

## Supporting Information

### **Folding-upon-Repair DNA Nanoswitches for Monitoring the Activity of DNA Repair Enzymes**

*Nada Farag<sup>+</sup>, Rosanna Mattossovich<sup>+</sup>, Rosa Merlo, Łukasz Nierzwicki, Giulia Palermo, Alessandro Porchetta,\* Giuseppe Perugino,\* and Francesco Ricci\**

anie\_202016223\_sm\_miscellaneous\_information.pdf

## Supporting Information

## Table of Contents

<b>Experimental Procedures .....</b>	<b>3</b>
Reagents and Materials .....	3
Oligonucleotides .....	3
DNA constructs, protein expression and purification .....	4
SDS-PAGE gel-imaging AGT assay .....	4
Fluorescence measurements .....	4
Fluorescence data analysis .....	5
Molecular Dynamics (MD) simulations .....	6
Simulation protocol.....	6
Single-stranded DNA folding to triplex conformation .....	6
Conventional MD simulations and analysis .....	7
Free energy simulation.....	7
<b>Supplementary Figures .....</b>	<b>9</b>
<b>Figure S1.</b> Relative FRET emission plots.....	9
<b>Figure S2.</b> Melting curve experiments.....	10
<b>Figure S3.</b> Urea denaturation experiments .....	11
<b>Figure S4.</b> Circular Dichroism spectral analysis .....	12
<b>Figure S5-S8.</b> Top and side-views of Methylated Triplex nanoswitches .....	13-16
<b>Figure S9.</b> Analysis of G <sub>4</sub> :C <sup>+</sup> <sub>32</sub> Hoogsteen interaction in un-methylated Triplex in protonated and deprotonated states .....	17
<b>Figure S10.</b> Opening distribution of G <sub>4</sub> :C <sup>+</sup> <sub>32</sub> Hoogsteen interaction in triplex nanoswitches.....	18
<b>Figure S11.</b> Distance distribution of G <sub>4</sub> :C <sup>+</sup> <sub>32</sub> Hoogsteen interaction in triplex nanoswitches .....	19
<b>Figure S12.</b> Analysis of G <sub>7</sub> :C <sup>+</sup> <sub>35</sub> Hoogsteen interaction in non-methylated Triplex in protonated and deprotonated states.....	20
<b>Figure S13.</b> Free energy simulation of the triplex nanoswitches.....	21
<b>Figure S14.</b> SDS-PAGE gel-imaging of Triplex switches vs. SNAP-Vista Green® .....	22
<b>Figure S15.</b> Time-course analysis of the enzymatic reaction.....	23
<b>Figure S16.</b> Analysis of hMGMT repair activity in serum .....	24
<b>Tables S1-S4.</b> DNA base pair parameters along with MD simulations .....	25-28
<b>Tables S5-S8.</b> Topologies and parameters for protonated cytosine and O6-methylated guanine.....	29-32
<b>References.....</b>	<b>33</b>

---

**Author contributions..... 34**

## Experimental Procedures

### Reagents and Materials

All reagent-grade chemicals, including Na<sub>2</sub>HPO<sub>4</sub>, NaCl, Phosphate buffered saline, Dimethyl sulfoxide (DMSO) (all from Sigma-Aldrich, Italy), were used as received. SNAP-Vista Green<sup>®</sup> fluorescent substrate was from New England Biolabs (Ipswich, MA). Synthetic oligonucleotides were from Eurofins (Milan, Italy); *Pfu* DNA polymerase was from NZYTech (Portugal). The Bio-Rad protein assay kit (Bio-Rad Pacific) was used for the determination of the protein concentration, using purified BSA as standard. The inhibitor *O*<sup>6</sup>-benzyl-guanine (*O*<sup>6</sup>-BG) was from Sigma-Aldrich, (Italy). The 4-azido-N-(4-(hydroxymethyl) benzyl) butanamide (BGN3) and the (N-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-4-azidobutanamide (BGSN) were synthesized by Prof. Alberto Minassi (University of Piemonte Orientale, Novara). Lomeguatrib (5 mg) was obtained as a dry white powder from Sigma-Aldrich, St. Louis, Missouri.

### Oligonucleotides

Oligonucleotides and DNA-based receptors employed in this work were synthesized, labelled, and purified (HPLC and reverse-phase) by IBA GmbH (Göttingen, Germany) and used without further purification. Unless otherwise stated the labeled oligonucleotides were dissolved in phosphate buffer at a concentration of 100 μM. The final concentration of the oligonucleotides was confirmed using Tecan Infinite M200pro (Mannedorf, Switzerland) through the NanoQuant Plate. The sequences of the DNA constructs are reported below.

### FRET-labelled DNA nanoswitches:

Triplex nanoswitch

5'-AAG GAA GAA G **TTT**(Cy-3) CTT CTT CCT T CTTTG *TTC CTT CTT C*(Cy-5)-3'

1-Me Triplex nanoswitch

5'-AAG GAA (*O*<sup>6</sup>-Me-G)AA G **TTT**(Cy-3) CTT CTT CCT T CTTTG *TTC CTT CTT C*(Cy-5)-3'

2-Me Triplex nanoswitch

5'-AAG (*O*<sup>6</sup>-Me-G)AA (*O*<sup>6</sup>-Me-G)AA G **TTT**(Cy-3) CTT CTT CCT T CTTTG *TTC CTT CTT C*(Cy-5)-3'

For all the sequences above, the bases in bold represent the loop of the duplex portion and the underlined bases represent the loop of the parallel triplex region. The bases in italics represent the triplex-forming portion.

**DNA constructs, protein expression and purification**

The commercial hMGMT cDNA (from a glycerol stock MGC Human MGMT Sequence-Verified Dharmacon) was cloned in the Quiagen pQE31<sup>TM</sup> (Hilden, Germany) as previously described for SsOGT and its relative mutant H5,<sup>[1]</sup> by using specific synthesized primers (Eurofins): MGMT-sense 5'-AATGATGGATCCAATGGACAAGGATTGTG-3' and MGMT-antisense 5'-TTCGATCAAGCTTATCAGTTTCGGCCAG CAGG 3' which possess the BamH I and Hind III (New England Biolabs, Ipswich, MA) cloning sites, respectively. The insertion of fragment (628 bp) in the pQE31 vector was in frame and downstream of a hexahistidine tag, for the subsequent purification procedures.

The cloning procedures for E. coli Ada-C gene followed a different strategy. The genomic DNA of E. coli strain DH5 $\alpha$  was used as the DNA template for the gene amplification by using primers Lig5-AdaC (5'-TCAGCAAGGGCTGAGGGCCATGGCGGCTAAACAATTCC-3') and Lig3-AdaC (5'-CCTCAGCGGAAGCTGAGGTTACCTCTCCTCATTTCAGC-3'). A DNA fragment of 580 bp was obtained and directly cloned into the expression vector pHTP1 (NZYTech, Portugal). The ligation mixture was entirely used to transform commercial E. coli DH5 $\alpha$  cells (NZY5 $\alpha$  Competent Cells- NZYTech, Portugal). For both the cloning procedures positive clones were confirmed by DNA sequencing. All AGT proteins were expressed and purified by following a previously described protocol.<sup>[1,2]</sup>

**SDS-PAGE *gel-imaging* AGT assay**

This method is used for the evaluation of AGTs activity with nanoswitches by using a fluorescein derivative of the O<sup>6</sup>-BG (SNAP-Vista Green<sup>®</sup>) as substrate. In each sample, 5.0  $\mu$ M of protein (0,1 mg/mL) was incubated at a specific temperature (30°C for hMGMT and Ada-C; 40°C for SsOGT and SsOGT-H<sup>5</sup>) with 10  $\mu$ M of each nanoswitch in a total volume of 10  $\mu$ L of Reaction Buffer (50 mM phosphate, 150 mM NaCl; pH 6.5) for 60 min. Then, SNAP-Vista Green<sup>®</sup> was added to the solution at a final concentration of 5  $\mu$ M and incubated again at the same temperature and time. Samples were boiled and directly loaded on 15% acrylamide SDS-PAGE. Bands were detected by direct *gel-imaging* using the VersaDoc 4000<sup>TM</sup> system (Bio-Rad), performing a double acquisition by applying as excitation/emission parameters a BLUE LED/530 and GREEN LED/605 bandpass filters, for the determination of fluorescent-labeled proteins and Cy5-based triplexes, respectively. Then, gels underwent *Coomassie staining* for the determination and correction of the protein amount loaded.

**Fluorescence measurements**

Fluorescence measurements were carried out on a Cary Eclipse Fluorimeter (Varian), setting excitation wavelength to  $\lambda_{ex} = 530$  nm ( $slit_{ex} = 5$  nm) and acquisition between 545 and 700 nm ( $slit_{em} = 10$  nm) using quartz cuvettes of microvolume (100  $\mu$ L). All measurements were performed at T = 25 °C in 50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer, 250 mM NaCl at pH 5.0. For detection of methyltransferase activity, nanoswitches were first diluted in 50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer, 150 mM NaCl at pH 7.5 to a concentration

of 1  $\mu\text{M}$ . Then in a 10  $\mu\text{L}$  solution of 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 150 mM NaCl, we prepared an enzymatic reaction mixture of 0.5  $\mu\text{M}$  nanoswitches in the absence and presence of 5  $\mu\text{M}$  AGTs (0.1  $\mu\text{g}/\mu\text{l}$ ). The reaction mixtures were incubated at a certain temperature and time for each enzyme; (60 minutes at 30°C for hMGMT and *E. coli* Ada-C) and (30 minutes at 70 °C for the thermostable SsOGT and SsOGT-H<sup>5</sup>). All the mixtures were heat-inactivated after the enzymatic reaction by incubating at 70 °C for 2 minutes (when using the thermostable SsOGT and SsOGT-H<sup>5</sup> heat inactivation was performed at 90 °C for 10 minutes). Fluorescence measurements were conducted at 25 °C by diluting the reaction mixtures to 100  $\mu\text{l}$  using 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 250 mM NaCl at pH 5.0. Test of inhibitor activity was performed by first dissolving all the inhibitors in DMSO to a concentration of 10 mM. Inhibition reactions were conducted in an equimolar concentration of hMGMT (0.125  $\mu\text{g}/\mu\text{l}$ ) and different inhibitors (5  $\mu\text{M}$ ) at 30 °C for 60 minutes in a 10  $\mu\text{l}$  solution of 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 150 mM NaCl at pH 7.5. 0.5  $\mu\text{M}$  of 2-Me triplex nanoswitch was then added to the reaction mixtures to interact with the remaining active hMGMT and incubated for another 60 minutes at 30 °C. The reaction mixture was heat-inactivated after the enzymatic reaction by incubating at 70 °C for 2 minutes. Fluorescence measurements were conducted at 25 °C by diluting the reaction mixtures to 100  $\mu\text{l}$  using 50 mM  $\text{Na}_2\text{HPO}_4$  buffer, 250 mM NaCl at pH 5.0.

### Fluorescence data analysis

The ratiometric FRET has been calculated as follows:

$$\text{Rat. FRET} = F_{\text{Cy5}} / F_{\text{Cy3}}$$

Where  $F_{\text{Cy5}}$  is the maximum fluorescence emission of Cy5 ( $\lambda_{\text{em}} = 670 \text{ nm}$ ) and  $F_{\text{Cy3}}$  is the maximum fluorescence emission of Cy3 ( $\lambda_{\text{em}} = 565 \text{ nm}$ ). The pH titration curves were obtained by plotting

$$\text{Triplex Fraction} = \text{Ratiometric FRET} + \left( \frac{[\text{H}^+] * (\text{Rat. FRET}_{\text{Triplex}} - \text{Rat. FRET}_{\text{Duplex}})}{[\text{H}^+] * K_D} \right)$$

Rat.FRET vs pH, and fitting the data with the following Langmuir-type equation:

Where  $\text{Rat.FRET}_{\text{Triplex}}$  and  $\text{Rat.FRET}_{\text{Duplex}}$  represent the FRET signal of the Triplex Switch in the triplex (closed, pH = 5.0) and duplex state (open, pH = 8.5), respectively, and where  $[\text{H}^+]$  represents the total concentration of hydrogen ions and  $K_D$  is the observed acid constant for the switch.

Inhibition percentage for different inhibitors was calculated as follows:

$$\% \text{inhibition} = \left[ 1 - \frac{(\text{Rat. FRET}_{\text{inh}} - \text{Rat. FRET}_{2\text{-Me Triplex}})}{(\text{Rat. FRET}_{2\text{-Me Triplex, unmet hylated}} - \text{Rat. FRET}_{2\text{-Me Triplex}})} \right] \%$$

Where  $\text{Rat.FRET}_{\text{inh}}$  is the FRET signal of the 2-Me triplex nanoswitch measured after the pre-incubation of different inhibitors with (0.1  $\mu\text{g}/\mu\text{l}$ ) hMGMT at pH 5.0,  $\text{Rat.FRET}_{2\text{-Me triplex, unmethylated}}$ , and

Rat.FRET<sub>2-Me Triplex</sub> represent the FRET signal of the 2-Me Triplex switch at pH = 5.0) after the incubation with (0.1 µg/µl) hMGMT and in the absence of hMGMT at pH = 5.0, respectively.

## Molecular Dynamics (MD) simulations

### Simulation protocol

All simulated model systems were composed of a single DNA molecule sequence 5'–AAGGAAGAAG [TTT] CTTCTTCCTT [CTTTG] TTCCTTCTTC–3' placed in 8 nm x 8 nm x 13 nm rectangular box, solvated with ~27,250 water molecules and sodium ions to neutralize the system. Based on this sequence, four triplex molecules were considered, composed of (a) conventional DNA bases; (b) protonated cytosines (at positions shown in italics and underlined in the sequence above); (c) protonated cytosines and O<sup>6</sup>-methylated guanine at position 7; (d) protonated cytosines and O<sup>6</sup>-methylated guanines at position 4 and 7 (bold, underlined). The initial structure of single-stranded DNA was obtained using the x3DNA package.<sup>[3]</sup> The DNA has been represented with Amber ff12SB force field, which includes the ff99bsc0 corrections for DNA.<sup>[4,5]</sup> This is a well-established force field for nucleic acid, which we have successfully employed in several recent articles.<sup>[6,7]</sup> The TIP3P model was used to represent water.<sup>[8]</sup> The electrostatic potential around each of the modified nucleobases (protonated cytosine and O<sup>6</sup>-methylated guanine) were computed from the geometries optimized at HF/6-31G\* level of theory with Merz-Kollman ESP fitting with Gaussian 09.<sup>[9]</sup> Next, this potential was used to compute atomic partial charges of the bases with Amber 16 package.<sup>[10]</sup> Topologies and parameters for both protonated cytosine and O<sup>6</sup>-methylated guanine are available in SI (Tables S5-S8). MD simulations were performed with Gromacs 5 package<sup>[11]</sup> with Plumed 2.2.3 plugin.<sup>[12]</sup> The simulations were performed in NPT ensemble with the temperature kept at 310 K with v-rescale thermostat<sup>[13]</sup> and pressure at 1 bar with Parrinello-Rahman algorithm.<sup>[14]</sup> Periodic boundary conditions were applied and the particle mesh Ewald algorithm<sup>[15]</sup> was used to compute long-range electrostatic interactions with a real-space cut-off of 1 nm. All covalent bonds including hydrogen were restrained with the LINCS algorithm.<sup>[16]</sup> The leap-frog verlet algorithm was used to integrate equations of motion with a time step of 2 fs.

### Single-stranded DNA folding to triplex conformation

The structure of the DNA triplex was generated through *de-novo* design over ~60 ns of steered MD simulation of single-stranded DNA (ssDNA) with deprotonated cytosines. As a reference structure, the example of triplex topology from x3DNA was used. DNA folding was performed in four consecutive steps: initially, the root mean squared deviation (RMSD) with respect to the duplex structure was used as a reaction coordinate for heavy atom positions of nucleobases 1 to 10 and 14 to 23. During the first ~20 ns of the simulation, the center of a one-sided harmonic potential was moved with a constant speed from the initial value of the reaction coordinate (2.2 nm) to 0 nm with a force constant gradually increased from 500 kJ nm<sup>-2</sup> to 7000 kJ nm<sup>-2</sup>. At the same time, the RMSD to the initial position of

nucleobase heavy atoms of residues 14 to 38 was restrained using a harmonic biasing potential with a force constant of  $7000 \text{ kJ nm}^{-2}$ . Next, the restraint for nucleobases 14 to 38 was removed and RMSD with respect to the nucleobase heavy atom position in the triplex structure was used for residues 1 to 10, 14 to 23 and 29 to 31 was used as a new reaction coordinate. In the next 8 ns of the simulation, the center of a one-sided harmonic potential was moved with a constant speed from the initial value of the reaction coordinate (2.3 nm) to 0 nm with a force constant gradually increased from  $500 \text{ kJ nm}^{-2}$  to  $7000 \text{ kJ nm}^{-2}$ . This step was further repeated for residues 1 to 10, 14 to 23, 29 to 34 and for residues for residues 1 to 10, 14 to 23, 29 to 38. Finally, additional 16 ns of simulation with restraint was produced in order to relax the position of the DNA backbone. The comparison of the final structure obtained with steered MD simulations to the reference triplex structure from x3dna resulted in RMSD for the nucleobase heavy atom position below 0.08 nm.

### Conventional MD simulations and analysis

The initial structures of the protonated and methylated triplex switches were obtained by substituting the modified bases in the final structure of deprotonated triplex and subsequent minimization and 1 ns long simulation, during which RMSD of the initial position of nucleobase heavy atoms for residues 1 to 10, 14 to 23 and 29 to 38 was restrained using a harmonic biasing potential with a force constant of  $7000 \text{ kJ nm}^{-2}$ . For each of the considered systems, 3 replicas of 500 ns conventional MD simulation were performed. DNA structural parameters from the trajectories were computed using do\_x3dna plugin<sup>[17]</sup> to VMD. All molecular images were created using VMD.<sup>[18]</sup> The graphs were prepared using matplotlib library of python.<sup>[19]</sup>

### Free energy simulations

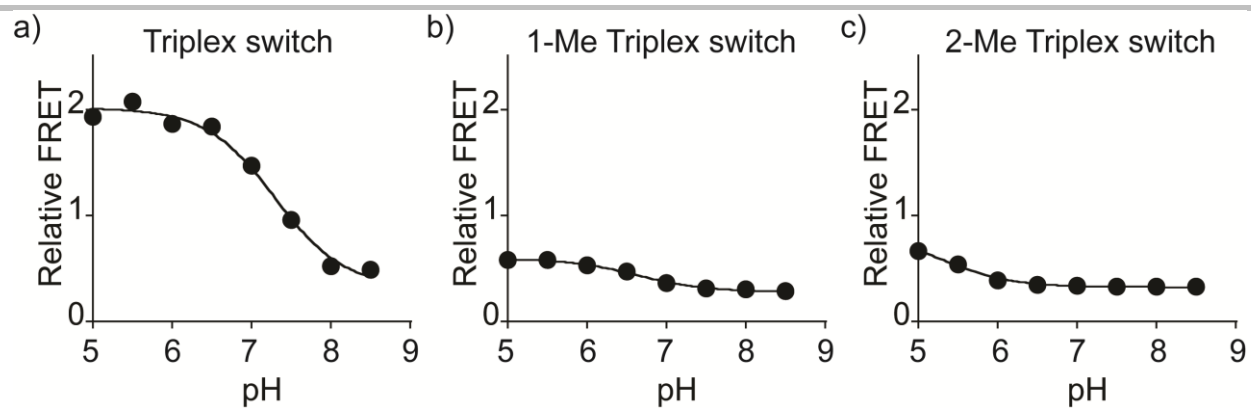
The free energy profiles for the triplex to duplex transition were computed in four considered systems (i.e., protonated unmethylated triplex; protonated 1-Me triplex; protonated 2-Me triplex; deprotonated unmethylated triplex) using replica-exchange umbrella sampling (REUS).<sup>[20]</sup> We used as a reaction coordinate (RC) the distance between two center of mass (COM): (i) the heavy atoms of 3'-terminal cytidine phosphate and (ii) the nucleobases that co-form the triplex plane with the terminal cytosine (panel A, Figure S13). The initial configurations for REUS simulations were obtained by performing ~600 ns-long steered-MD simulations, in which the distance between terminal cytidine and the nucleobases was increased from 1.15 nm to 8.35 nm using a one-sided harmonic potential with a force constant of  $2500 \text{ kJ nm}^{-2}$ . The reaction coordinate was divided into 38 windows, with 0.1 nm spacing in the region of the RC from 1.15 to 1.35 nm, and with 0.2 nm spacing in the remaining region of the RC. A force constant of  $1000 \text{ kJ nm}^{-2}$  was used to restrain the RC at the given distance. 500 ns long trajectories were obtained for each REUS window. For each window, the first ~100 ns were discarded from the analysis (being part of the system's equilibration) and the final free energy profiles were obtained considering the equilibrated runs using the weighted histogram analysis method (WHAM).<sup>[21]</sup>



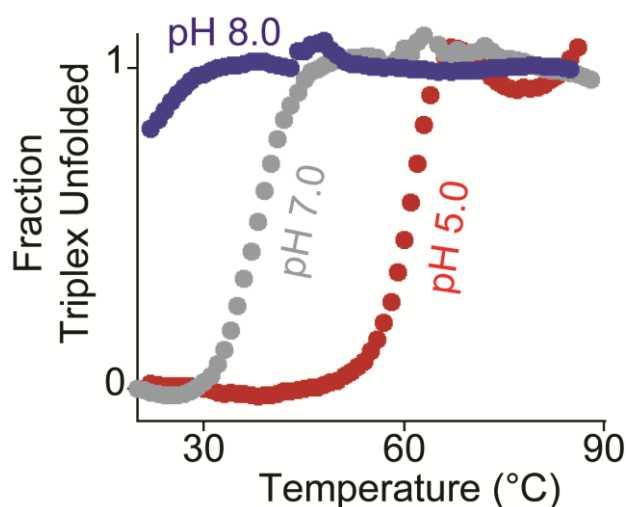
Monte Carlo bootstrap method was used to estimate the uncertainties of the free energy profiles taking into account time series correlations.

To characterize the impact of both protonation and methylation on the stability of the triplexes, the free energy profiles for the triplex to duplex transition have been computed for (i) protonated unmethylated triplex; (ii) protonated and 1-Me triplex; (iii) protonated and 2-Me triplex and (iv) deprotonated unmethylated triplex (Figure S13). For the protonated unmethylated triplex (black line), the unfolding process is strongly disfavored in the whole range of RC. In the 1-Me (green) and 2-Me protonated (blue) triplexes, we observe an initial increase of the free energy (i.e., at RC  $\sim$ 1.2 nm), corresponding to the cost for the contact break between the 3' terminal cytosine and the corresponding guanine. Then, the free energy profile reaches a plateau up to RC  $\sim$ 2.7 nm, corresponding to the unfolding process. This suggests that the unfolding could proceed spontaneously in this region. Next, above RC  $\sim$ 4.3 nm, the free energy rapidly increases with the RC for the 1-Me triplex (green), showing that the process is strongly disfavored, similarly to the protonated triplex (black). On the other hand, in the case of the 2-Me system (blue), we observe a lower increase of the free energy with the RC, suggesting that further unfolding is easier up to RC  $\sim$ 6.15 nm. For the deprotonated unmethylated triplex (red), we initially observe a drop in free energy up to RC  $\sim$ 1.6 nm, which suggests that the unfolding is initially favored, as the deprotonated terminal cytosine does not form efficient Hoogsteen interactions with the corresponding guanine. The subsequent region of the free energy profile shows similarity to the 2-Me triplex (blue), yet the free energy cost is lower. These results indicate that the stability of the triplex systems is as it follows: the most stable is the protonated unmethylated triplex (black), followed by the protonated 1-Me (green), then by the protonated 2-Me triplex (blue), while the least stable is deprotonated unmethylated triplex (red).

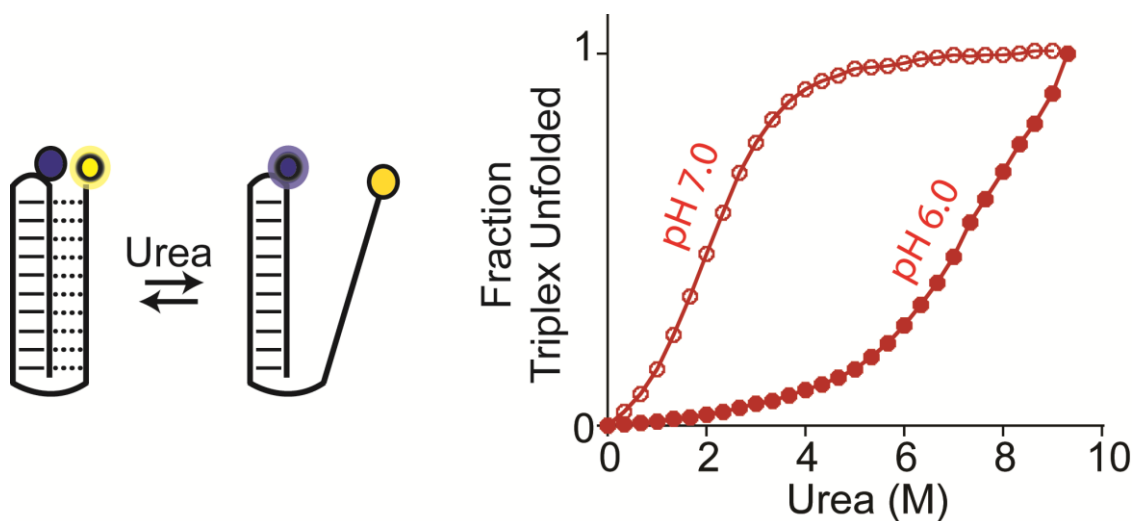
It is important to note that the computed values of the free energy profiles cannot be considered as absolute, due to the limitations of the force fields in describing single stranded DNA and the inability of the simulations to capture the protonation-deprotonation equilibria for cytosine. Nevertheless, the relative energetics depict a reasonable scenario in consistency with the experiments reported here. Accordingly, while the unfolding of the protonated triplex is highly disfavored, the cost for the unfolding of 1-Me triplex is moderately low in the region of RC  $\sim$ 1 to  $\sim$ 3.5 nm and the further unfolding is strongly disfavored. On the other hand, the free energy profiles for the 2-Me triplex and the deprotonated unmethylated triplex are comparable with each other, showing the lowest stability in the whole region of RC.



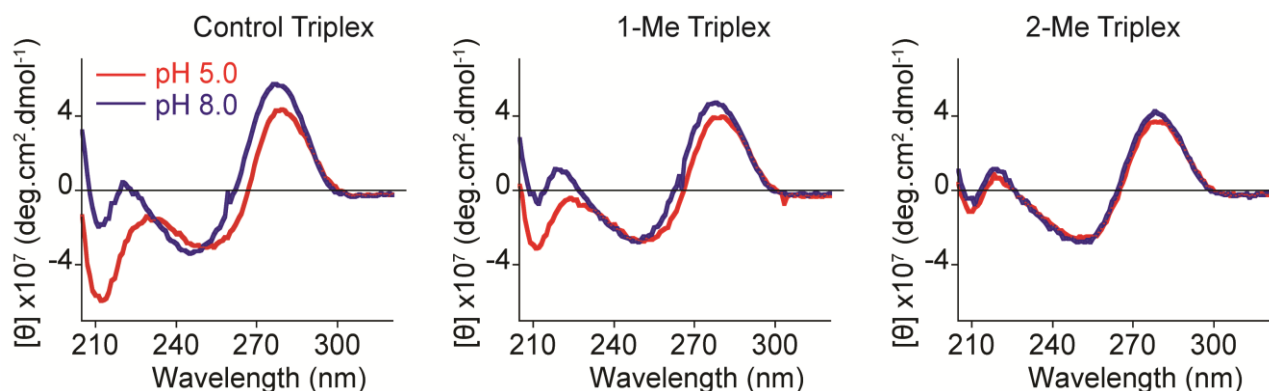
**Figure S1.** Relative FRET emission plots for triplex nanoswitches showing the triplex to duplex transition of the DNA nanoswitch as a function of pH changes in the solution. The experimental conditions were the same as described in Figure 1.



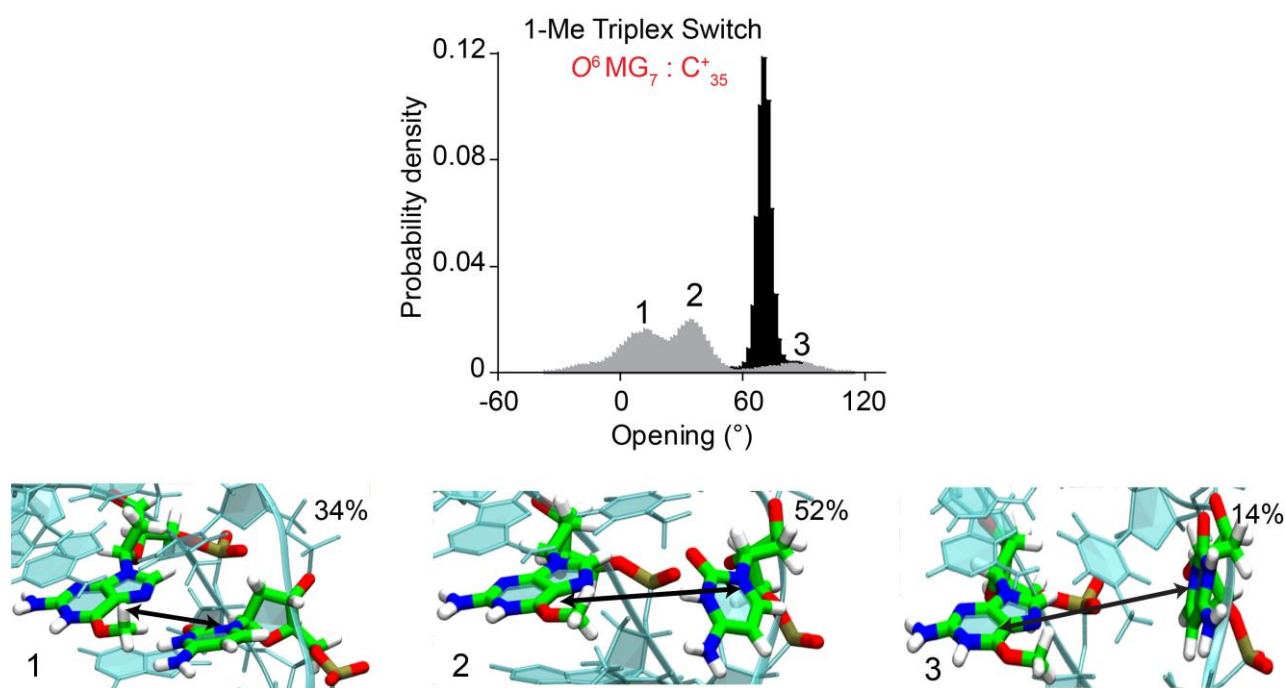
**Figure S2.** Melting curve experiments: At pH 5.0, the low initial fluorescence signal supports the formation of the triplex structure. As the temperature increases, we observe a sharp transition at about  $T_m \approx 65$  °C due to the triplex unfolding. At pH 7.0, the shift in the triplex to duplex transition is shifted to lower temperatures. At pH 8.0, most DNA probes populate the unfolded state, and no transition can be observed as the intramolecular duplex state is expected to be highly stable under these experimental conditions. The experiment was conducted using (50nM) Triplex nanoswitch in 50 mM  $\text{Na}_2\text{HPO}_4$ , 150 mM NaCl buffer at different pHs (pH=5.0, 7.0, and 8.0). The temperature was ramped at a rate of 1°C/min from 20°C to 90°C.



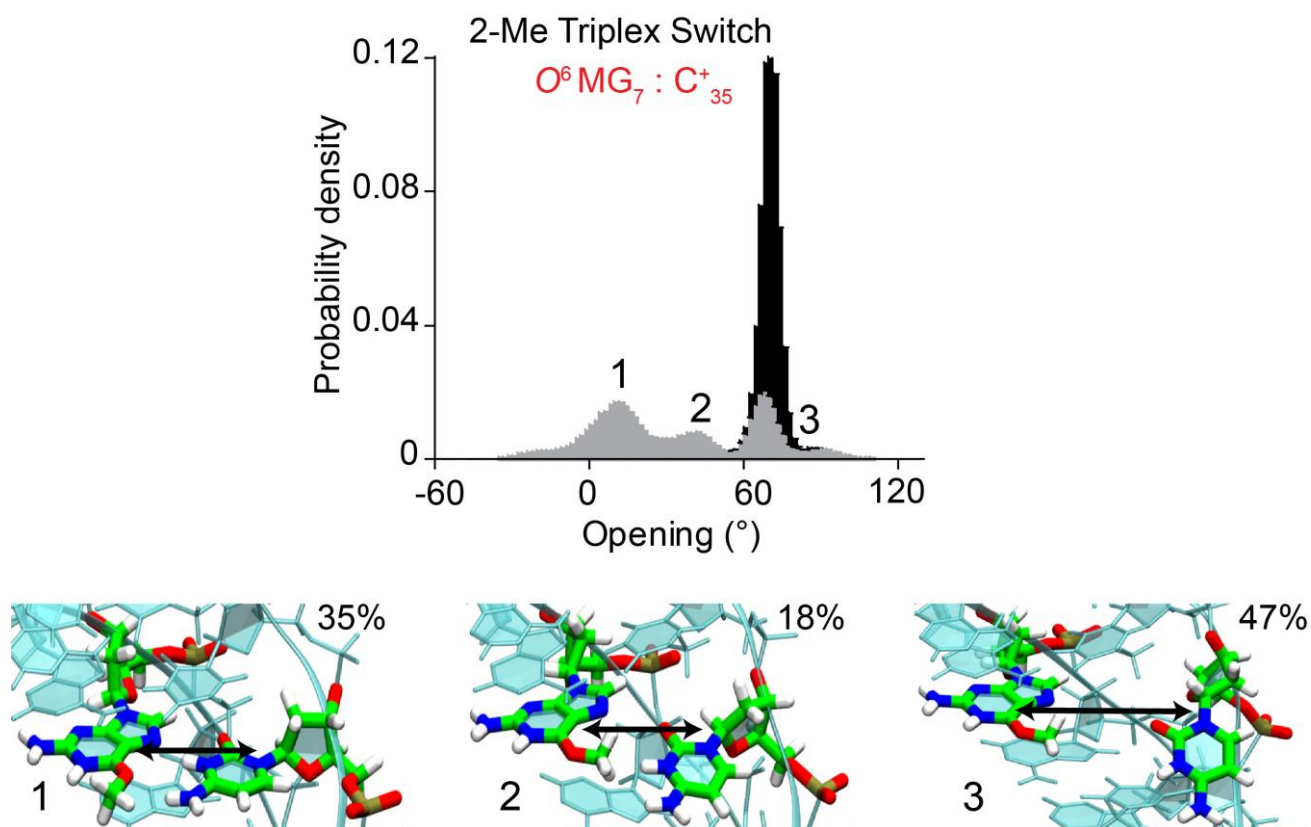
**Figure S3.** Urea denaturation experiments for the unmethylated triplex nanoswitch at 2 different pHs (6.0 and 7.0) showing a pH-dependent transition that supports the formation of pH-dependent Hoogsteen interactions. Urea titration curves were obtained at 25°C using a solution containing the triplex switch (50nM) and sequentially increasing the urea concentration by adding increasing volumes of a 10 M urea solution prepared in the working buffer solution (50mM Phosphate buffer) and containing the same concentration of the triplex switch (50nM).



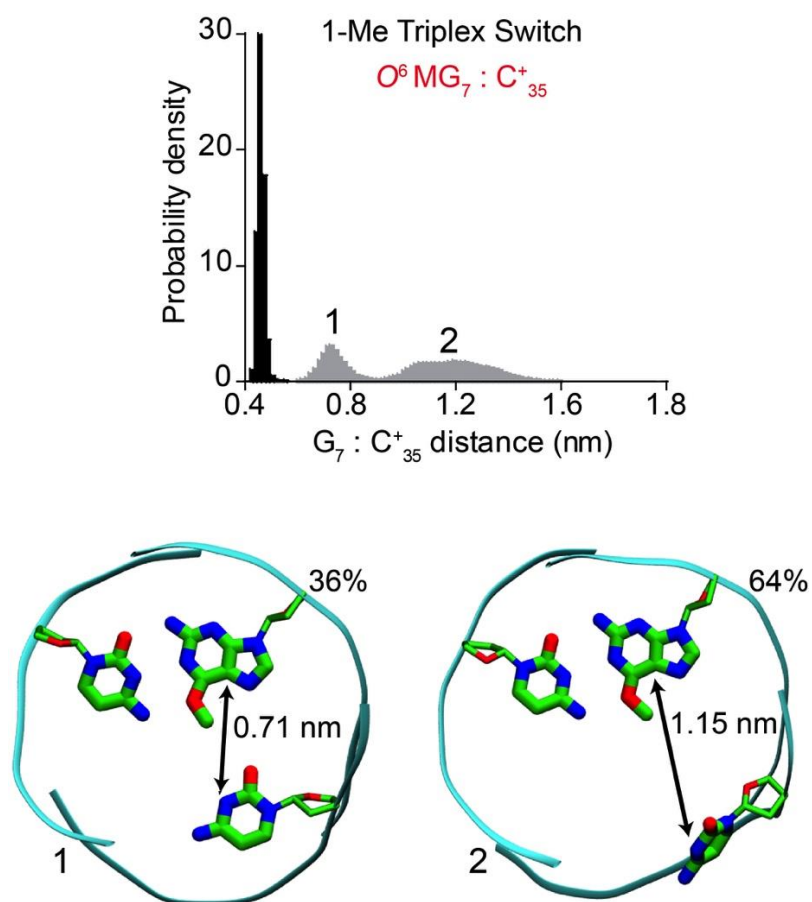
**Figure S4.** CD spectral analysis for the triplex nanoswitches at two pHs (pH 5.0 and 8.0). The unmethylated triplex at pH 8.0 shows the classic CD spectrum of a duplex DNA with two positive peaks around 280 and 220 nm and a negative one around 245 nm.<sup>[22]</sup> At pH 5.0, a sharp negative signal at 210 nm characteristic of triplex formation can also be observed. The introduction of one *O*<sup>6</sup>-MeG in the 1-Me switch strongly destabilizes the triplex formation confirmed by a significant decrease in the negative peak intensity at 210 nm indicating a partially unfolded triplex structure. 2-Me Triplex switch shows no significant difference in the CD spectra confirming the switch unfolding into a duplex structure. The experiment was performed at 25°C using (10 μM) triplex nanoswitches in 50mM Na<sub>2</sub>HPO<sub>4</sub> buffer at pHs (5.0 and 8.0).



**Figure S5.** Side-view for the  $G_7:C^+_{35}$  Hoogsteen interaction of other representative structures of 1-Me Triplex switch. The % propensity for each configuration is also indicated.

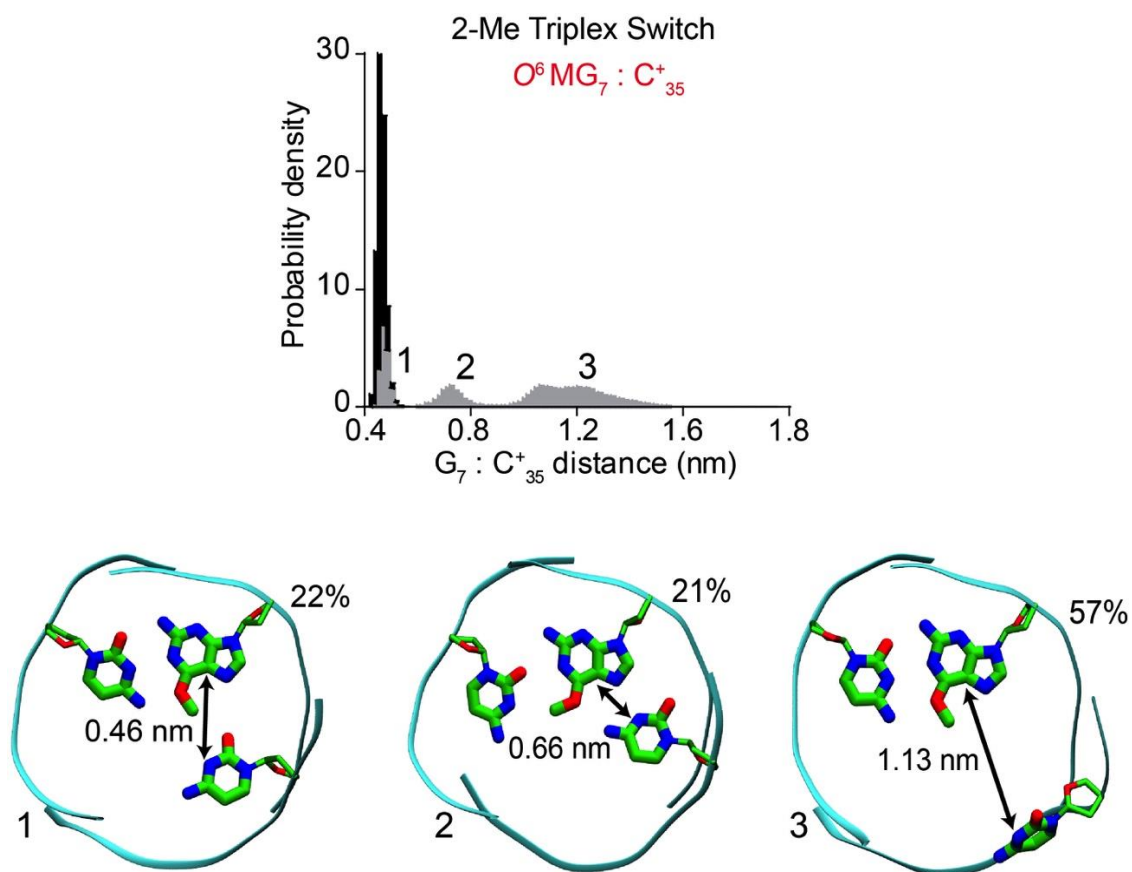


**Figure S6.** Side-view for the  $G_7:C^+_{35}$  Hoogsteen interaction of other representative structures of 2-Me Triplex switch. The % propensity for each configuration is also indicated.

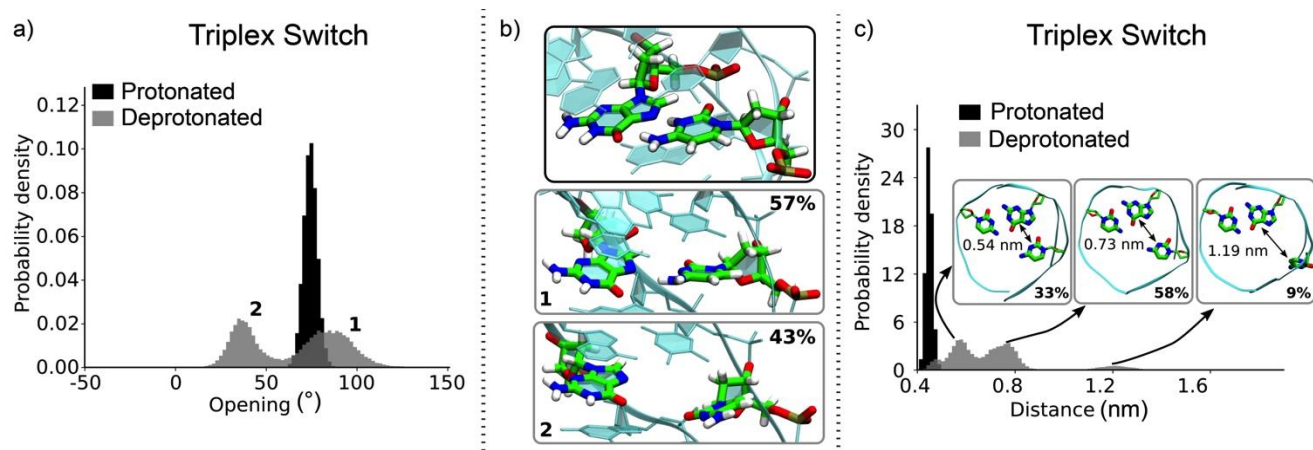


**Figure S7.** Top-view for the  $G_7:C_{35}^+$  Hoogsteen interaction of other possible structures of 1-Me Triplex switch. The % propensity for each configuration is also shown.

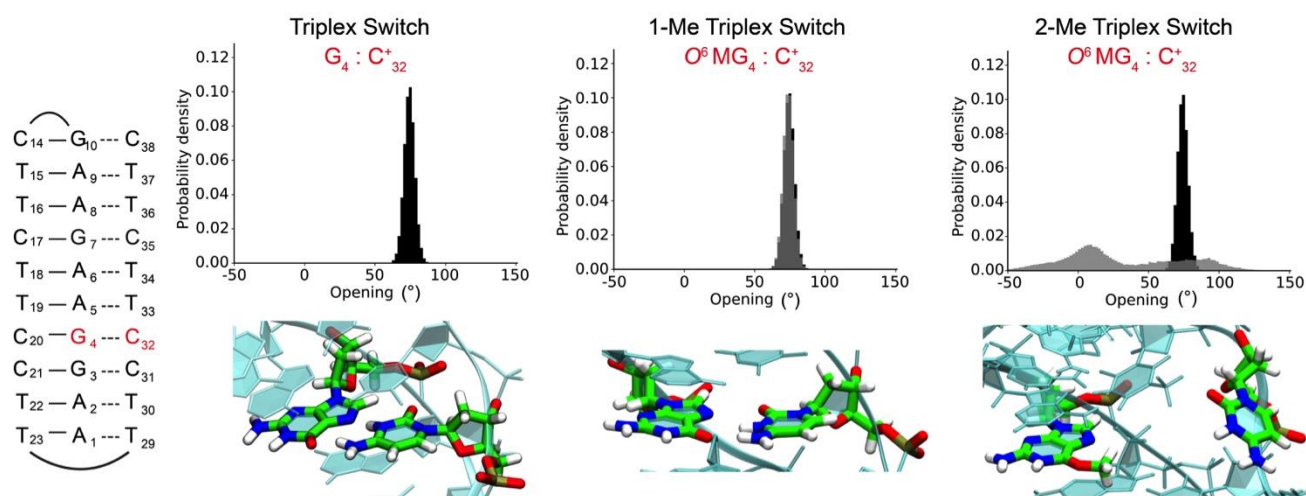




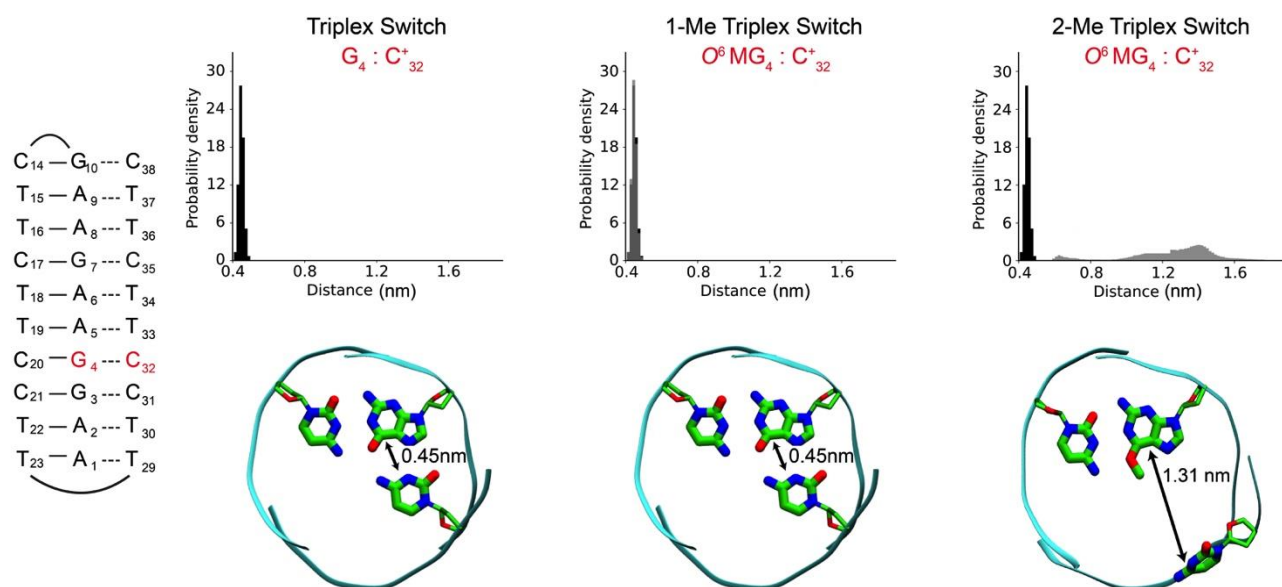
**Figure S8.** Top-view for the  $G_7:C^+_{35}$  Hoogsteen interaction of other possible structures of 2-Me Triplex switch. The % propensity for each configuration is also shown.



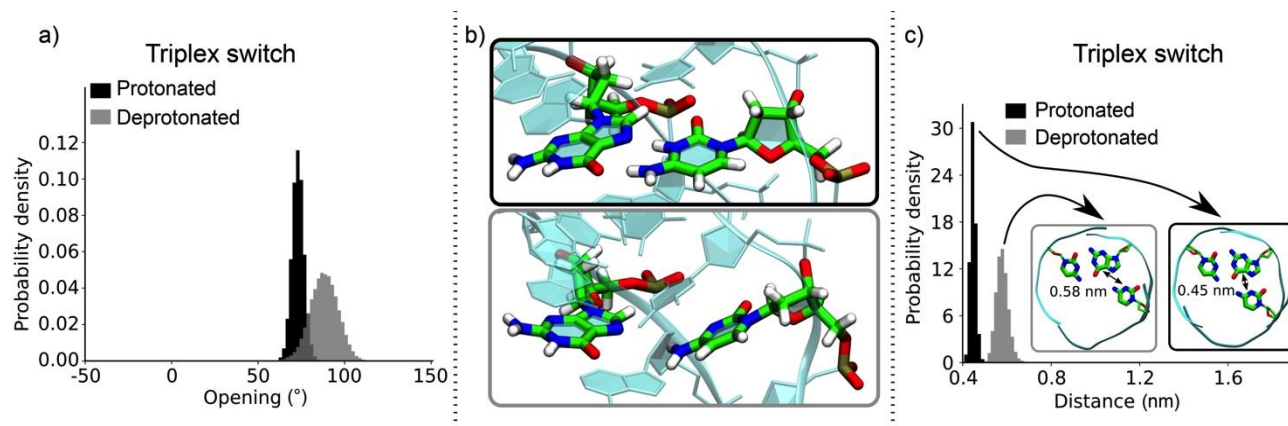
**Figure S9.** (a) Opening distribution for the G<sub>4</sub>:C<sup>+</sup><sub>32</sub> Hoogsteen interaction for the Triplex switch under protonated conditions (black) and deprotonated conditions (grey). (b) snapshot of the side view with highlighted nucleotides is also shown for both protonated (black) and deprotonated (grey) states. (c) Distance parameter for the G<sub>4</sub>:C<sup>+</sup><sub>32</sub> Hoogsteen interaction for the Triplex switch with a snapshot of the top view of the triplet most probable configuration is also shown. The distance distribution of the non-methylated protonated Triplex switch is shown (black) and deprotonated (grey).



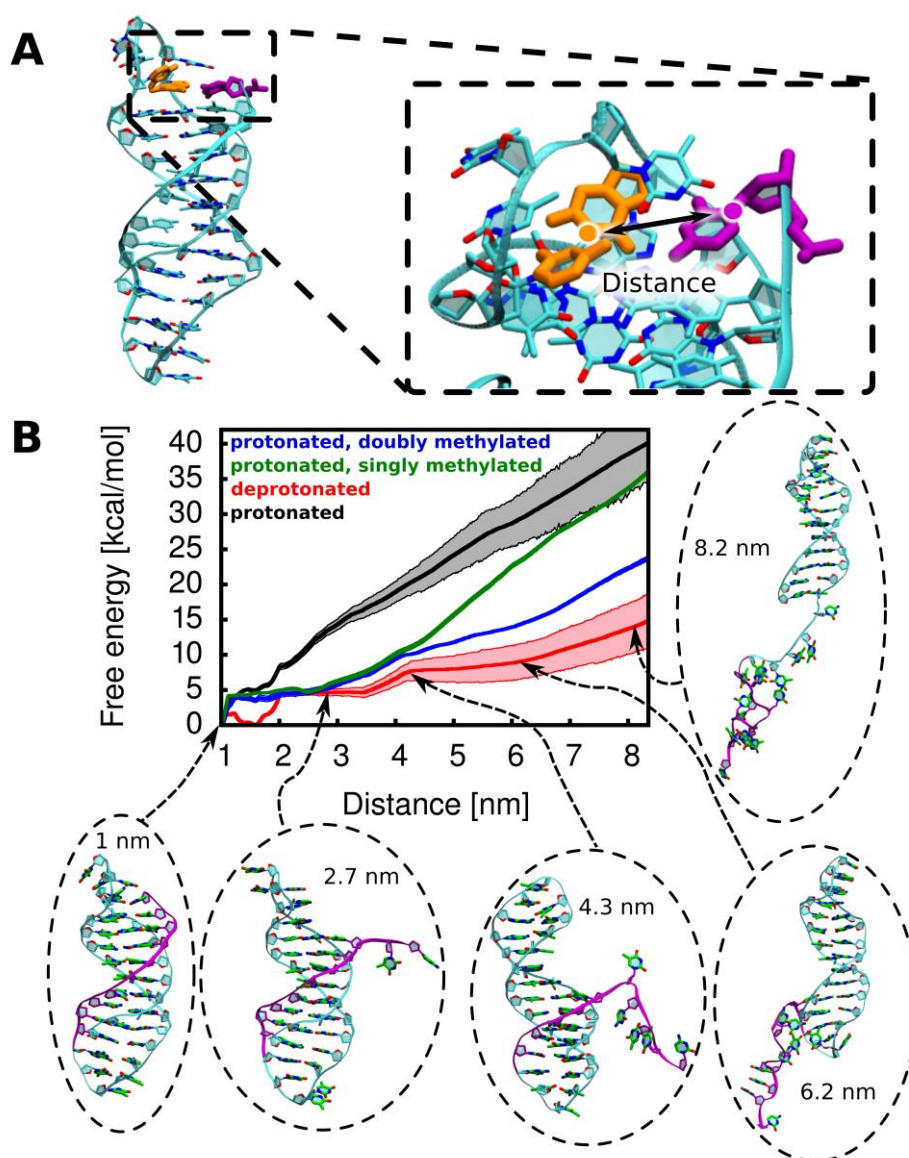
**Figure S10.** Opening distribution for the  $G_4:C_{32}^+$  Hoogsteen interaction for the Triplex switch, 1-Me Triplex switch, and 2-Me Triplex switch in non-methylated (black) and methylated (grey) states. The opening distribution of the non-methylated Triplex switch is shown (black) in the graphs of methylated switches as a comparison. Below each graph, a snapshot of the side view with highlighted nucleotides is also shown.



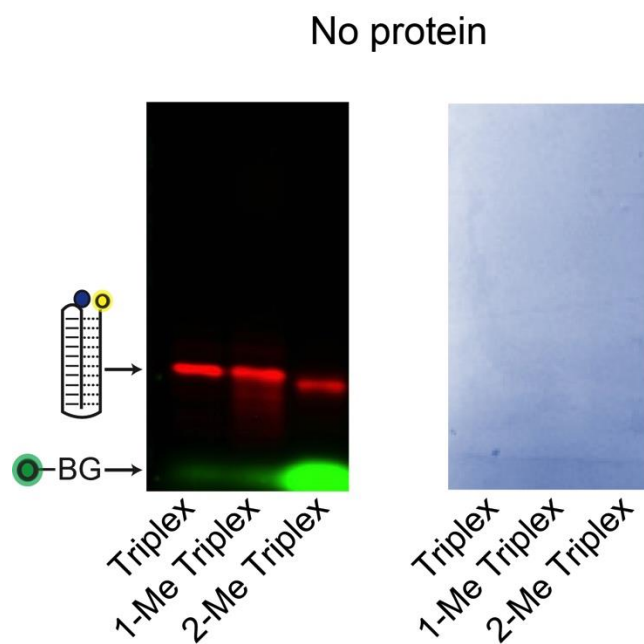
**Figure S11.** Distance distribution for the  $G_4:C_{32}^+$  Hoogsteen interaction for the Triplex switch, 1-Me Triplex switch, and 2-Me Triplex switch in non-methylated (black) and methylated (grey) states. Below each graph, a snapshot of the top view of the triplet most probable configuration is also shown. The distance distribution of the non-methylated Triplex switch is shown (black) in the graphs of methylated switches as a comparison.



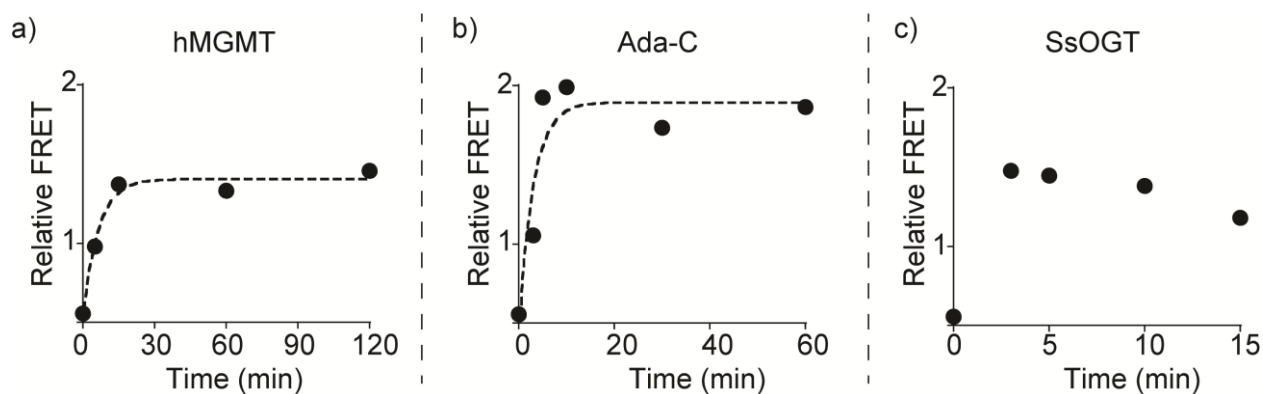
**Figure S12.** (a) Opening distribution for the  $G_7:C^+_{35}$  Hoogsteen interaction for the Triplex switch under protonated conditions (black) and deprotonated conditions (grey). (b) snapshot of the side view with highlighted nucleotides is also shown for both protonated (black) and deprotonated (grey) states. (c) Distance parameter for the  $G_7:C^+_{35}$  Hoogsteen interaction for the Triplex switch with a snapshot of the top view of the triplet most probable configuration is also shown. The distance distribution of the non-methylated protonated Triplex switch is shown (black) and deprotonated (grey).



**Figure S13. (A)** Free energy simulations of the triplex nanoswitches. The reaction coordinate (RC) has been defined as the distance between two center of mass (COM): (i) the heavy atoms of 3'-terminal cytidine phosphate and (ii) the nucleobases that co-form the triplex plane with the terminal cytosine. **(B)** Free energy profiles for the triplex to duplex transition for (i) protonated unmethylated triplex (black); (ii) protonated 1-Me triplex (green); (iii) protonated 2-Me triplex (blue); (iv) deprotonated control unmethylated triplex (red). Representative structures of the protonated triplex at selected RC values are shown in circles.

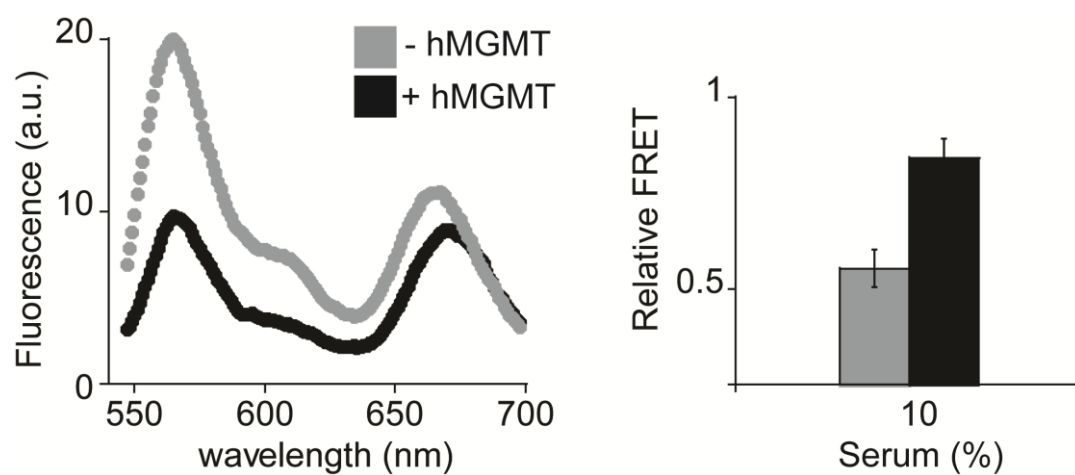


**Figure S14.** SDS-PAGE gel-imaging (left) showing visible bands for both O6-BG fluorescent derivative (SNAP-Vista Green®) and the three DNA nanoswitches tested. It demonstrates that no interaction occurs between the nanoswitches and the fluorescent-labelled substrate in the absence of any protein. Fluorescence-labelled triplex nanoswitches appear as red bands. 2-Me Triplex shows a slightly, but significant, higher mobility. Presumably, it adopts a partial folding in denaturing conditions.



**Figure S15.** Time-course analysis of the enzymatic reaction with the 2-Me Triplex nanoswitch using three different methyltransferase enzymes (a) hMGMT, (b) *E. coli* Ada-C, and (c) SsOGT. The experiment was performed as described in Figure 4.





**Figure S16.** Analysis of hMGMT repair activity in 10% serum. Relative FRET signals obtained with (0.5  $\mu\text{M}$ ) 2-Me triplex nanoswitches in the absence and presence of 5  $\mu\text{M}$  (0.1  $\mu\text{g}/\mu\text{l}$ ) hMGMT. The experiment was conducted under the same conditions described in Figure 4.

**Table S1.** DNA base pair parameters (Shear, Stretch, Stagger, Buckle, Propeller, Opening) for the triplex switch, computed along ~1.5  $\mu$ s of molecular dynamics simulations. DNA parameters were computed between the base pairs reported in the first column. The base pairs forming a duplex are reported in the top portion of the table, while the base pairs forming a triplex are reported in the bottom.

Triplex switch (strands #1 and #2, forming a duplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 23	-0.07±0.29	-0.06±0.14	-0.57±0.50	-6.39±10.77	-10.44±8.41	1.49±5.45
2 - 22	0.03±0.27	-0.02±0.12	0.13±0.42	-1.58±9.37	-10.22±7.03	-1.46±4.54
3 - 21	-0.23±0.30	-0.12±0.10	-0.02±0.37	-3.21±8.79	-10.81±6.90	-0.78±2.99
4 - 20	-0.16±0.31	-0.08±0.11	-0.26±0.38	-6.68±9.44	-12.19±8.14	-0.09±3.46
5 - 19	0.06±0.27	0.02±0.12	-0.13±0.42	-2.30±10.15	-12.37±7.29	0.06±5.00
6 - 18	0.03±0.28	-0.03±0.11	0.15±0.42	0.33±8.99	-12.17±6.49	0.55±4.83
7 - 17	-0.21±0.30	-0.11±0.10	-0.13±0.37	0.57±8.12	-10.01±7.39	-1.18±3.06
8 - 16	0.02±0.27	0.02±0.12	0.02±0.42	-3.37±8.88	-14.81±7.04	-1.90±4.65
9 - 15	0.11±0.27	-0.01±0.12	-0.06±0.42	-4.70±8.65	-12.20±8.16	-2.46±4.66
10 - 14	-0.09±0.37	-0.07±0.17	-0.15±0.60	7.98±11.72	12.95±16.48	-1.58±5.72
Triplex switch (strands #1 and #3, forming a triplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 29	0.93±0.42	-3.68±0.22	0.13±0.66	3.39±9.63	3.25±9.42	75.54±5.48
2 - 30	0.69±0.35	-3.56±0.19	0.47±0.62	2.62±8.14	4.90±8.38	71.70±4.79
3 - 31	0.37±0.33	-3.36±0.26	-0.35±0.43	10.76±6.54	-3.63±8.60	74.77±3.57
4 - 32	0.68±0.38	-3.56±0.26	-0.02±0.43	5.03±6.97	7.49±8.69	74.38±3.93
5 - 33	0.86±0.34	-3.66±0.20	0.38±0.54	1.39±7.78	12.89±7.96	74.34±4.56
6 - 34	0.67±0.33	-3.58±0.20	0.25±0.66	4.99±8.80	8.64±8.38	72.09±5.17
7 - 35	0.57±0.32	-3.52±0.23	-0.14±0.48	6.12±7.28	5.52±8.20	73.11±3.43
8 - 36	0.90±0.33	-3.68±0.19	0.68±0.56	-0.07±7.85	9.65±6.61	72.64±4.74
9 - 37	0.74±0.29	-3.60±0.19	-0.56±0.52	14.89±8.73	1.24±7.12	76.49±3.77
10 - 38	0.45±0.34	-3.43±0.29	0.17±0.49	12.24±8.41	16.95±8.66	69.06±4.88

**Table S2.** DNA base pair parameters (Shear, Stretch, Stagger, Buckle, Propeller, Opening) for the 1Me-triplex switch, computed along  $\sim 1.5 \mu\text{s}$  of molecular dynamics simulations. DNA parameters were computed between the base pairs reported in the first column. The base pairs forming a duplex are reported in the top portion of the table, while the base pairs forming a triplex are reported in the bottom.

1Me-triplex switch (strands #1 and #2, forming a duplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 23	-0.04±0.29	-0.06±0.14	-0.59±0.49	-5.48±10.83	-8.71±8.33	1.80±5.41
2 - 22	0.03±0.28	-0.01±0.12	0.11±0.42	-0.04±9.27	-8.63±7.70	-1.36±4.65
3 - 21	-0.25±0.31	-0.12±0.11	-0.04±0.38	-1.90±8.88	-9.59±7.21	-0.59±3.09
4 - 20	-0.20±0.30	-0.08±0.11	-0.28±0.38	-6.31±9.47	-11.53±8.34	-0.20±3.38
5 - 19	0.07±0.28	0.03±0.13	-0.22±0.45	-8.25±9.72	-14.69±7.76	2.62±5.71
6 - 18	0.10±0.28	-0.01±0.12	0.03±0.42	-5.22±10.28	-5.22±7.65	-0.69±4.89
7 - 17	-0.41±0.43	0.00±0.15	0.07±0.47	3.36±9.57	-8.36±8.00	1.95±3.93
8 - 16	0.03±0.27	0.00±0.13	-0.19±0.48	-6.04±9.93	-15.14±7.61	-0.95±5.24
9 - 15	0.12±0.29	0.00±0.12	-0.03±0.43	-4.75±9.26	-11.78±8.49	-1.27±5.10
10 - 14	0.17±1.47	-0.22±0.78	0.03±0.63	4.81±11.90	6.91±13.34	-1.83±6.32
1Me-triplex switch (strands #1 and #3, forming a triplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 29	0.83±0.39	-3.65±0.22	0.26±0.64	1.19±9.78	-0.05±10.45	73.82±5.90
2 - 30	0.75±0.34	-3.58±0.19	0.60±0.63	0.61±8.57	3.92±8.97	72.95±5.02
3 - 31	0.39±0.28	-3.36±0.24	-0.25±0.45	9.16±6.97	-4.39±9.36	74.42±3.41
4 - 32	0.64±0.37	-3.54±0.26	0.05±0.46	2.83±7.70	5.85±9.30	73.93±3.84
5 - 33	0.97±0.38	-3.72±0.23	0.60±0.63	-3.63±9.89	11.60±7.80	75.78±5.06
6 - 34	-0.58±2.30	-5.26±2.59	0.25±1.17	-6.86±13.03	18.17±11.11	89.15±22.88
7 - 35	-7.54±2.45	-2.63±4.65	-3.11±4.09	0.33±34.42	-18.14±26.87	28.16±29.67
8 - 36	0.81±0.33	-3.63±0.20	0.28±0.68	2.64±9.33	6.81±9.27	73.60±5.28
9 - 37	0.78±0.32	-3.62±0.20	-0.19±0.76	9.65±11.89	4.75±9.70	75.71±4.31
10 - 38	0.50±0.45	-3.44±0.35	0.31±0.78	7.42±14.74	17.09±9.51	70.44±6.75

**Table S3.** DNA base pair parameters (Shear, Stretch, Stagger, Buckle, Propeller, Opening) for the 2Me-triplex switch, computed along ~1.5  $\mu$ s of molecular dynamics simulations. DNA parameters were computed between the base pairs reported in the first column. The base pairs forming a duplex are reported in the top portion of the table, while the base pairs forming a triplex are reported in the bottom.

2Me-triplex switch (strands #1 and #2, forming a duplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 23	0.06+-0.31	0.01+-0.15	-0.27+-0.56	0.18+-14.29	-11.42+-7.83	1.66+-5.77
2 - 22	-0.02+-0.27	-0.02+-0.12	-0.01+-0.42	3.95+-10.43	-11.44+-7.72	1.60+-5.21
3 - 21	-0.37+-0.29	-0.13+-0.11	-0.07+-0.37	0.90+-8.61	-4.82+-7.95	-0.54+-3.01
4 - 20	-0.52+-0.42	-0.01+-0.15	-0.20+-0.46	-5.76+-11.63	-12.66+-8.45	3.41+-4.12
5 - 19	0.08+-0.28	0.02+-0.13	-0.30+-0.46	-10.57+-9.89	-14.37+-7.56	2.75+-5.73
6 - 18	0.12+-0.28	-0.00+-0.12	0.02+-0.42	-7.62+-9.51	-7.11+-7.64	-0.15+-5.06
7 - 17	-0.94+-1.02	0.20+-0.37	0.08+-0.51	0.71+-9.65	-11.80+-9.32	10.41+-14.33
8 - 16	0.07+-0.27	0.02+-0.13	-0.16+-0.47	-7.11+-9.72	-14.77+-7.45	0.41+-6.08
9 - 15	0.11+-0.28	-0.00+-0.12	-0.07+-0.42	-4.62+-8.64	-10.41+-9.02	-1.31+-5.09
10 - 14	-0.12+-0.38	-0.08+-0.17	-0.18+-0.63	6.92+-11.37	11.84+-18.22	-1.83+-5.29
2Me-triplex switch (strands #1 and #3, forming a triplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 29	-0.45+-2.81	-3.67+-1.73	1.68+-3.35	-10.78+-32.71	-13.91+-22.51	77.00+-47.56
2 - 30	0.77+-0.35	-3.61+-0.21	0.52+-0.63	-2.66+-9.57	0.14+-11.09	73.54+-5.20
3 - 31	0.21+-1.87	-3.82+-1.98	0.63+-1.43	-12.53+-14.35	9.36+-11.97	73.23+-7.24
4 - 32	-6.11+-5.15	-9.11+-6.10	-0.50+-6.88	-12.26+-50.74	-23.81+-35.54	30.39+-43.65
5 - 33	0.85+-0.51	-3.69+-0.41	0.38+-0.91	-2.59+-13.02	2.96+-11.16	75.79+-6.86
6 - 34	0.78+-0.35	-3.63+-0.22	0.59+-0.83	-4.62+-12.55	11.90+-10.57	73.71+-5.30
7 - 35	-5.69+-4.36	-2.76+-3.86	-2.61+-4.08	0.31+-28.30	-19.31+-26.71	36.03+-33.57
8 - 36	0.85+-0.33	-3.65+-0.20	0.48+-0.60	0.04+-8.77	6.73+-8.85	73.92+-5.15
9 - 37	0.79+-0.32	-3.63+-0.21	0.02+-0.84	6.50+-12.86	5.16+-9.93	75.52+-4.44
10 - 38	0.48+-0.36	-3.47+-0.29	0.57+-0.70	2.58+-13.68	16.65+-9.06	69.20+-5.00

**Table S4.** DNA base pair parameters (Shear, Stretch, Stagger, Buckle, Propeller, Opening) for the deprotonated Triplex switch, computed along  $\sim 1.5 \mu\text{s}$  of molecular dynamics simulations. DNA parameters were computed between the base pairs reported in the first column. The base pairs forming a duplex are reported in the top portion of the table, while the base pairs forming a triplex are reported in the bottom.

Triplex switch, deprotonated (strands #1 and #2, forming a duplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 23	0.04+-0.29	-0.02+-0.16	-0.61+-0.49	-8.02+-12.27	-8.21+-8.38	5.66+-6.07
2 - 22	0.06+-0.29	0.01+-0.13	-0.05+-0.43	2.82+-10.06	-6.81+-9.31	-1.31+-5.40
3 - 21	-0.22+-0.31	-0.10+-0.11	-0.10+-0.40	-0.47+-9.51	-6.21+-7.94	-0.67+-3.12
4 - 20	-0.10+-0.32	-0.06+-0.12	-0.34+-0.41	-3.56+-10.68	-8.08+-8.52	0.26+-3.73
5 - 19	0.12+-0.27	0.02+-0.12	-0.14+-0.45	-2.25+-9.66	-14.52+-7.28	0.91+-5.44
6 - 18	0.01+-0.28	-0.02+-0.11	-0.01+-0.42	0.18+-9.10	-9.58+-7.12	0.38+-5.00
7 - 17	-0.14+-0.30	-0.06+-0.11	-0.18+-0.39	-1.54+-8.92	-5.06+-7.80	-0.36+-3.10
8 - 16	0.08+-0.27	0.02+-0.12	-0.17+-0.45	-3.91+-10.27	-12.66+-7.97	0.91+-5.51
9 - 15	0.06+-0.28	0.01+-0.13	-0.08+-0.42	-0.23+-10.37	-6.18+-8.64	-1.40+-5.70
10 - 14	-0.16+-0.36	-0.05+-0.17	0.11+-0.46	11.90+-11.57	0.97+-11.11	-0.99+-4.08
Triplex switch, deprotonated (strands #1 and #3, forming a triplex)						
base pair	Shear [Å]	Stretch [Å]	Stagger [Å]	Buckle [°]	Propeller [°]	Opening [°]
1 - 29	0.77+-0.37	-3.67+-0.22	-0.03+-0.68	4.57+-11.38	0.26+-10.79	72.19+-6.47
2 - 30	0.77+-0.50	-3.76+-0.33	0.52+-0.76	2.47+-11.00	12.31+-11.98	69.65+-7.54
3 - 31	0.24+-3.18	-4.64+-1.20	0.64+-1.53	-4.10+-15.08	8.05+-19.90	66.69+-22.97
4 - 32	-1.85+-3.80	-5.12+-2.77	1.36+-1.52	-2.14+-13.31	13.08+-26.03	64.51+-28.31
5 - 33	0.87+-0.35	-3.67+-0.21	0.31+-0.78	3.25+-10.58	6.98+-9.49	73.91+-5.43
6 - 34	0.84+-0.33	-3.71+-0.20	-0.02+-0.68	4.22+-10.10	8.58+-9.54	73.55+-5.82
7 - 35	2.72+-0.43	-5.25+-0.30	0.83+-0.67	-10.40+-8.93	3.05+-9.19	88.34+-8.10
8 - 36	0.76+-0.60	-3.64+-0.25	1.20+-0.77	-10.34+-9.54	9.67+-9.36	72.87+-6.35
9 - 37	0.88+-0.44	-3.64+-0.30	1.53+-1.24	-14.84+-16.65	17.93+-13.72	76.30+-7.63
10 - 38	-5.54+-4.68	-10.13+-7.81	3.49+-4.77	-2.11+-27.26	-11.34+-36.81	81.29+-93.28

**Table S5.** Topology and computed partial charges for protonated cytosine.

atoms	Atom types	Atomic charges	bonds	
P	P	1.165900	P	O1P
O1P	O2	-0.776100	P	O2P
O2P	O2	-0.776100	P	O5'
O5'	OS	-0.495400	O5'	C5'
C5'	CI	-0.006900	C5'	H5'1
H5'1	H1	0.075400	C5'	H5'2
H5'2	H1	0.075400	C5'	C4'
C4'	CT	0.162900	C4'	H4'
H4'	H1	0.117600	C4'	O4'
O4'	OS	-0.369100	C4'	C3'
C1'	CT	-0.011600	O4'	C1'
H1'	H2	0.196300	C1'	H1'
N1	N*	0.974980	C1'	N1
C6	CM	-0.761836	C1'	C2'
H6	H4	0.588626	N1	C6
C5	CM	-0.024512	N1	C2
H5	HA	0.181781	C6	H6
C4	CA	0.289000	C6	C5
N4	N2	-0.958833	C5	H5
H41	H	0.467388	C5	C4
H42	H	0.467388	C4	N4
N3	N*	0.070370	C4	N3
C2	C	-0.286788	N4	H41
O2	O	-0.373286	N4	H42
C3'	CT	0.071300	N3	C2
H3'	H1	0.098500	C2	O2
C2'	CT	-0.085400	C3'	H3'
H2'1	HC	0.071800	C3'	C2'
H2'2	HC	0.071800	C3'	O3'
HP	H	0.302622	C2'	H2'1
O3'	OS	-0.523200	C2'	H2'2
			N3	HP
			-O3'	P
impropers				
C2	C6	N1	C1'	
N1	N3	C2	O2	
C4	H41	N4	H42	
N1	C5	C6	H6	
C6	C4	C5	H5	
C5	N4	C4	N3	

**Table S6.** Topology and computed partial charges for 3'-terminal protonated cytosine.

atoms	Atom types	Atomic charges		bonds	
P	P	1.165900		P	O1P
O1P	O2	-0.776100		P	O2P
O2P	O2	-0.776100		P	O5'
O5'	OS	-0.495400		O5'	C5'
C5'	CI	-0.006900		C5'	H5'1
H5'1	H1	0.075400		C5'	H5'2
H5'2	H1	0.075400		C5'	C4'
C4'	CT	0.162900		C4'	H4'
H4'	H1	0.117600		C4'	O4'
O4'	OS	-0.369100		C4'	C3'
C1'	CT	-0.011600		O4'	C1'
H1'	H2	0.196300		C1'	H1'
N1	N*	0.974980		C1'	N1
C6	CM	-0.761836		C1'	C2'
H6	H4	0.588626		N1	C6
C5	CM	-0.024512		N1	C2
H5	HA	0.181781		C6	H6
C4	CA	0.289000		C6	C5
N4	N2	-0.958833		C5	H5
H41	H	0.467388		C5	C4
H42	H	0.467388		C4	N4
N3	N*	0.070370		C4	N3
C2	C	-0.286788		N4	H41
O2	O	-0.373286		N4	H42
C3'	CT	0.071300		N3	C2
H3'	H1	0.098500		C2	O2
C2'	CT	-0.085400		C3'	H3'
H2'1	HC	0.071800		C3'	C2'
H2'2	HC	0.071800		C3'	O3'
O3'	OH	-0.654900		C2'	H2'1
HP	H	0.302622		C2'	H2'2
H3T	HO	0.43960		O3'	H3T
				N3	HP
				-O3'	P
impropers					
C2	C6	N1	C1'		
N1	N3	C2	O2		
C4	H41	N4	H42		
N1	C5	C6	H6		
C6	C4	C5	H5		
C5	N4	C4	N3		

**Table S7.** Topology and computed partial charges for methylated guanine.

atoms	Atom types	Atomic charges	bonds	
P	P	1.165900	P	O1P
O1P	O2	-0.776100	P	O2P
O2P	O2	-0.776100	P	O5'
O5'	OS	-0.495400	O5'	C5'
C5'	CI	-0.006900	C5'	H5'1
H5'1	H1	0.075400	C5'	H5'2
H5'2	H1	0.075400	C5'	C4'
C4'	CT	0.162900	C4'	H4'
H4'	H1	0.117600	C4'	O4'
O4'	OS	-0.369100	C4'	C3'
C1'	CT	0.035800	O4'	C1'
H1'	H2	0.174600	C1'	H1'
N9	N*	0.861452	C1'	N9
C8	CK	-0.224093	C1'	C2'
H8	H5	0.370032	N9	C8
N7	NB	-0.636871	N9	C4
C5	CB	0.701865	C8	H8
C6	C	0.160093	C8	N7
O6	OS	-0.256184	N7	C5
N1	NA	-0.120981	C5	C6
H1	H	0.229271	C5	C4
C2	CA	0.360794	C6	O6
N2	N2	-0.692175	C6	N1
H21	H	0.369539	N1	H1
H22	H	0.369539	N1	C2
N3	NC	-0.272559	C2	N2
C4	CB	-0.583189	C2	N3
C3'	CT	0.071300	N2	H21
H3'	H1	0.098500	N2	H22
C2'	CT	-0.085400	N3	C4
H2'1	HC	0.071800	C3'	H3'
H2'2	HC	0.071800	C3'	C2'
Cm	CT	0.004817	C3'	O3'
HM1	H1	0.089950	C2'	H2'1
HM2	H1	0.089950	C2'	H2'2
HM3	H1	0.089950	O6	Cm
O3'	OS	-0.523200	Cm	HM1
			Cm	HM2
			Cm	HM3
			-O3'	P
impropers				
C4	C8	N9	C1'	
C5	N1	C6	O6	
C6	C2	N1	H1	
C2	H21	N2	H22	
N9	N7	C8	H8	
N2	N1	C2	N3	



**Table S8.** Bonded parameters for the modified bases<sup>[4,5]</sup>

Bonds				Function type	Eq. value [nm]	Force constant [kJ nm <sup>-2</sup> ]
CA	N*			1	0.13390	404174.4
Angles				Function type	Eq. value [°]	Force constant [kJ rad <sup>-2</sup> ]
CB	C	OS		1	128.800	669.440
N*	C	N*		1	118.600	585.760
NA	C	OS		1	120.600	669.440
CM	CA	N*		1	121.500	585.760
N2	CA	N*		1	119.300	585.760
CA	N*	H		1	119.200	418.400
C	N*	CA		1	120.500	585.760
Improper dihedrals				Function type	Eq. value [°]	Force constant [kJ rad <sup>-2</sup> ]
CM	N2	CA	N*	4	180.00	4.60240
CB	NA	C	OS	4	180.00	43.93200
Dihedrals				Function type	Eq. value [°]	Force constant [kJ rad <sup>-2</sup> ]
X	CA	N*	X	9	180.0	6.90360

## References

- [1] G. Perugino, A. Vettone, G. Illiano, A. Valenti, M. C. Ferrara, M. Rossi, M. Ciaramella, *J. Biol. Chem.* **2012**, *287*, 4222–4231.
- [2] G. Perugino, R. Miggiano, M. Serpe, A. Vettone, A. Valenti, S. Lahiri, F. Rossi, M. Rossi, M. Rizzi, M. Ciaramella, *Nucleic Acids Res.* **2015**, *43*, 8801–8816.
- [3] X. J. Lu, W. K. Olson, *Nat. Protoc.* **2008**, *3*, 1213–1227.
- [4] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Loughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817–3829.
- [5] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins Struct. Funct. Bioinforma.* **2010**, *78*, 1950–1958.
- [6] C. Ma, C. Lin, Y. Wang, X. Chen, *TrAC - Trends Anal. Chem.* **2016**, *77*, 226–241.
- [7] K. W. East, J. C. Newton, U. N. Morzan, Y. B. Narkhede, A. Acharya, E. Skeens, G. Jogl, V. S. Batista, G. Palermo, G. P. Lisi, *J. Am. Chem. Soc.* **2020**, *142*, 1348–1358.
- [8] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [9] Frisch, M. J. E. A. et al. Gaussian 09, revision D.01 (Gaussian, Inc., **2014**).
- [10] D. A. Case Ross C Walker, T. Darden Junmei Wang, *Amber 2016 Reference Manual Principal Contributors to the Current Codes*, **n.d.**
- [11] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, *SoftwareX* **2015**, *1–2*, 19–25.
- [12] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- [13] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
- [14] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [15] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- [16] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [17] R. Kumar, H. Grubmüller, *Bioinformatics* **2015**, *31*, 2583–2585.
- [18] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.
- [19] J. D. Hunter, *Comput. Sci. Eng.* **2007**, *9*, 99–104.
- [20] Y. Sugita, A. Kitao, Y. Okamoto, *J. Chem. Phys.* **2000**, *113*, 6042–6051.
- [21] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1339–1350.
- [22] D. M. Gray, S. H. Hung, K. H. Johnson, *Methods Enzymol.* **1995**, *246*, 19–34.

## Author Contributions

A. P., G. P., and F. R. conceived the idea and designed the experiments; N. F. performed the fluorescence experiments; R. Ma. purified SsOGT, hMGMT and *E. coli* AdaC. and performed the all fluorescent competition assays on SDS-PAGE, R. Me. Purified the SsOGT-H<sup>5</sup> mutant and tested the BGN3 and BGSN inactivators; Ł. N. and G. P. performed the Molecular Simulation experiments; A. P., G. P. and F.R. supervised the research; A.P., G. P. and, F. R. wrote the manuscript.