Supporting information

DNA triplex-based fluorescence turn-on sensors for adenosine using fluorescent molecular rotor 5-(3-Methylbenzofuran-2-yl)deoxyuridine

Takashi Kanamori, Yoshiaki Masaki, Yuki Oda, Hiroki Ohzeki, Akihiro Ohkubo, Mitsuo Sekine and Kohji Seio

Department of Life Science and Technology, Tokyo Institute of Technology, 4259 J2-16 Nagatsuta-cho, Midori-ku, Yokohama, Japan.
Corresponding author's email address: kseio@bio.titech.ac.jp

Contents

Oligonucleotide synthesis ---------- S1

UV-melting temperatures measurements ---------- S4

Fluorescence measurements ---------- S5

Fluorescence spectra of triplexes with or without nucleosides and nucleotides---------- S5

Dissociation constants (K_Ds) calculations ---------- S6

Molecular dynamics (MD) simulations of triplex based seosor ODN-1/FODN-1/TFO-1 and ODN-1/FODN-2/TFO-1 ---------- S7

Oligonucleotide synthesis (FODN-1, FODN-2 and TFO-1)

The phosphoramidite units of dU^{MBF} [1] and dP[2] in which the amino group was protected with phenoxyacetyl group were prepared according to previously reported procedures.[2d] The synthesis of oligodeoxynucleotides was performed on an ABI 392 DNA synthesizer using the standard 1.0 μmol-scale DNA phosphoramidite approach, which consists of detritylation, coupling, capping, and iodine oxidation steps. The synthesis was carried out using a DNA synthesis protocol installed in the synthesizer. The synthesized oligomers were released from CPG supports by treatment with a solution of aq. NH_3, at room temperature for 2 h, then heated at 55 °C for 6 h. The solution was evaporated under reduced pressure at room temperature to remove NH_3, and the residue was diluted with 0.1 M NH_4OAc. The solution was placed on a C18 cartridge column, and the failure sequences were eluted using 10% MeCN/0.1 M NH_4OAc as an eluent. After washing with 0.1 M NH_4OAc and water, the column was treated with aq. 2% TFA to remove the DMTr group and further washed with

water. The target oligodeoxynucleotide was eluted using 20% MeCN/water, and the fractions containing the target were concentrated. The residues were purified by RP HPLC (30 °C with a linear gradient of solvent A (30 mM NH_4OAc buffer) in solvent B (acetonitrile) was used at a flow rate of 1.0 mL/min for 40 min) to afford the pure materials at the yield of 25% of **FODN**-1, 32% of **FODN**-2, respectively. The structures were confirmed by MALDI-TOF mass spectroscopy. MALDI-TOF mass m/z calcd. for **FODN-1** $[M+H]^+$ 4863.9, found 4865.0. MALDI-TOF mass m/z calcd. for **FODN-2** $[M+H]^+$ 4863.9, found 4868.9. MALDI-TOF mass m/z calcd. for **TFO-1** $[M+H]^+$ 3965.7, found 3966.4.

1) T. Kanamori, H. Ohzeki, Y. Masaki, A. Ohkubo, M. Takahashi, K. Tsuda, T. Ito, M. Shirouzu, K. Kuwasako, Y. Muto, M. Sekine and K. Seio, *ChemBioChem*, 2015, **16**, 167.
2) (*a*) P. J. Bates, C. A. Laughton, T. C. Jenkins, D. C. Capaldi, P. D. Roselt, C. B. Reese and S. Neidle, *Nucleic Acids Res.*, 1996, **24**, 4176. (b) S. Hildbrand and C. Leumann, *Angew. Chem. Int. Ed.*, 1996, **35**, 1968. (c) S. A. Cassidy, P. Slickers, J. O. Trent, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Neidle and K. R. Fox, *Nucleic Acids Res.*, 1997, **25**, 4891; (*d*) M. Nakahara, Y. Hari and S. Obika, *Heterocycles*, 2012, **86**, 1135.

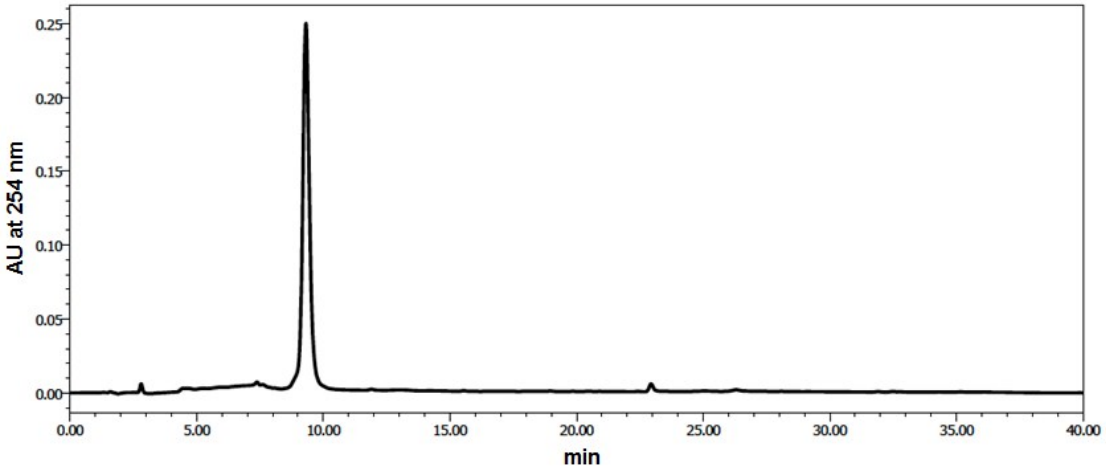

Figure S1. Reversed-phase HPLC profile of **FODN-1** after purification. Conditions: 30 °C with a linear gradient of 0-40% solvent B (acetonitrile) in solvent A (30 mM NH_4OAc buffer) at a flow rate of 1.0 mL/min for 40 min.

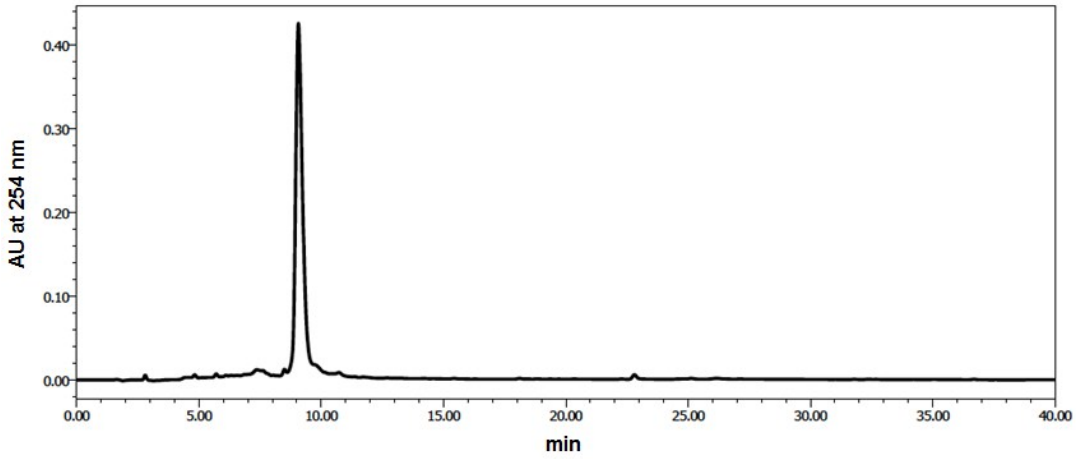

Figure S2. Reversed-phase HPLC profile of **FODN-2** after purification. Conditions: 30 °C with a linear gradient of 0-40% solvent B (acetonitrile) in solvent A (30 mM NH_4OAc buffer) at a flow rate of 1.0 mL/min for 40 min.

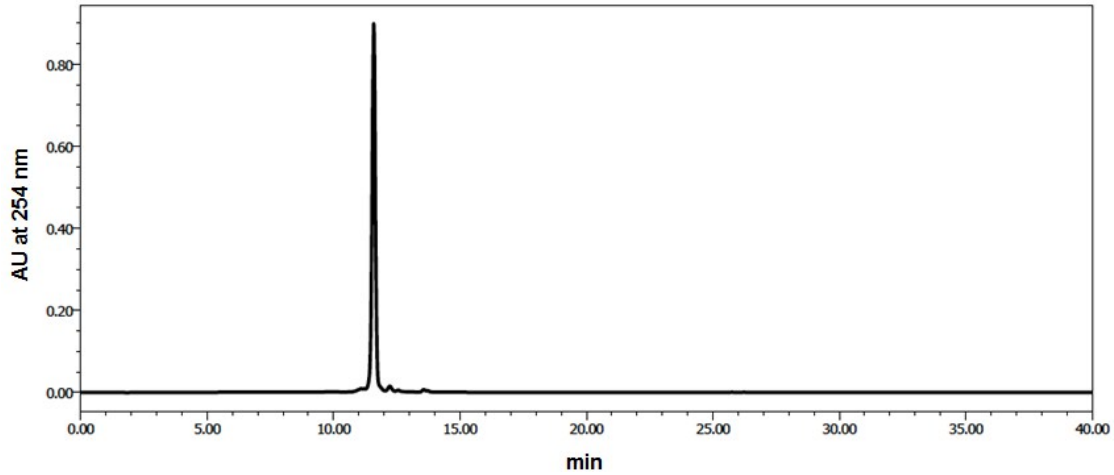

Figure S3. Reversed-phase HPLC profile of **TFO-1** after purification. Conditions: 30 °C with a linear gradient of 0-40% solvent B (acetonitrile) in solvent A (30 mM NH_4OAc buffer) at a flow rate of 1.0 mL/min for 40 min.

UV-melting temperatures measurements

FODN-1 or **FODN-2**, **ODN-1** and **TFO-1** were dissolved in 10 mM sodium cacodylate buffer (pH 7.0) containing 500 mM NaCl, 10 mM $MgCl_2$. The final concentration was 2.0 μM for each oligonucleotide. The concentration of the oligonucleotides were calculated[3] by assuming that the extinction coefficients of dU^{MBF} and dP were identical to those of thymidine and deoxycytidine, respectively. First, the solutions were kept at 80 °C for 30 min, and then cooled to 5 °C at the rate of 0.5 °C/min. Subsequently, the temperature was raised to 80 °C at the same rate during which the UV absorbances at 260 nm were measured at every 1 °C. The data were plotted against temperature and smoothed using Stavitzky-Golay method (5 point). The melting curves were differentiated and the melting temperatures (T_ms) were determined as the temperatures which gave the maximum of each differential curve.

3) Borer,P.N. (1975) Optical properties of nucleic acids, absorption and circular dichroism spectra. In Fasman,G.D. (ed.), Handbook of Biochemistry and Molecular Biology, 3rd Edn. CRC Press, Cleveland, OH, USA, Vol. I, pp. 589-595.

Fluorescence measurements

FODN-1 or **FODN-2**, **ODN-1** and **TFO-1** were dissolved in 10 mM sodium cacodylate buffer (pH 7.0) containing 500 mM NaCl, 10 mM $MgCl_2$. The final concentration was 0.2 μM for each oligonucleotide. The fluorescence measurements were performed at 5 °C in the presence of various concentrations of adenosine with a FP-8300 Fluorescence Spectrometer (JASCO). The excitation wavelength was set to 350 nm.

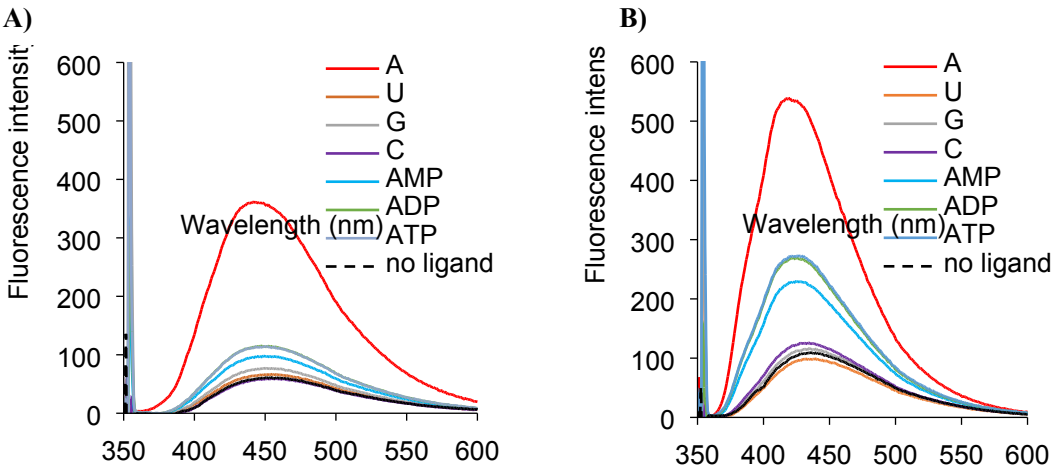

Figure S4. Fluorescence spectra of A) **ODN-1**/**FODN-1**/**TFO-1** triplex and B) **ODN-1**/**FODN-2**/**TFO-1** triplex in the presence and absence of several 20 μM ribonucleosides and adenosine phosphate derivatives. λ_{ex} = 350 nm, 0.2 μM **ODN-1**, **FODN-1**, and **TFO-1** for **ODN-1**/**FODN-1**/**TFO-1** triplex and 0.2 μM **ODN-1**, **FODN-2**, and **TFO-1** for **ODN-1**/**FODN-2**/**TFO-1** triplex, 10 mM cacodylate buffer (pH 7.0), 500 mM NaCl, 10 mM $MgCl_2$, at 5 °C. Each fluorescence spectrum of the triplex in the presence of 20 μM adenosine and no ligand (whithout adenosine) is identical to that in Fig. 4.

Dissociation constants (K_Ds) calculations

Dissociation constants (K_D) between **ODN-1**/**FODN-1**/**TFO-1** triplex and adenosine and between **ODN-1**/**FODN-2**/**TFO-1** triplex and adenosine were determined from the change in fluorescence intensity (shown in Figure 4 and S4) at 440 nm and 420 nm, respectively. The plots shown in Fig. S4 were fitted to the following theoretical equation to obtain the K_D values.[4]

$$\Delta\, Flu = \frac{[TripAde]}{[Trip]_0} = \frac{[Trip]_0 + [Ade]_0 + K_D - \sqrt{([Trip]_0 + [Ade]_0 + K_D)^2 - 4[Trip]_0[Ade]_0}}{2[Trip]_0}$$

Where ΔFlu is normalized fluorescence intensity at 440 nm or 420 nm, [TripAde] is the concentration of adenosine bound to triplex, $[Trip]_0$ is the total concentration of triplex, $[Ade]_0$ is the total concentration of adenosine, and $[TripAde]/[Ade]_0$ means the ratio of adenosine bound to triplex, which corresponds to the normalized fluorescence intensity (ΔFlu).

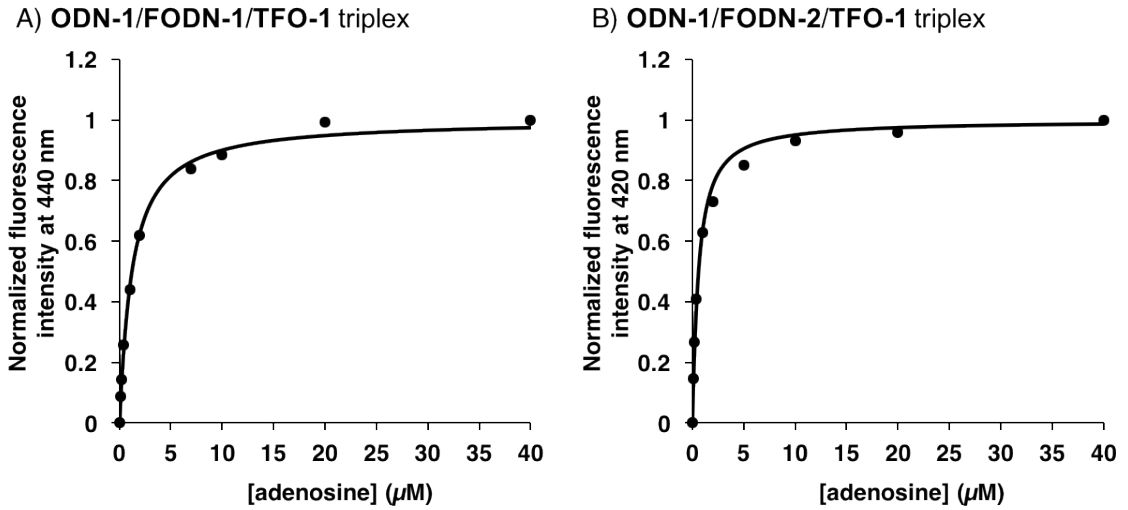

Figure S5. Plot of fluorescence intensity of A) **ODN-1**/**FODN-1**/**TFO-1** triplex and B) **ODN-1**/**FODN-2**/**TFO-1** triplex in the presence of several concentration of adenosine (0, 0.1, 0.2, 0.4, 1, 2, 7, 10, 20, 40 μM for **ODN-1**/**FODN-1**/**TFO-1**, 0, 0.1, 0.2, 0.4, 1, 2, 5, 10, 20, 40 μM for **ODN-1**/**FODN-2**/**TFO-1**). λ_{ex} = 350 nm, 0.2 μM **ODN-1**, **FODN-1**, and **TFO-1** for **ODN-1**/**FODN-1**/**TFO-1** triplex and 0.2 μM **ODN-1**, **FODN-2**, and **TFO-1** for **ODN-1**/**FODN-2**/**TFO-1** triplex, 10 mM cacodylate buffer (pH 7.0), 500 mM NaCl, 10 mM $MgCl_2$, at 5 °C.

4) T. D. Pollard, *Mol Biol Cell*, 2010, **21**, 4061.

Molecular dynamics (MD) simulations of ODN-1md/FODN-1md/TFO-1md and ODN-1md/FODN-2md/TFO-1md triplexes

Protocols for molecular dynamics simulation

Molecular dynamics (MD) simulations were performed using AMBER 14 program package.[5,6] The sequences used in this study were shown in Fig.2. The charges of modified nucleoside such as U^{MBF}, C3 (propylene linker), and P (N2-protonated 2-aminopyridine) were determined by RESP charge fitting method with Gaussian output files (HF/6-31G(d), iop(6/33=2)). As the force fields parameters, the ff14SB AMBER force fields were used and non-canonical ones were taken from GAFF as shown in Fig. S6, S9. The initial triplex structures were constructed by using Pymol and xleap program. The initial structures were solvated in periodic box with a 10.0 Å of explicit water molecules (TIP3PBOX). Na^+ and Cl^- ions were added to give ca. 0.1 M NaCl solution. MD simulations were performed following our previous protocols.[7,8] Further, to study the time course of change in torsion angle between uracil ring and methyl benzofuran ring, these torsion angles were analyzed by cpptraj program from AmberTools 15. For the visualization of 3D structures, Pymol program was used.

5) D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A. Kollman (2015), AMBER 2015, University of California, San Francisco.

6) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179.

7) Y. Masaki, R. Miyasaka, A. Ohkubo, K. Seio and M. Sekine, *J. Phys. Chem. B*, 2010, **114**, 2517.

8) Y. Masaki, R. Miyasaka, K. Hirai, H. Tsunoda, A. Ohkubo, K. Seio and M. Sekine, *Chem. Commun*., 2012, **48**, 7313.

Parameters for molecular dynamics simulation

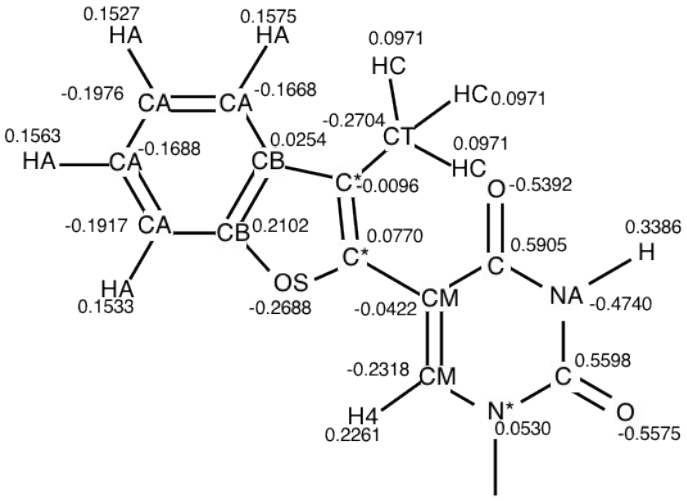

Figure S6. RESP charges and atom types for U^{MBF}.

```
remark goes here
MASS

BOND
CM-C*  346.5    1.490       same as cp-cp
C*-C*  504.0    1.371       same as cc-cd
C*-OS  376.1    1.370       same as cc-os
CB-OS  392.60   1.357       same as c2-os

ANGLE
CM-CM-C*   70.22      117.21    same as c2-c2-cc
CM-C*-C*   69.09      117.02    same as c2-cc-cd
CM-C*-OS   70.10      120.79    same as c2-cc-os
C -CM-C*   67.930     120.700   same as c -c2-c2
C*-C*-CB   69.09      117.02    same as c2-cc-cd
C*-C*-CT   64.81      119.45    same as c3-cc-cd
C*-OS-CB   65.950     113.140   same as c2-os-c2
C*-C*-OS   69.96      120.30    same as cc-cd-os
C*-CB-CB   70.22      117.21    same as c2-c2-cc
CB-CB-OS   71.040     121.890   same as c2-c2-os
OS-CB-CA   71.040     121.890   same as c2-c2-os

DIHE
CM-CM-C*-C*   4    4.000       180.000          2.000      same as X -cp-cp-X obtained from gaff
CM-CM-C*-OS   4    4.000       180.000          2.000      same as X -cp-cp-X obtained from gaff
CM-C*-C*-CB   4   16.000       180.000          2.000      same as X -cc-cd-X statistic value of parm94
CM-C*-C*-CT   4   16.000       180.000          2.000      same as X -cc-cd-X statistic value of parm94
CM-C*-OS-CB   2    2.100       180.000          2.000      same as X -c2-os-X parm99
C -CM-C*-C*   4    4.000       180.000          2.000      same as X -cp-cp-X obtained from gaff
C -CM-C*-OS   4    4.000       180.000          2.000      same as X -cp-cp-X obtained from gaff
C*-OS-CB-CB   2    2.100       180.000          2.000      same as X -c2-os-X parm99
C*-OS-CB-CA   2    2.100       180.000          2.000      same as X -c2-os-X parm99
C*-C*-OS-CB   2    2.100       180.000          2.000      same as X -c2-os-X parm99
CB-C*-C*-OS   4   16.000       180.000          2.000      same as X -cc-cd-X statistic value of parm94
OS-C*-C*-CT   4   16.000       180.000          2.000      same as X -cc-cd-X statistic value of parm94

IMPROPER
CM-H4-CM-N*         1.1          180.0        2.0          General improper torsional angle (2 general atom types)
C -C*-CM-CM         1.1          180.0        2.0          Using default value
CM-NA-C -O         10.5          180.0        2.0          General improper torsional angle (2 general atom types)
C -C -NA-H          1.0          180.0        2.0          General improper torsional angle (2 general atom types)
N*-NA-C -O         10.5          180.0        2.0          General improper torsional angle (2 general atom types)
C*-CM-C*-OS         1.1          180.0        2.0          Using default value
C*-CB-C*-CT         1.1          180.0        2.0          Using default value
C*-CA-CB-CB         1.1          180.0        2.0          Using default value
CA-CB-CB-OS         1.1          180.0        2.0          Using default value
CA-CB-CA-HA         1.1          180.0        2.0          General improper torsional angle (2 general atom types)
CA-CA-CA-HA         1.1          180.0        2.0          General improper torsional angle (2 general atom types)

NONBON
```

Figure S7. Non-canonical force fields for U^{MBF} (taken from GAFF).

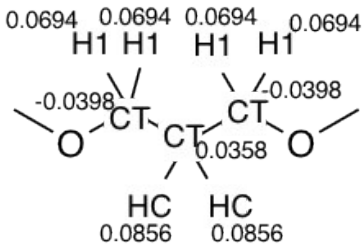

Figure S8. RESP charges and atom types for C3.

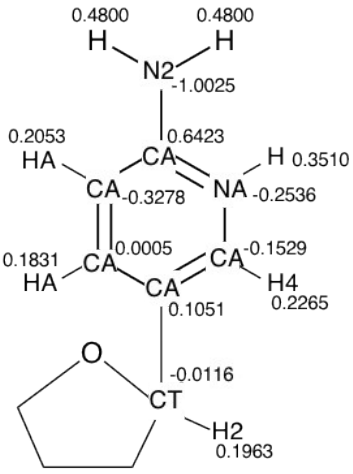

Figure S9. RESP charges and atom types for P (2-aminopyridine).

```
remark goes here
MASS

BOND

ANGLE
OS-CT-CA   68.45      108.480   same as c2-c3-os
H2-CT-CA   47.030     110.490   same as c2-c3-hc
CA-CA-NA   69.830     121.380   same as c2-c2-na
CA-CA-N2   69.830     121.380   same as c2-c2-na
CA-NA-CA   67.80      110.370   same as c2-na-c2
CA-N2-CA   67.80      110.370   same as c2-na-c2
NA-CA-H4   51.18      112.420   same as ha-c2-na

DIHE

IMPROPER

NONBON
```

Figure S10. Non-canonical force fields for P (2-aminopyridine) taken from GAFF.

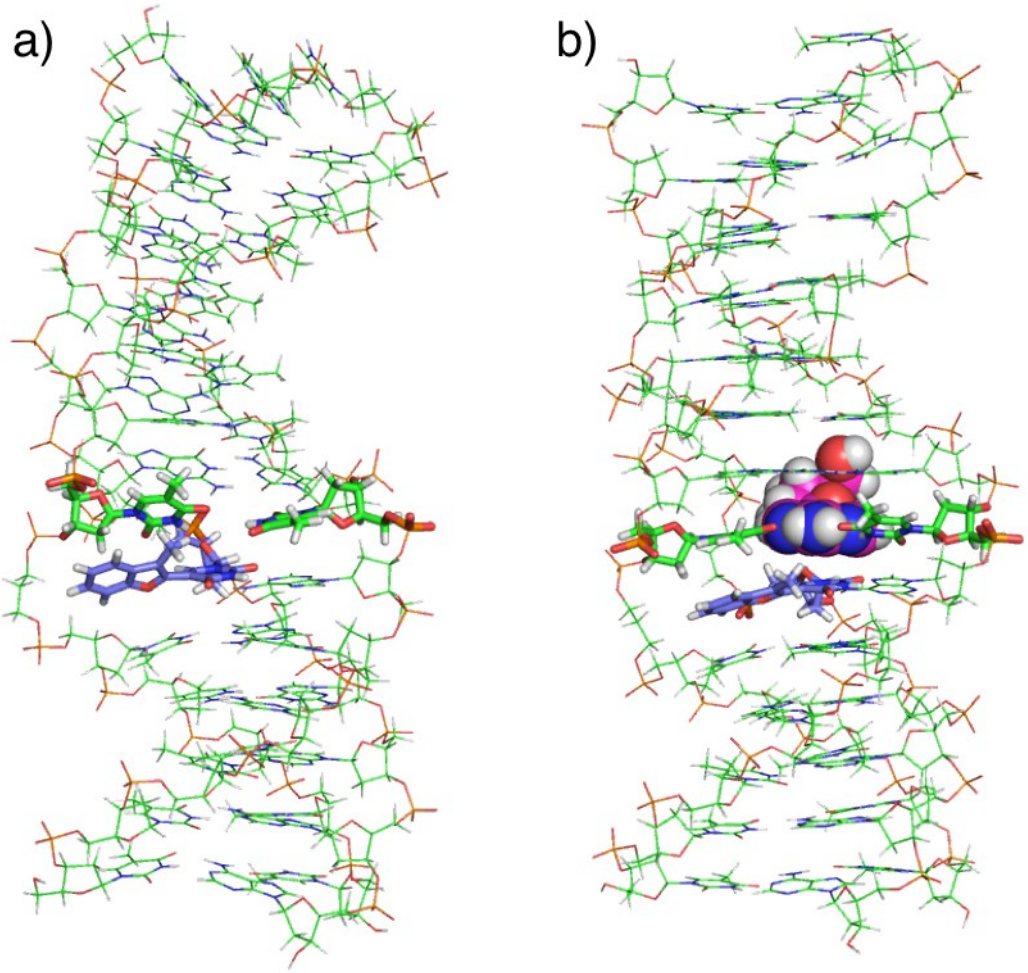

Figure S11. Full triplex structures of **ODN-1md/FODN-1md/TFO-1md** obtained after optimization of the last snapshot of 50 ns MD simulations a) in the absence and b) presence of adenosine.

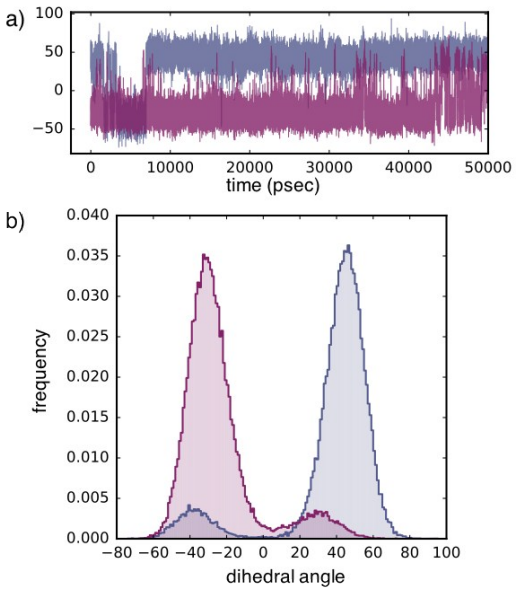

Figure S12. The time course a) plots and b) histograms of the dihedral angle between uracil and benzofuran ring during the 50 ns MD simulations of **ODN-1md/FODN-1md/TFO-1md** in the presence (purple line) and absence (gray line) of adenosine.

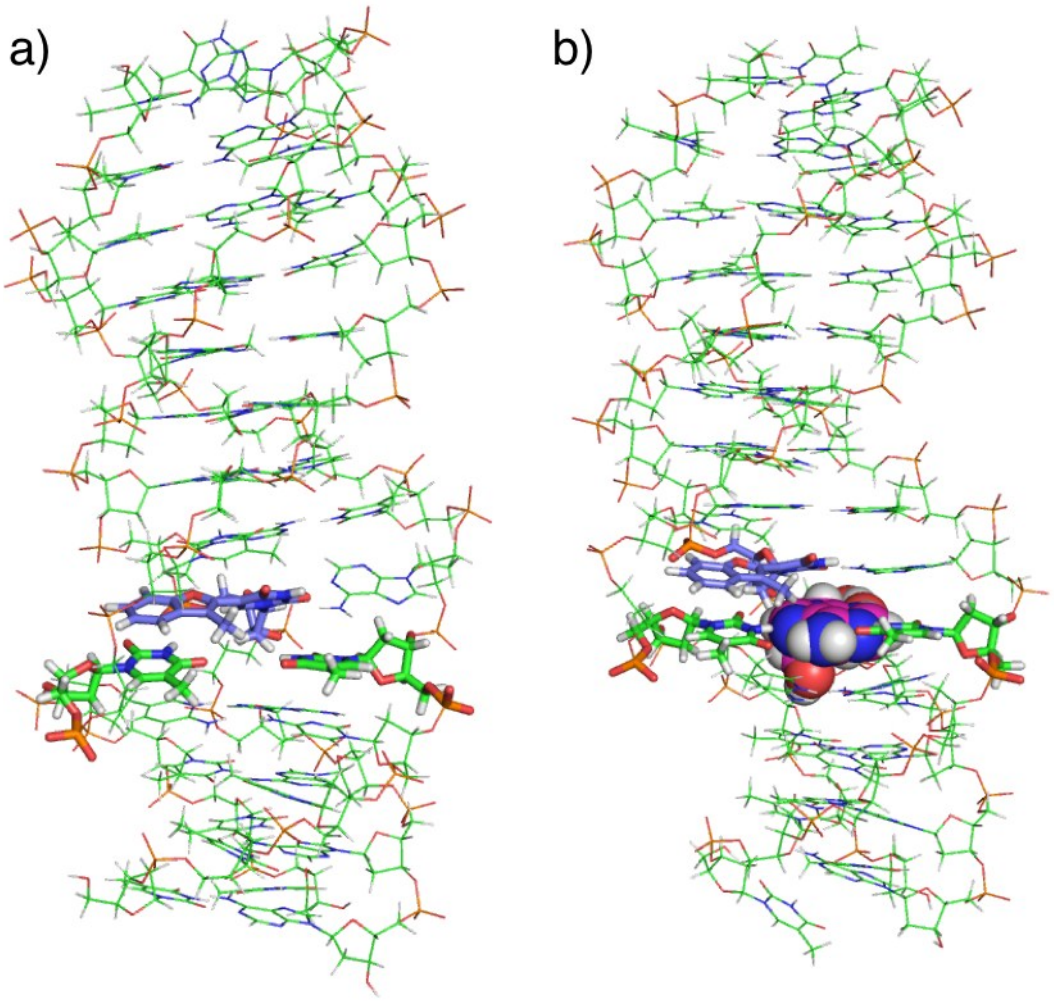

Figure S13. Full triplex structures of **ODN-1md/FODN-2md/TFO-1md** obtained after optimization of the last snapshot of 50 ns MD simulations a) in the absence and b) presence of adenosine.

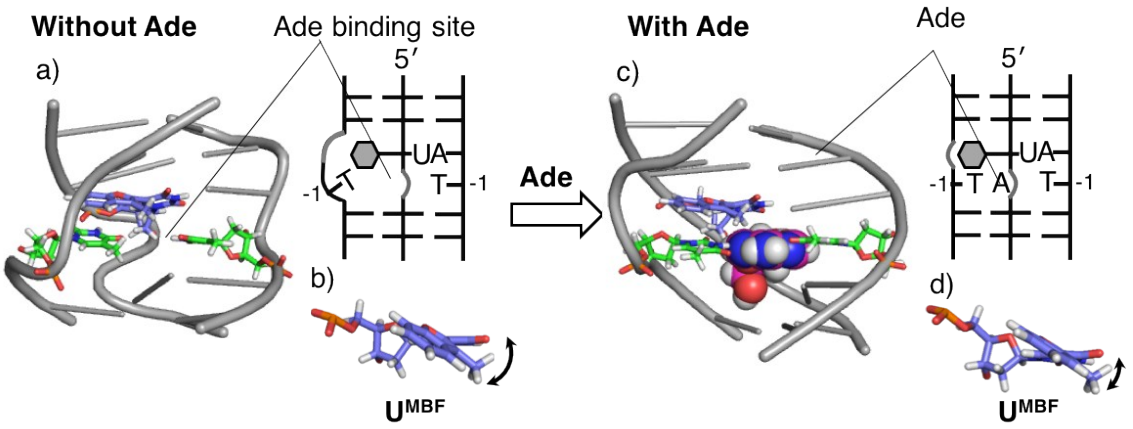

Figure S14. 50 ns MD simulations of **ODN-1md/FODN-2md/TFO-1md** in the presence a), d) and absence c), d) of adenosine. a) and c) are corresponding extracted structures from full triplexes obtained after optimization of the last snapshot of 50 ns MD simulations. b) and d) are U^{MBF} residues extracted from a) and c), respectively. All MD simulations were performed using AMBER 14 programs. See Fig. S12 for full triplex structures.

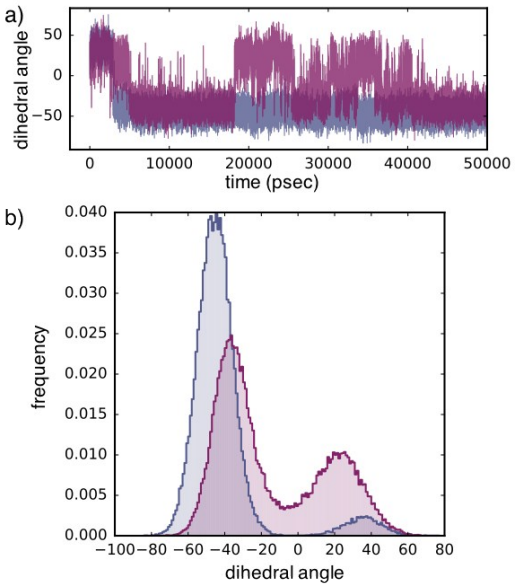

Figure S15. The time course a) plots and b) histograms of the dihedral angle between uracil and benzofuran ring during the 50 ns MD simulations of **ODN-1md/FODN-2md/TFO-1md** in the presence (purple line) and absence (gray line) of adenosine.