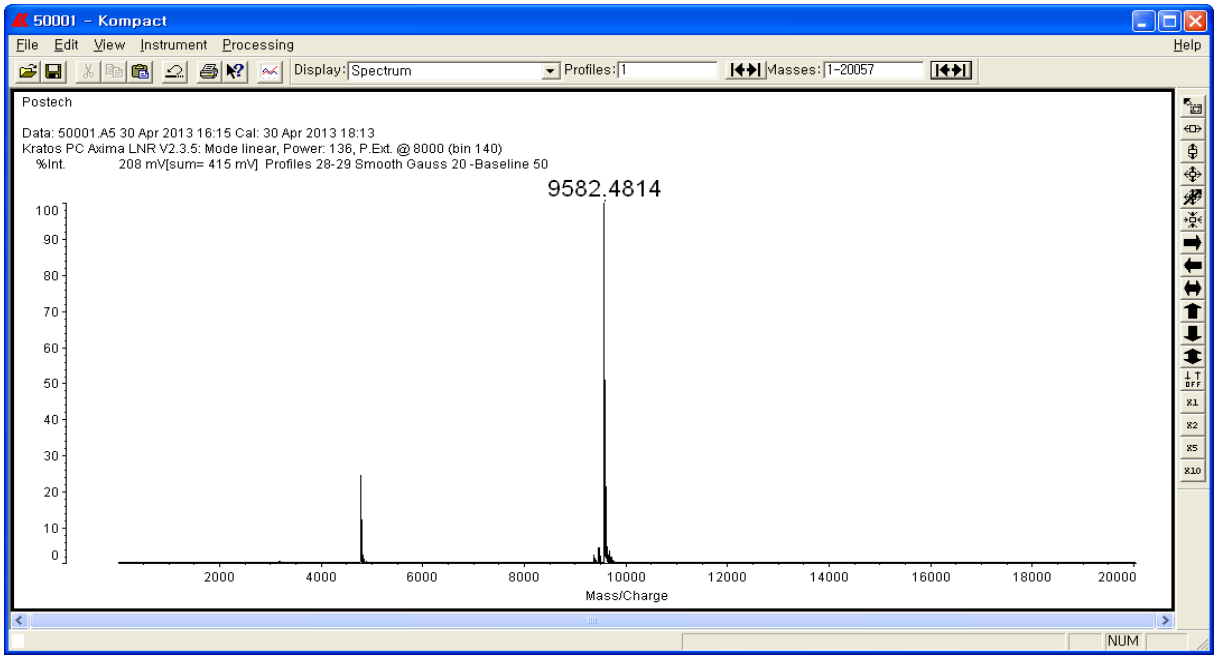
(b) P2; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



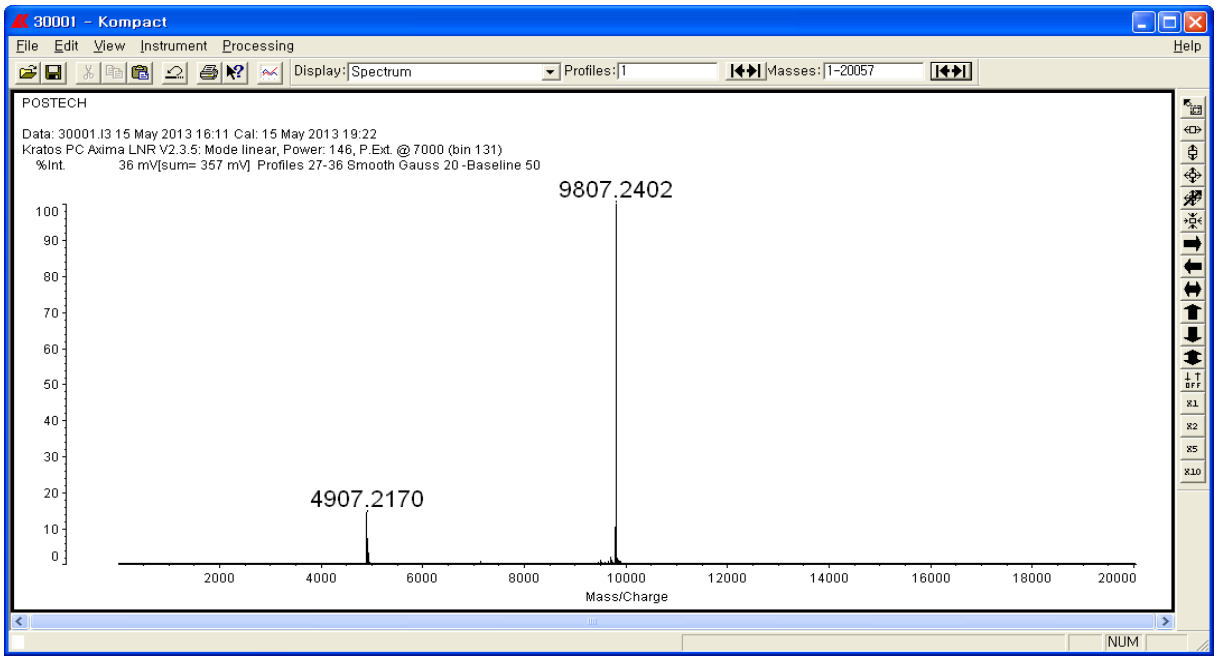
(c) P3; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



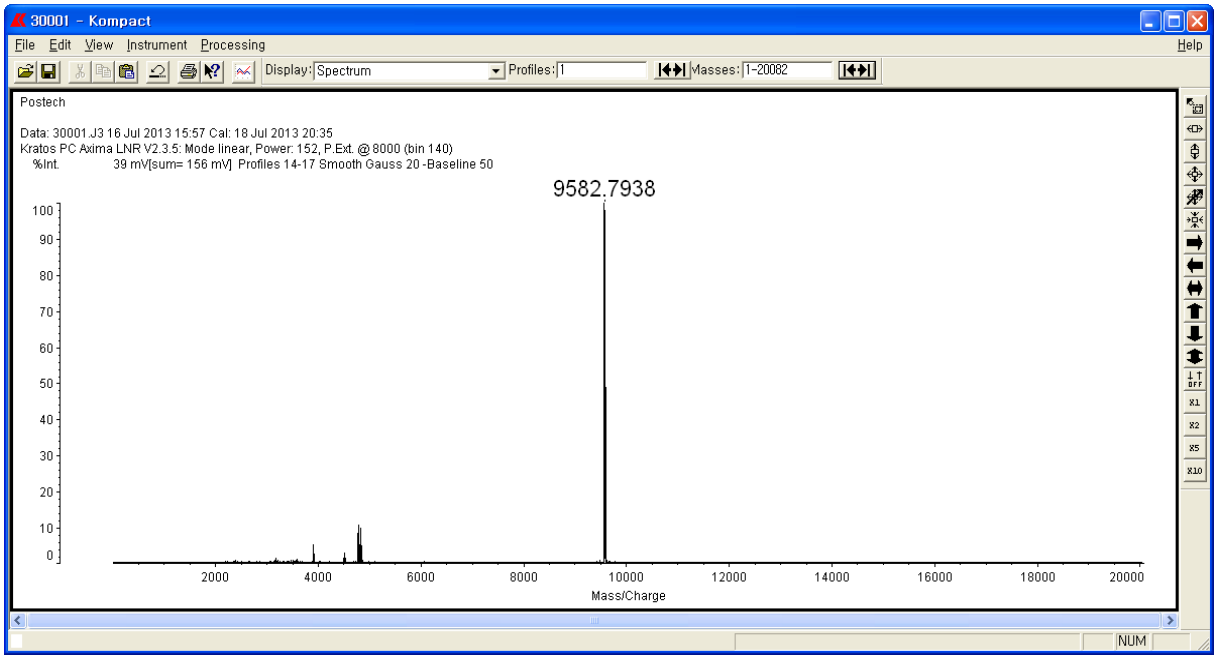
(d) P4; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



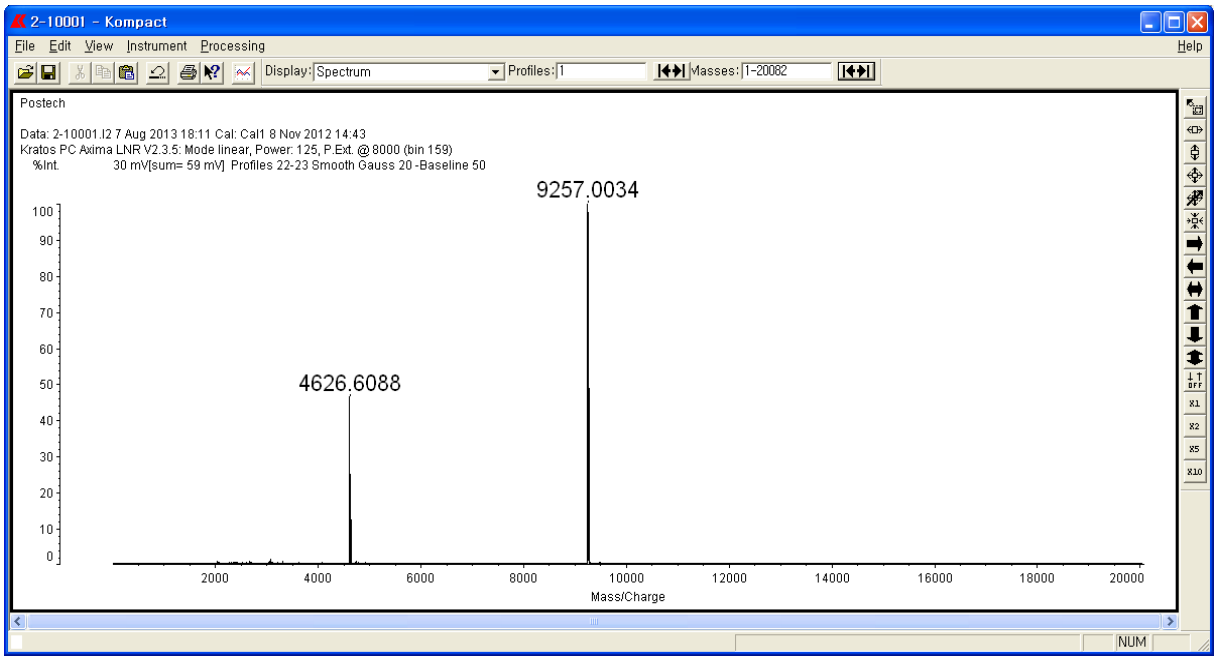
(e) P5; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



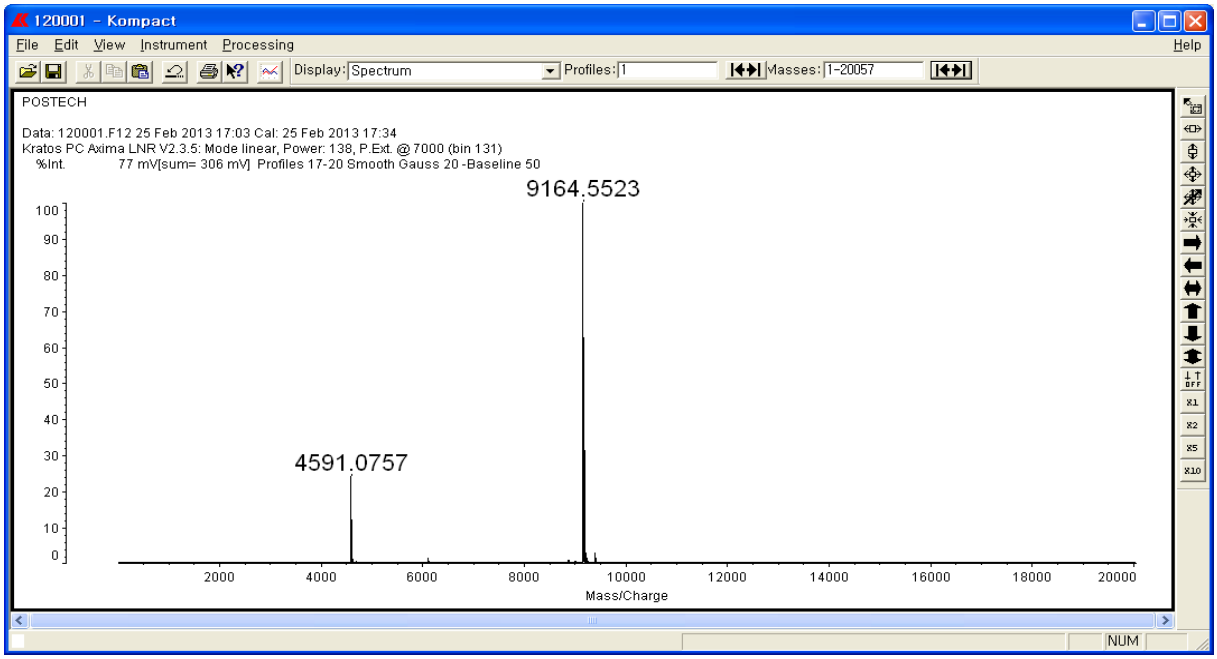
(f) P3L; calcd. Mass for $C_{332}H_{383}N_{100}O_{196}P_{29}$: (*m*/*z*) 9808.1261



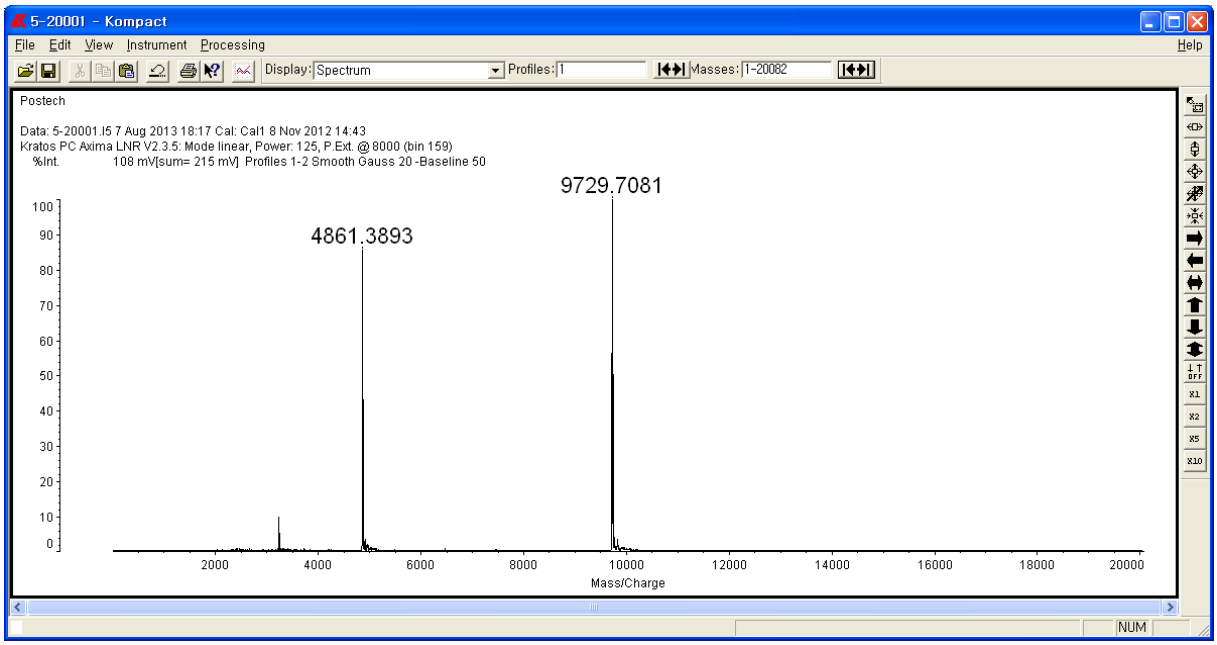
(g) P3X; calcd. Mass for $C_{324}H_{383}N_{100}O_{188}P_{29}$: (*m*/*z*) 9584.0453



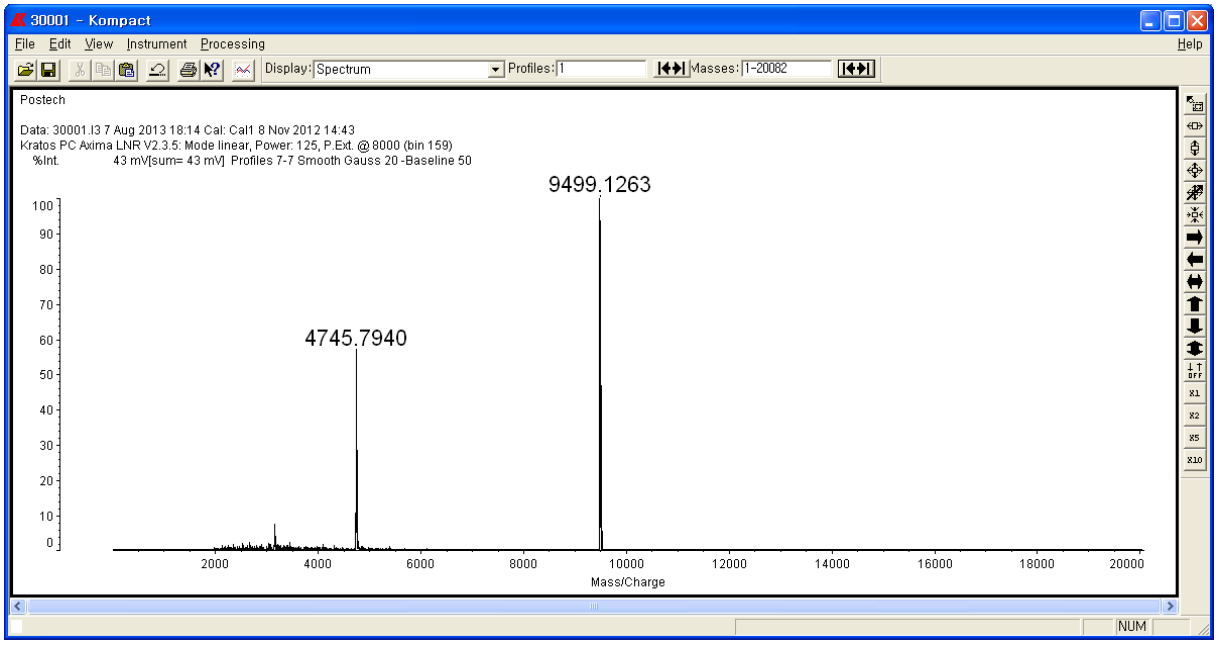
(h) ODN-A; calcd. Mass for $C_{290}H_{361}N_{130}O_{168}P_{29}$: (*m*/*z*) 9253.7193



(i) ODN-T; calcd. Mass for $C_{290}H_{371}N_{100}O_{188}P_{29}$: (*m*/*z*) 9163.5863

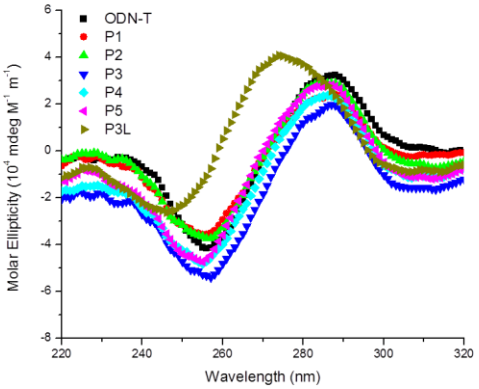


(j) ON-A; calcd. Mass for $C_{290}H_{361}N_{130}O_{198}P_{29}$: (*m*/*z*) 9733.7013

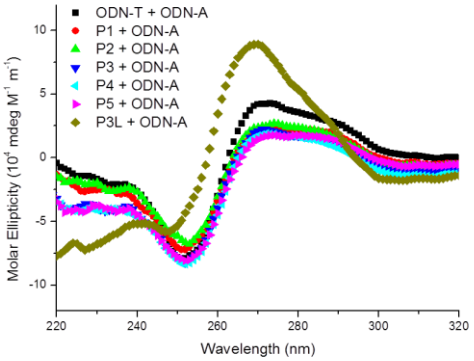


(k) ON-U; calcd. Mass for $C_{280}H_{351}N_{100}O_{218}P_{29}$: (*m/z*) 9503.3023

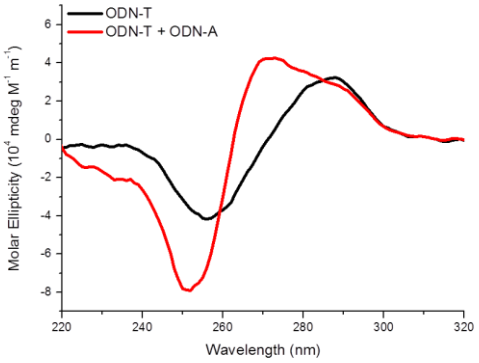
Fig. S2 CD spectra of (a) the hairpin states of the sequences **P1**–**P5**, **P3L**, and **ODN-T**; (b) the DNA/DNA duplex states of **P1**–**P5**, **P3L**, and **ODN-T**; (c–i) the hairpin and DNA/DNA hybrid state of (c) **ODN-T**, (d) **P1**, (e) **P2**, (f) **P3**, (g) **P4**, (h) **P5**, and (i) **P3L**; (j) the DNA/RNA duplex states of **P1**–**P5**, **P3L**, and **ODN-T**; (k–q) the hairpin and DNA/RNA hybrid states of (k) **ODN-T**, (l) **P1**, (m) **P2**, (n) **P3**, (o) **P4**, (p) **P5**, and (q) **P3L**; and (r) the duplex states of **P3L** and **ODN-T**. [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C]



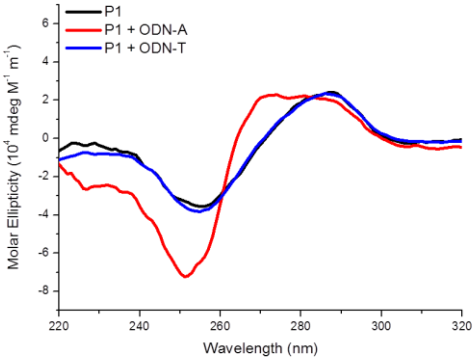
(a)



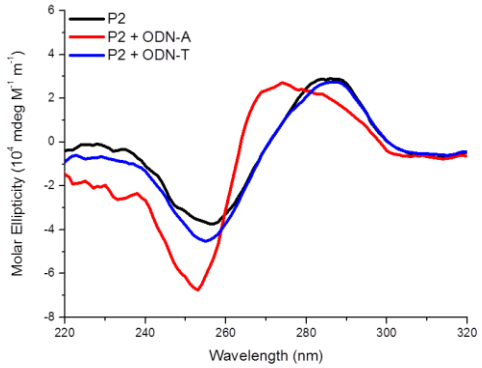
(b)



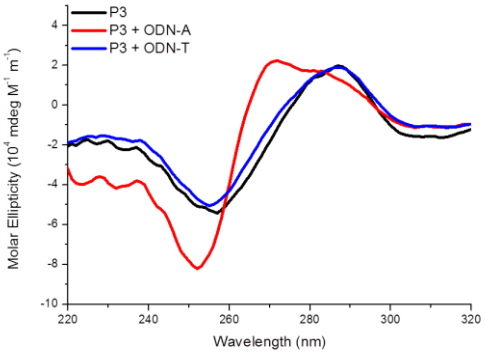
(c)



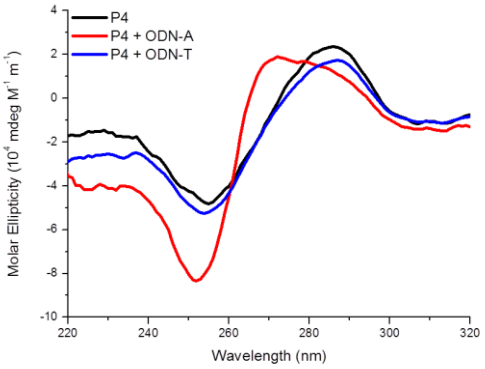
(d)


P2
P2 + ODN-A
P2 + ODN-T
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

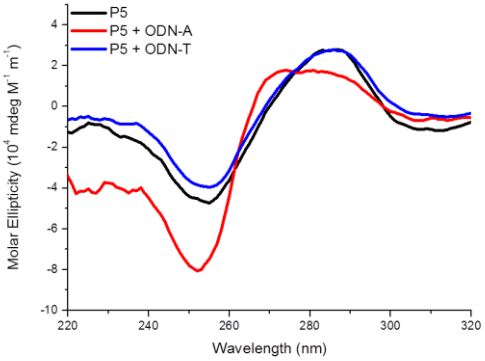
(e)


P3
P3 + ODN-A
P3 + ODN-T
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

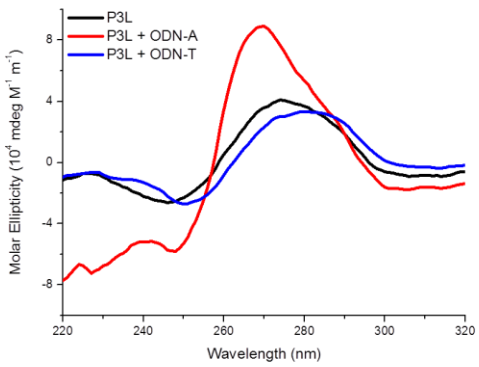
(f)


P4
P4 + ODN-A
P4 + ODN-T
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

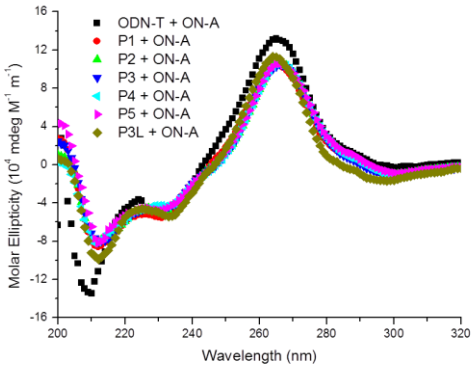
(g)


P5
P5 + ODN-A
P5 + ODN-T
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

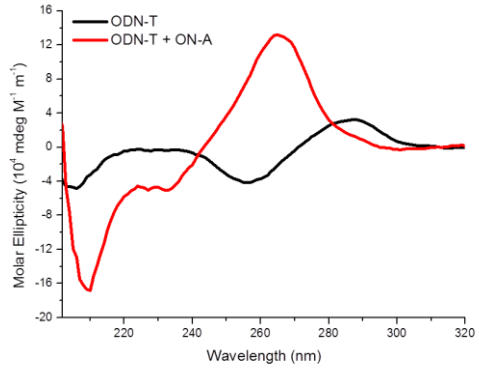
(h)


P3L
P3L + ODN-A
P3L + ODN-T
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

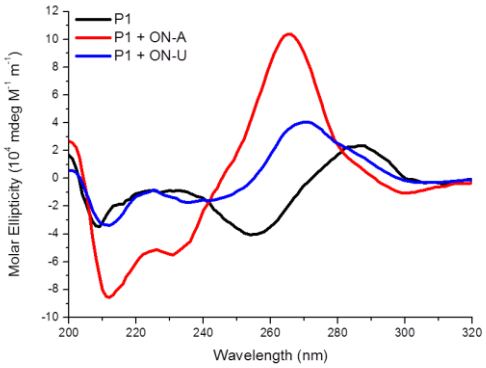
(i)


ODN-T + ON-A
P1 + ON-A
P2 + ON-A
P3 + ON-A
P4 + ON-A
P5 + ON-A
P3L + ON-A
Molar Ellipticity (10⁴ mdeg M⁻¹ m⁻¹)
Wavelength (nm)

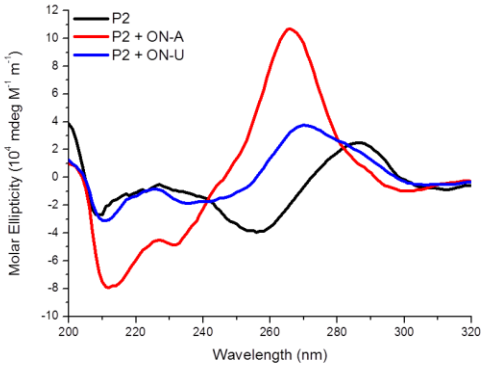
(j)


ODN-T
ODN-T + ON-A
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

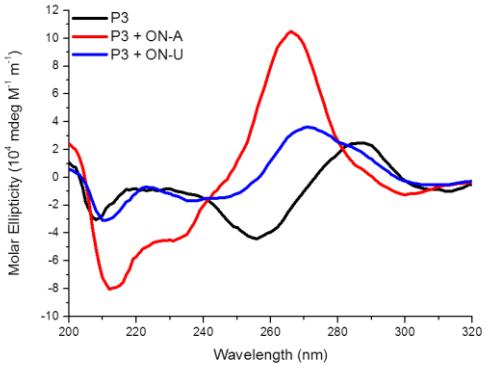
(k)


P1
P1 + ON-A
P1 + ON-U
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

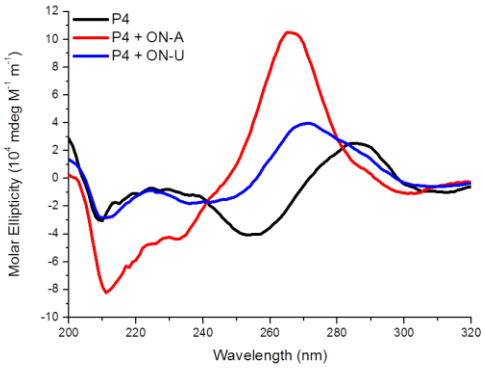
(l)


P2
P2 + ON-A
P2 + ON-U
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

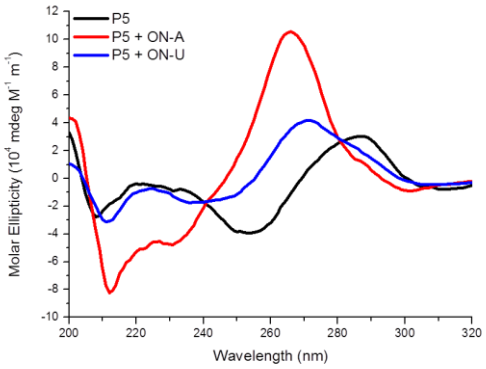
(m)


P3
P3 + ON-A
P3 + ON-U
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

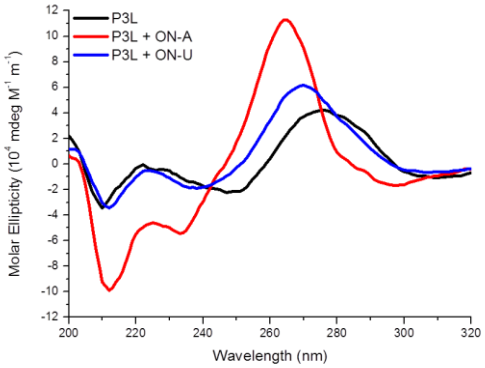
(n)


P4
P4 + ON-A
P4 + ON-U
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

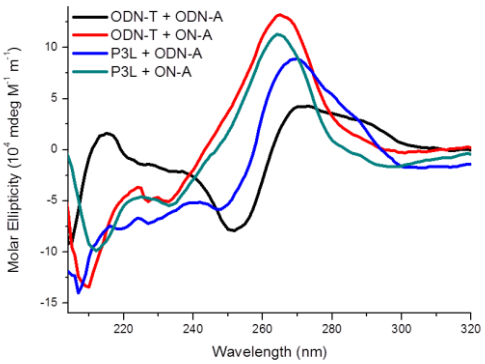
(o)


P5
P5 + ON-A
P5 + ON-U
Molar Ellipticity (10^4 mdeg M^{-1} m^{-1})
Wavelength (nm)

(p)



(q)



(r)

Fig. S3 Polyacrylamide gel electrophoresis (PAGE) results for the sequences **ODN-T**, **P1**–**P5**, and **P3L** (20% non-denaturing PAGE, 100 pmol samples, **L**: size marker confirmed by 5 b.p. ladder; lane 1: **ON-T**; 2: **ON-T** + **ODN-A** or **ON-A**; 3: **P1**; 4: **P1** + **ODN-A** or **ON-A**; 5: **P2**; 6: **P2** + **ODN-A** or **ON-A**; 7: **P3**; 8: **P3** + **ODN-A** or **ON-A**; 9: **P4**; 10: **P4** + **ODN-A** or **ON-A**; 11: **P5**; 12: **P5** + **ODN-A** or **ON-A**; 13: **P3L**; 14: **P3L** + **ODN-A** or **ON-A**).

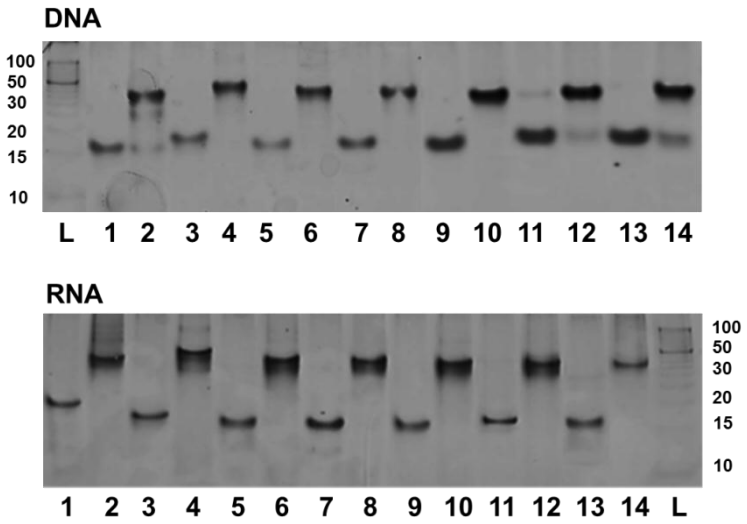
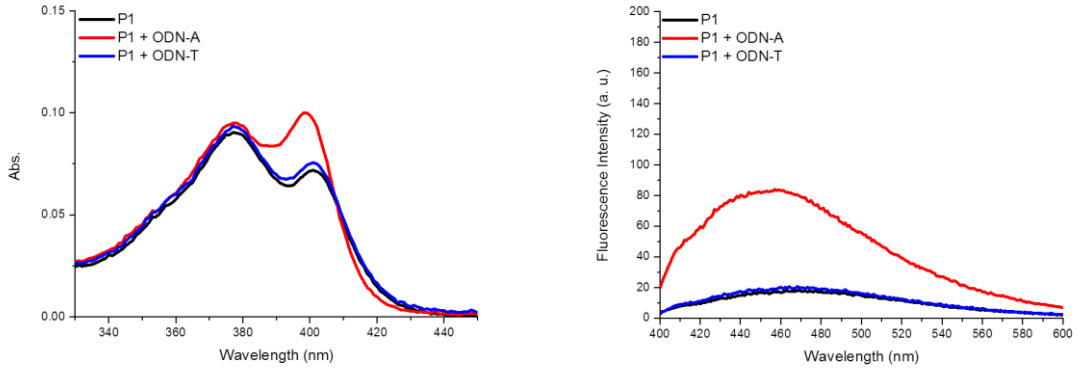
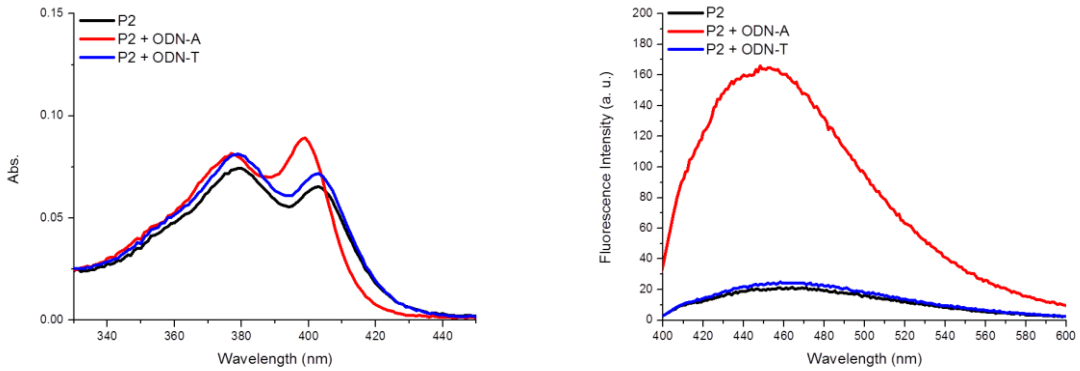


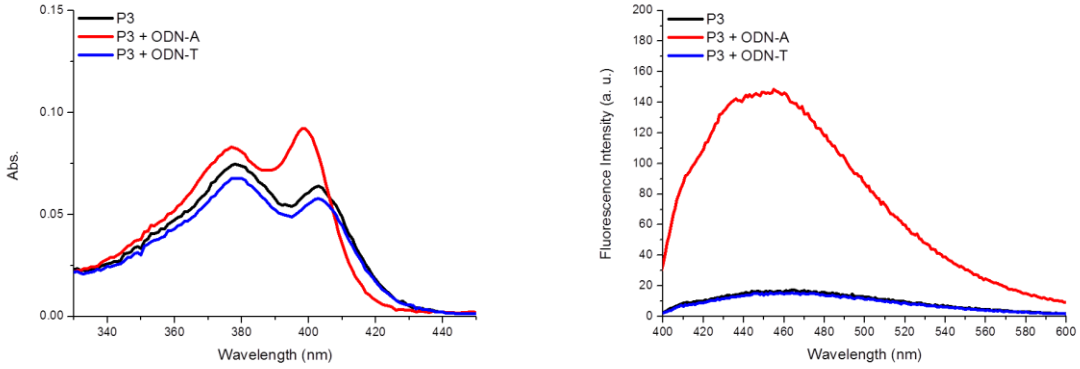
Fig. S4 UV absorption and fluorescence emission spectra of (a, g) **P1**, (b, h) **P2**, (c, i) **P3**, (d, j) **P4**, (e, k) **P5**, and (f, l) **P3L** with (a–f) DNA and (g–l) RNA targets. [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C]



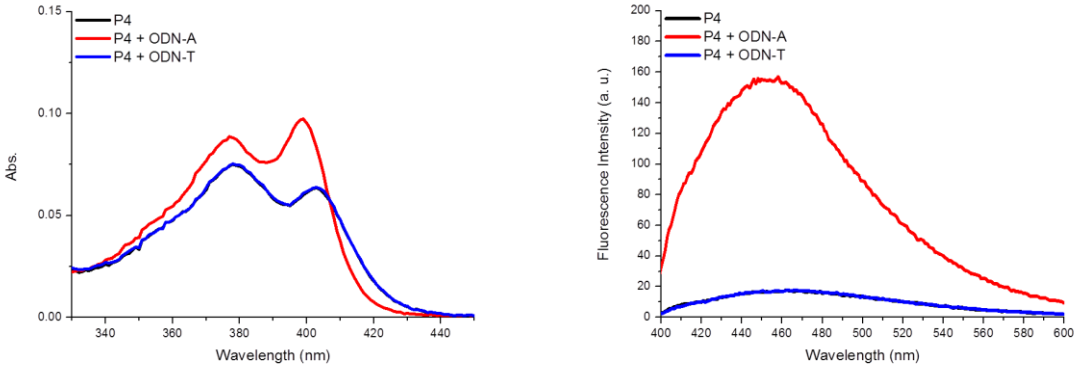
(a) P1 + DNA



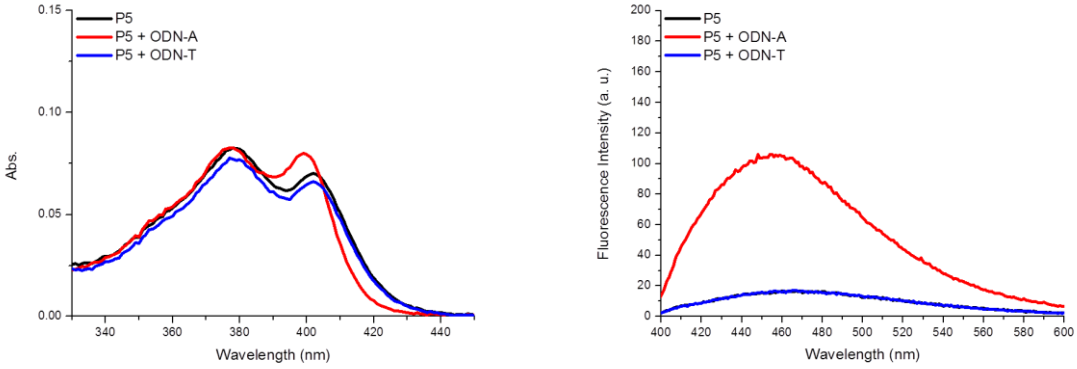
(b) P2 + DNA



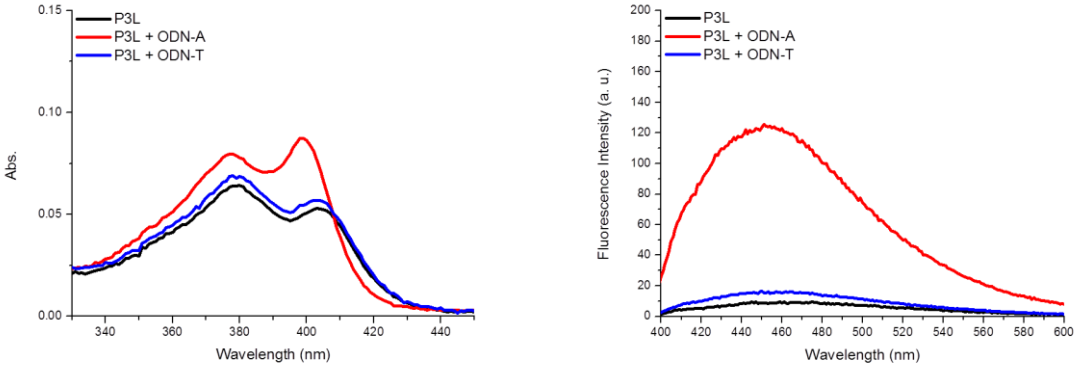
(c) P3 + DNA



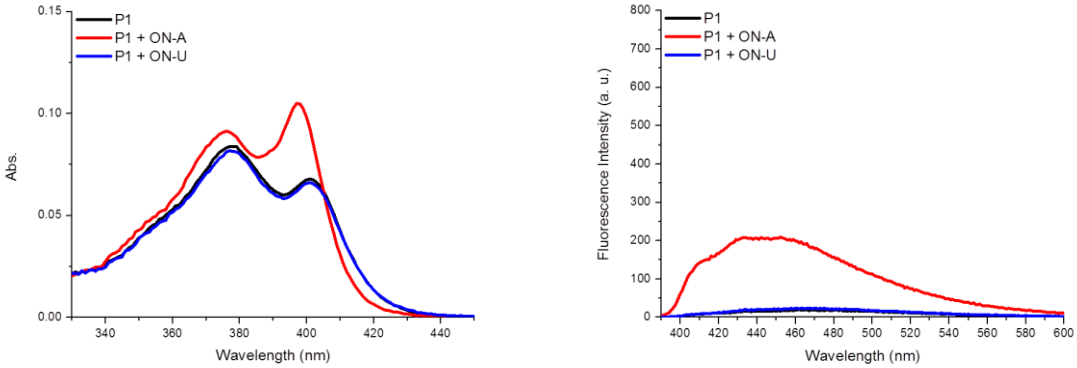
(d) P4 + DNA



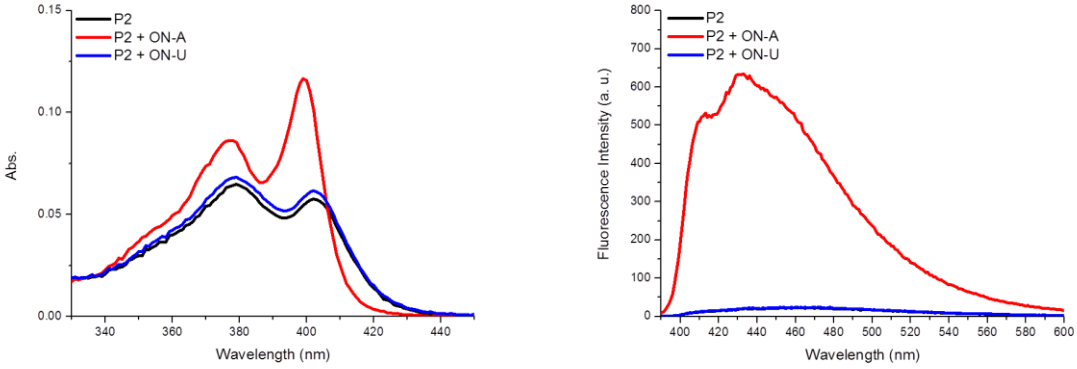
(e) P5 + DNA



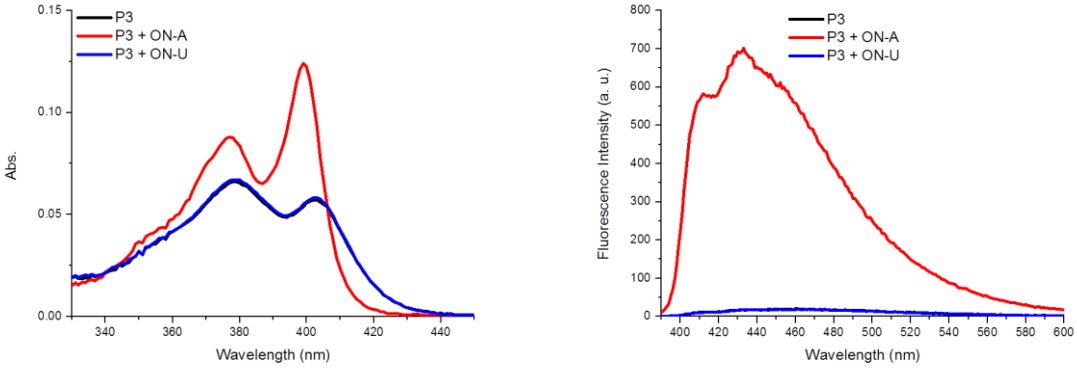
(f) P3L + DNA



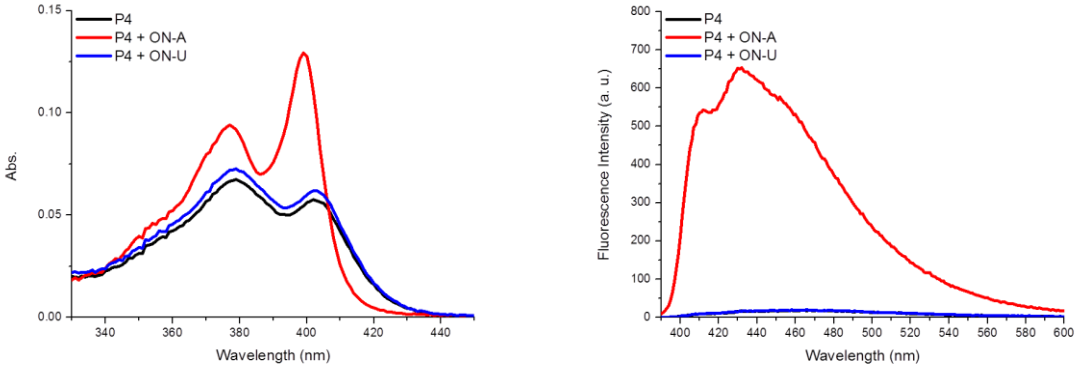
(g) P1 + RNA



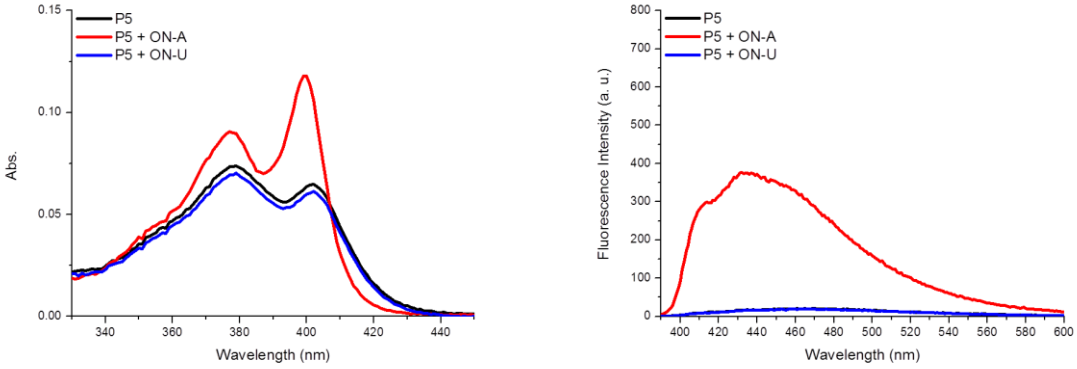
(h) P2 + RNA



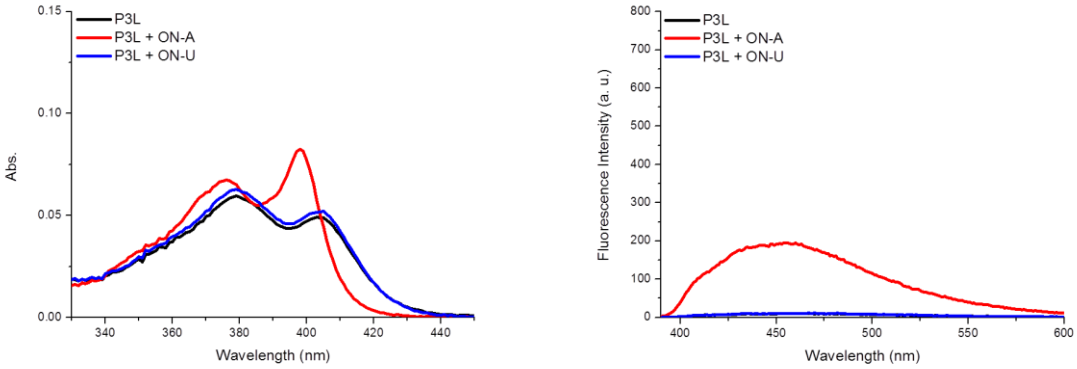
(i) P3 + RNA



(j) P4 + RNA

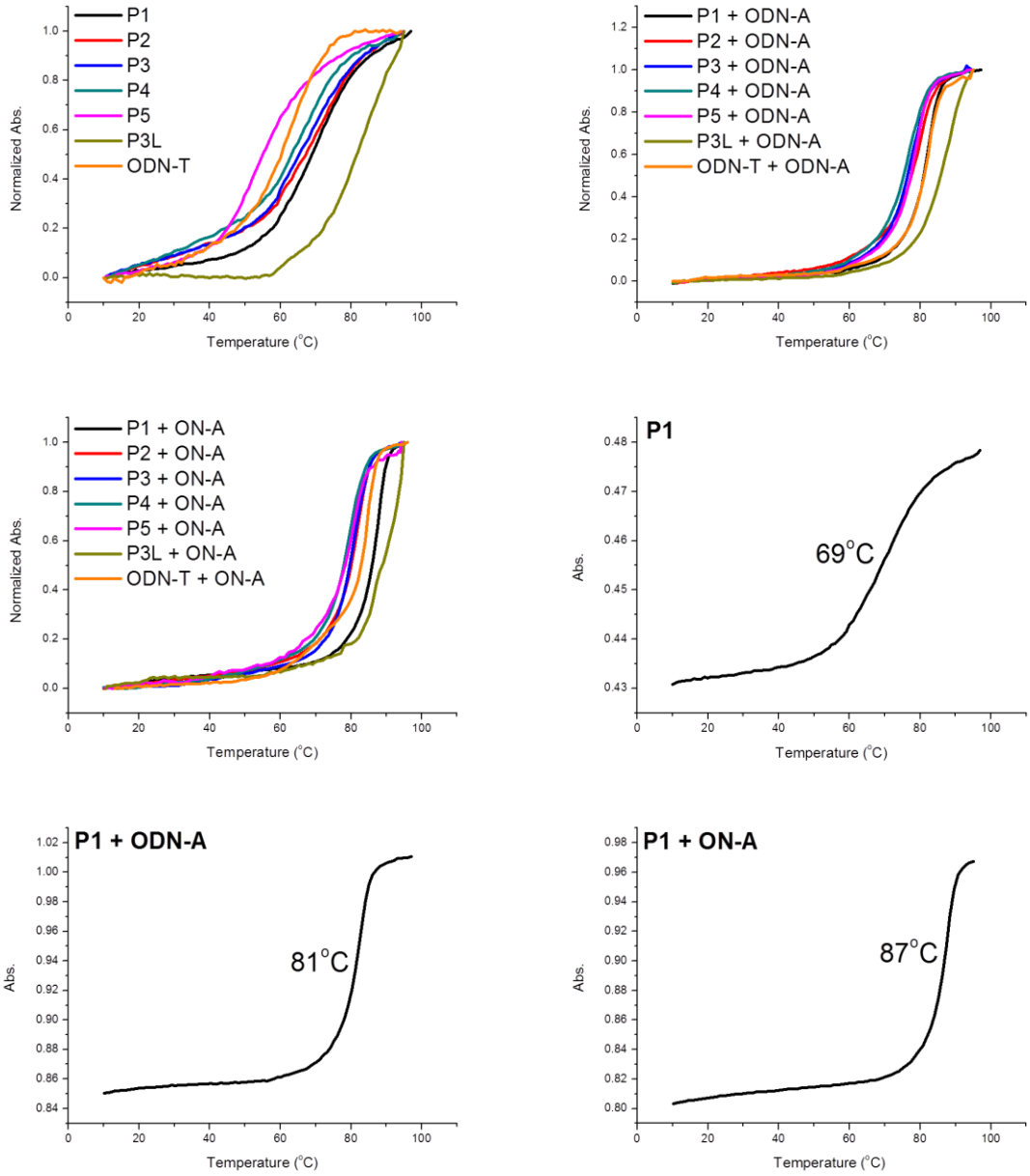


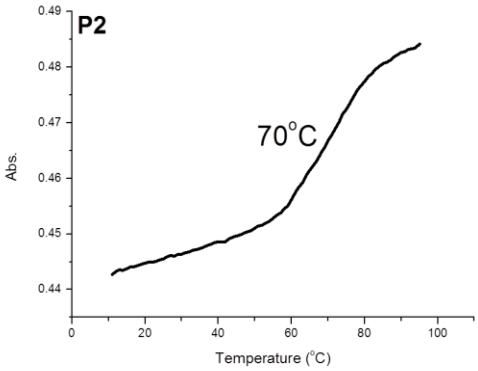
(k) P5 + RNA

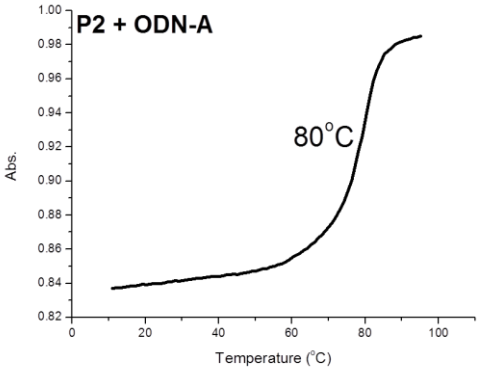


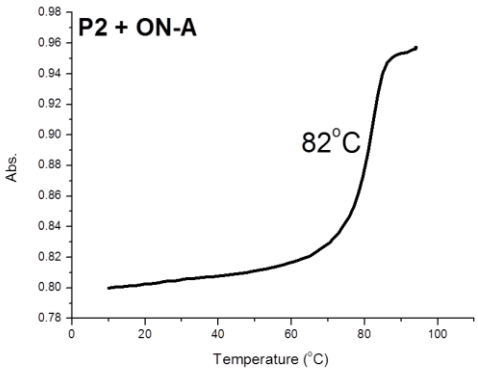
(l) P3L + RNA

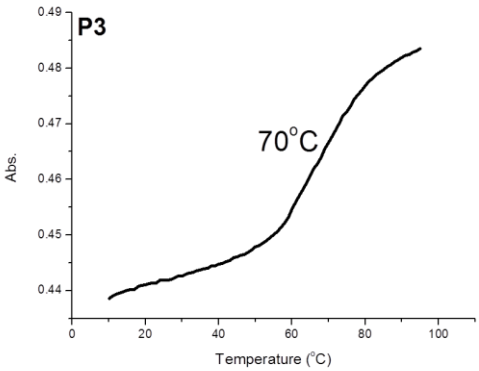
Fig. S5 Thermal stability measurements of all ODN samples. [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$]

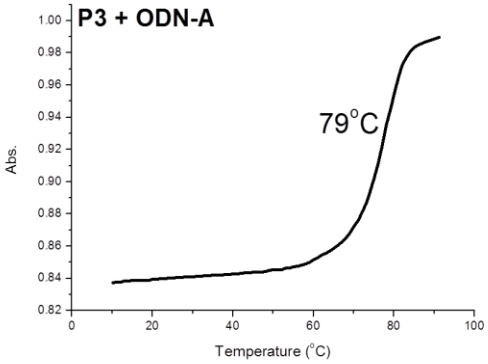


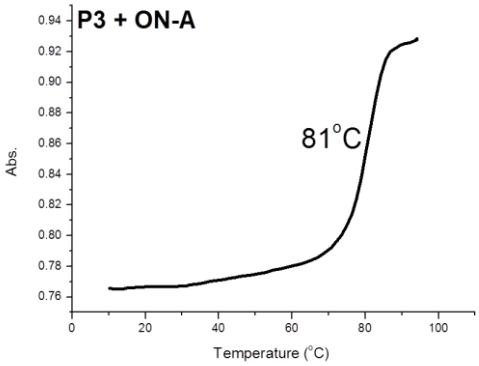

P2
70°C
Abs.
Temperature (°C)

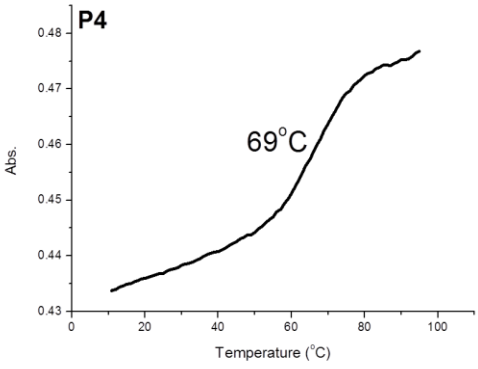

P2 + ODN-A
80°C
Abs.
Temperature (°C)

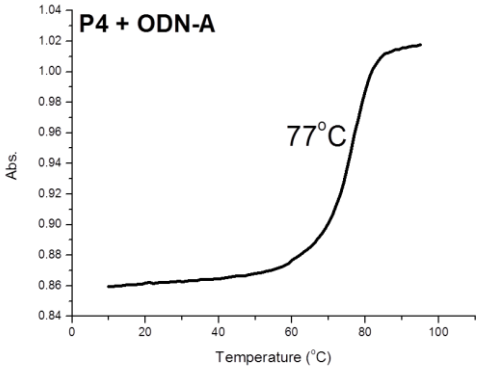

P2 + ON-A
82°C
Abs.
Temperature (°C)

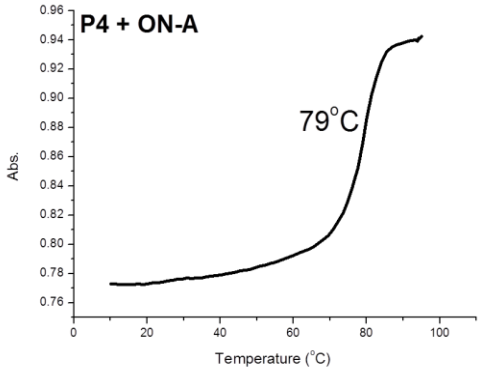

P3
70°C
Abs.
Temperature (°C)

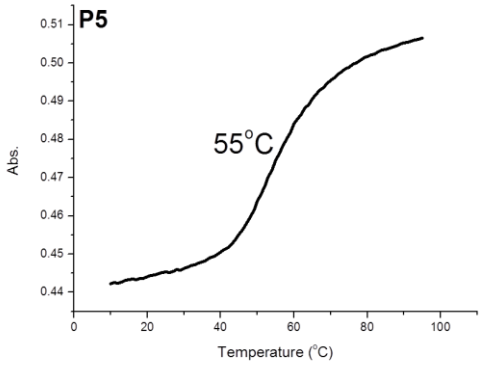

P3 + ODN-A
79°C
Abs.
Temperature (°C)

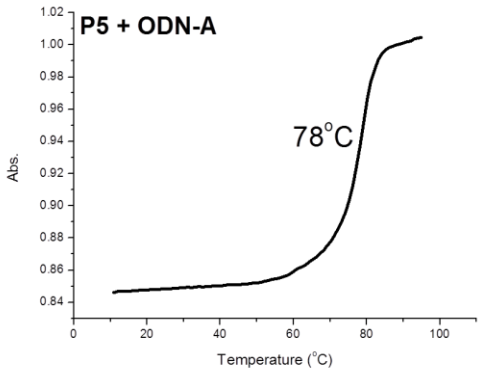

P3 + ON-A
81°C
Abs.
Temperature (°C)

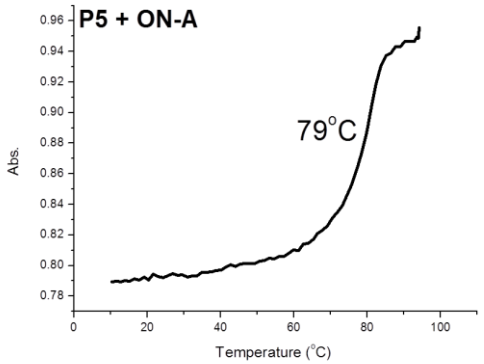

P4
69°C
Abs.
Temperature (°C)

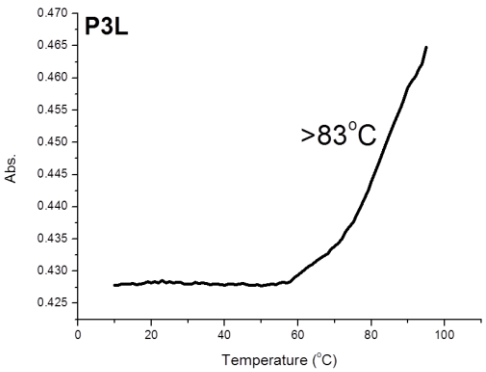

P4 + ODN-A
77°C
Abs.
Temperature (°C)

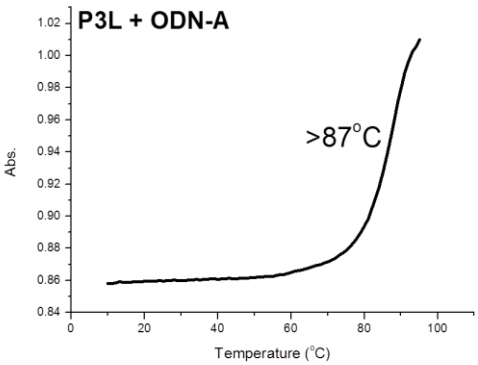


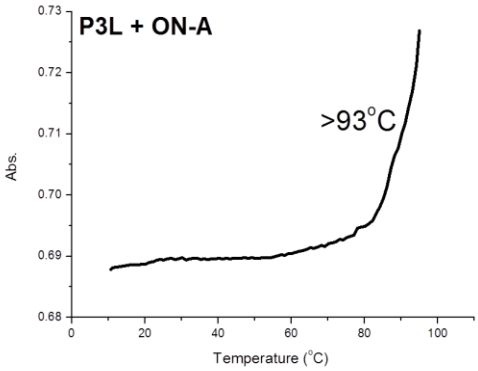


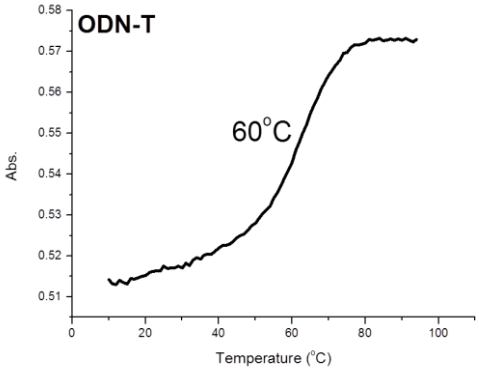


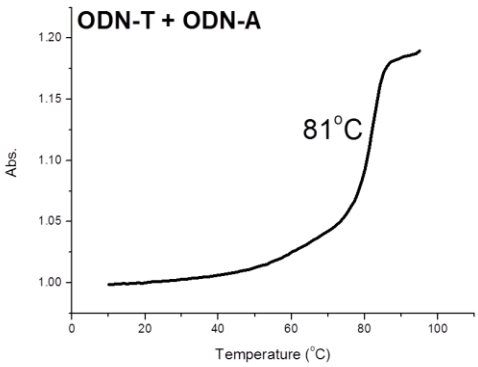







P3L + ON-A
>93°C
Abs.
Temperature (°C)


ODN-T
60°C
Abs.
Temperature (°C)


ODN-T + ODN-A
81°C
Abs.
Temperature (°C)

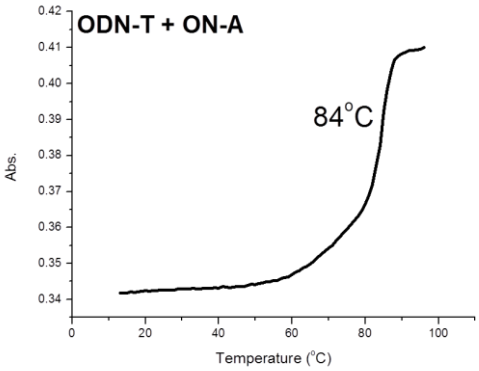

ODN-T + ON-A
84°C
Abs.
Temperature (°C)

Fig. S6 Normalized fluorescence spectra of (a) **P3** and ^{Py}**U** in DMSO or aqueous solution [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$ for aqueous solution, 25 °C] and (b) monomeric ^{Py}**U** in various solvents (inset: relative fluorescence intensities).

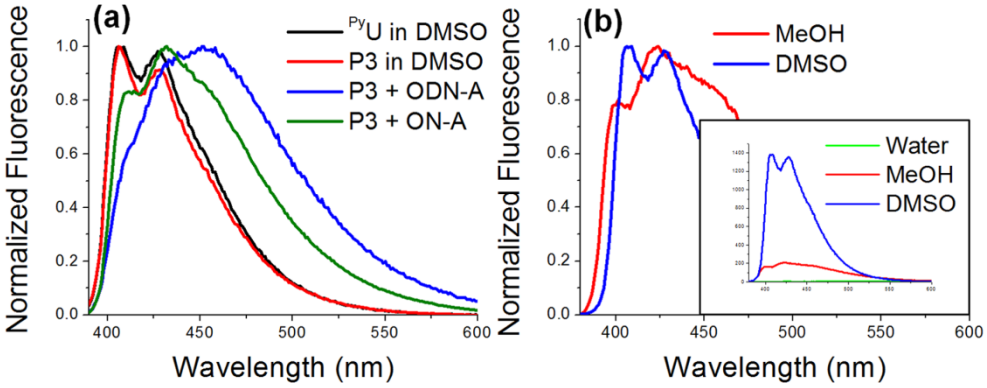


Fig. S7 Quantum yield (QY), brightness (B), and fluorescence enhancement (FE) of Qf-MBs for CAG repeats. In general, these values all decreased upon positioning the ^{Py}**U** units closer to ends of the stem. For **P1**, the QYs were the lowest among our tested systems, originating from the lower stability of the end positions (i.e., so-called "fraying"[5] caused the terminal ^{Py}**U** units to be readily exposed to the aqueous environment, leading to decreased fluorescence). Notably, the fluorescence of ^{Py}**U** units is very sensitive to the nature of the solvent and the pH (involved in proton-coupled intramolecular electron transfer).[6]

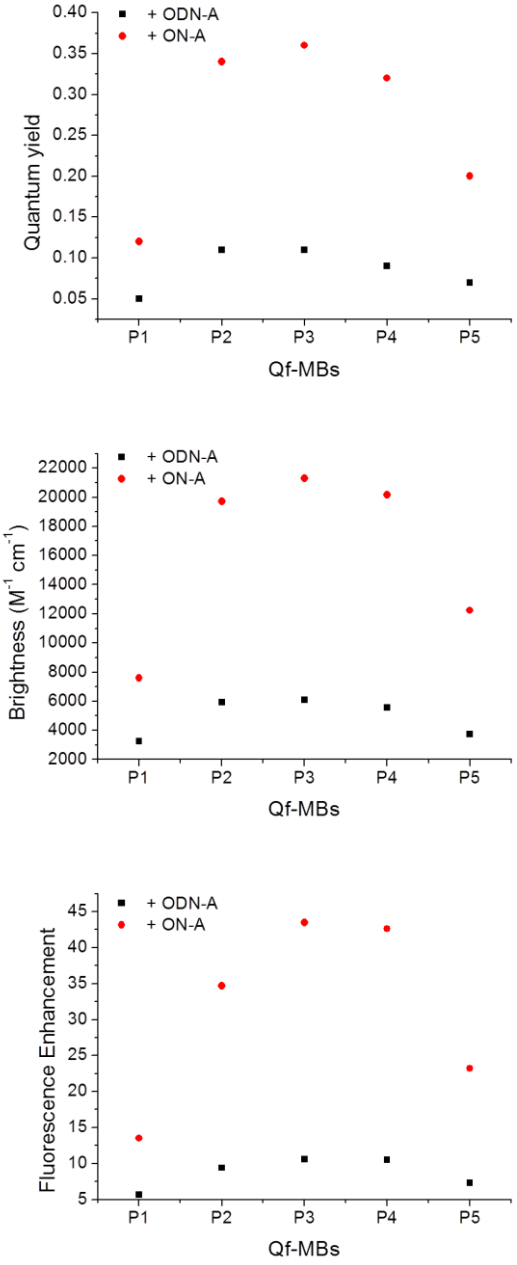


Fig. S8 The results from **P3L** probe. We confirmed that **P3L** also formed DNA duplexes and DNA/RNA hybrids with CAG target sequences, as evidenced from CD spectra and PAGE (Fig. S2 and S3). **P3L** exhibited a lower FE (decreasing from 43.5 to 21.6) in the presence of **ON-A** while it showed the increased FE (from 10.6 to 16) in the presence of **ODN-A** (Table 2, Fig. 3a and 3b in the main text). The LNA units in **P3L** appeared to stabilize its hairpin structure significantly (its melting temperature was 13 °C higher than that of **P3**; Table 2 in main text), despite the presence of several T•••T mismatched pairs. Through this effect, the two ^{Py}**U** units in **P3L** readily stacked together, inducing a lower background signal and a higher FE. However, the LNA/RNA duplex structure may have promoted the formation of ^{Py}**U** exciplexes, inducing a lower FE, despite forming an A-form helix (Fig. S2j). According to normalized fluorescence of the duplex forms of **P3** and **P3L** (figure below), **P3L** also exhibited its major fluorescence band near 450 nm, resulting from the exciplex of its ^{Py}**U** units. [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C]

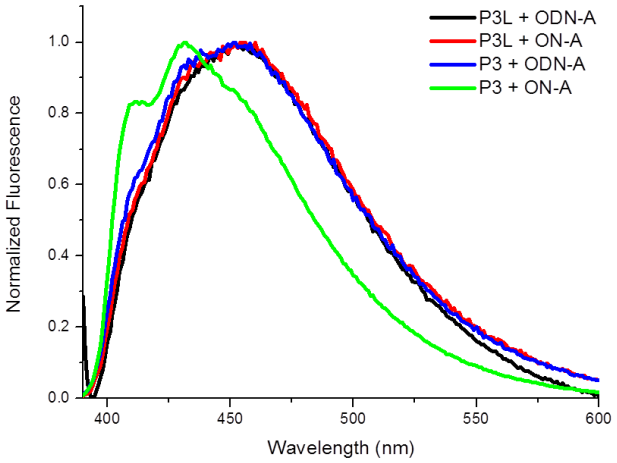


Fig. S9 Time-dependent fluorescence enhancement of the **P3** in the presence of target oligonucleotides. The fluorescence is saturated within 5 min (8.4- and 42.1-fold increase in fluorescence in the presence of ODN-A and ON-A, respectively) at 25 °C. In addition, the **P3** can detect all longer DNA CAG repeat ($(CAG)_{20}$, $(CAG)_{30,}$ and $(CAG)_{40}$) even in ambient conditions. [1.5 μM of sample, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C]

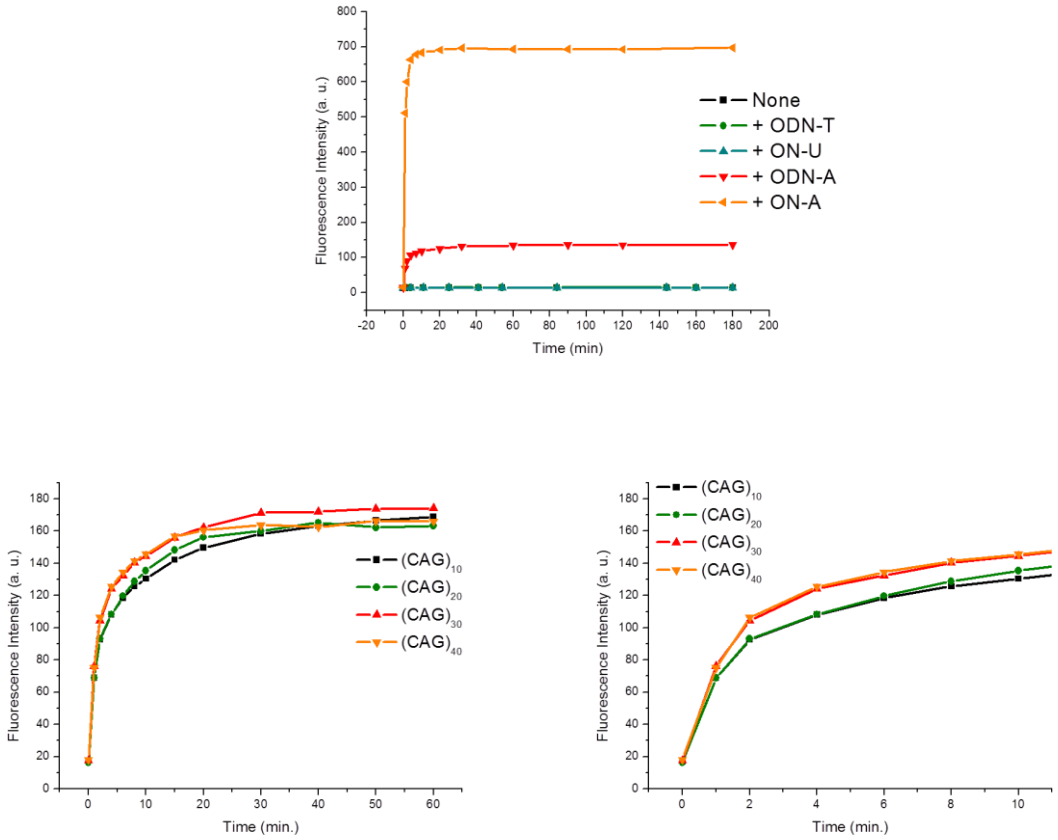
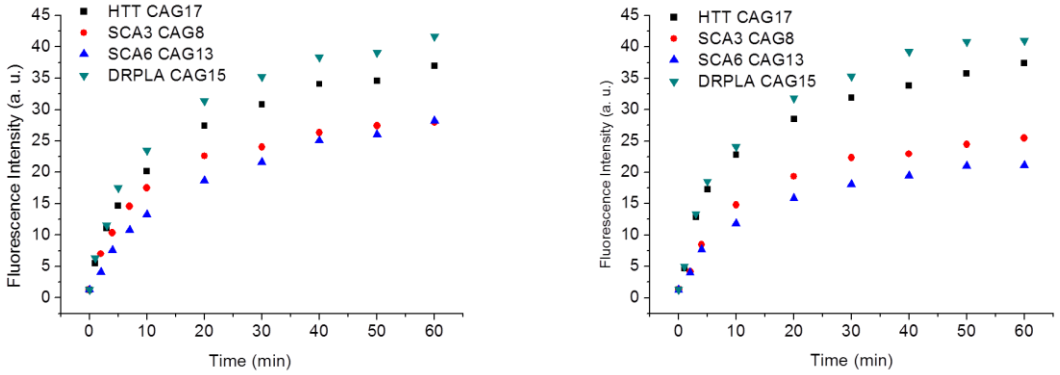


Fig. S10 Time-dependent fluorescence measurement of **P3** in the presence of mRNA sequences from Huntington diseases (HTT CAG17)[7], Spinocerebellar ataxia (SCA3 CAG8 and SCA6 CAG13), and Dentatorubropallidoluysian atrophy (DRPLA CAG15).[8] **P3** exhibited very fast fluorescence enhancement in the presence of both (a) nonannealed and (b) annealed mRNAs. In regardless of annealing, we observed significant fluorescence increase in the presence of target mRNAs. [100 nM of sample, 1 mL; 20 mM trizma buffer (pH 7.2), 80 mM NaCl, 20 mM $MgCl_2$, 37 °C]



(a) (b)

Name	Sequences (5´ to 3´)
HTT CAG17	GAG UCC CUC AAG UCC UUC CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAA CAG CCG CCA CCG CCG (87mer)
SCA3 CAG8	GGC CUA CUU UGA AAA ACA GCA GCA AAA GCA GCA ACA GCA GCA GCA GCA GCA GCA GGG GGC CUA UCA GGA CAG AGU UCA CAU CCA U (88mer)
SCA6 CAG13	AAG GCC GGC GGC UCG GGG CCC CCG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG GCG GUG GCC AGG CCG GGC CGG GCG (87mer)
DRPLA CAG15	CAU CAC CAC CAG CAA CAG CAA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAU CAC GGA AAC UCU GGG CCC (87mer)

References

(1) (a) G. T. Hwang, Y. J. Seo, S. J. Kim and B. H. Kim, *Tetrahedron Lett.*, 2004, **45**, 3543-3546; (b) A. D. Malakhov, E. V. Malakhova, S. V. Kuznitsova, I. V. Grechishnikova, I. A. Prokhorenko, M. V. Skorobogatyi, V. A. Korshun and Y. A. Berlin, *Russ. J. Bioorg. Chem. (Engl. Transl.)* 2000, **26**, 34-44.

(2) J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.

(3) A. M. Brouwer, *Pure Appl. Chem.*, 2011, **83**, 2213–2228.

(4) A. T. R. Williams and S. A. Winfield, *Analyst*, 1983, **108**, 1067-1071.

(5) D. J. Patel and C. W. Hilbers, *Biochemistry*, 1975, **14**, 2651–2656.

(6) (a) M. Rist, N. Amann and H.-A. Wagenknecht, *Eur. J. Org. Chem.*, 2003, 2498–2504; (b) M. Raytchev, E. Mayer, N. Amann, H.-A. Wagenknecht and T. Fiebig, *ChemPhysChem*, 2004, **5**, 706–712; (c) S. T. Gaballah, Y. H. A. Hussein, N. Anderson, T. T. Lian and T. L. Netzel, *J. Phys. Chem. A*, 2005, **109**, 10832–10845.

(7) K. T. Gagnon, H. M. Pendergraff, G. F. Deleavey, E. E. Swayze, P. Potier, J. Randolph, E. B. Roesch, J. Chattopadhyaya, M. J. Damha, C. F. Bennett, C. Montaillier, M. Lemaitre and D. R. Corey, *Biochemistry*, 2010, **49**, 10166-10178.

(8) G. Michlewski and W. J. Krzyzosiak, *J. Mol. Biol.*, 2004, **340**, 665-679.