Supporting Information for:

Electronic Tuning of Fluorescent 8-Aryl-Guanine Probes for Monitoring DNA Duplex-Quadruplex Exchange

Michael Sproviero, Kaila L. Fadock, Aaron A. Witham, Richard A. Manderville,* Purshotam Sharma and Stacey D. Wetmore*

Table of Contents:

Experimental Section	S2
Figure S1. DFT calculations of ^{Fur} dG	.S15
Figure S2. DFT calculations of ^{CNPh} dG	.S16
CD spectra of modified TBA duplex structures	S17
Figures S4-S12. NMR spectra of ^{CNPh} dG and 1-3	.S18
Table S1. ESI-MS Analysis of 8-Aryl-G Modified TBA Oligonucleotides.	S22
Figures S13-S15. ESI-MS spectrum of ^{Fur} dG and ^{CNPh} dG modified TBA	S23
S16, S17. Mol2 files of the ^{CNPh} dG and ^{Fur} dG nucleoside adducts	S24
Figure S18. Percent distribution of θ (degrees) and χ (degrees) dihedral and	gles
throughout the 20 ns MD simulations	S27
Table S2. Occupancies for the Hydrogen Bonds Between the G5-Adduct a	nd the Paired
Bases in G-quartet over the Duration of the MD Simulations	S28
Figures S19-S28. RMSDs for MD simulations	S27
	Experimental Section. Figure S1. DFT calculations of ^{Fur} dG Figure S2. DFT calculations of ^{CNPh} dG. CD spectra of modified TBA duplex structures. Figures S4-S12. NMR spectra of ^{CNPh} dG and 1-3 . Table S1. ESI-MS Analysis of 8-Aryl-G Modified TBA Oligonucleotides. Figures S13-S15. ESI-MS spectrum of ^{Fur} dG and ^{CNPh} dG modified TBA S16, S17. Mol2 files of the ^{CNPh} dG and ^{Fur} dG nucleoside adducts. Figure S18. Percent distribution of θ (degrees) and χ (degrees) dihedral and throughout the 20 ns MD simulations. Table S2. Occupancies for the Hydrogen Bonds Between the G ₅ -Adduct at Bases in G-quartet over the Duration of the MD Simulations. Figures S19-S28. RMSDs for MD simulations.

Experimental Section

Materials and Methods: Boronic acids, $Pd(OAc)_2$, 3,3',3"-phosphinidynetris-(benzenesulfonic acid) trisodium salt (TPPTS), and other commercial compounds were used as received. The synthesis of 8-bromo-2'-deoxyguanosine (8-Br-dG) was performed according to literature procedures by treating dG with N-bromosuccinimide in water–acetonitrile.¹ Suzuki cross-coupling reactions of boronic acids with 8-Br-dG to afford ^{Fur}dG,² and ^{CNPh}dG³ were performed as described previously by Western and coworkers.⁴ NMR spectra were recorded on Bruker 300 or 400 MHz spectrometers in either DMSO-d₆ or CDCl₃ referenced to TMS (0 ppm) or the respective solvent. All UV–vis and fluorescence emission spectra were recorded with baseline correction and stirring using 10 mm light path quartz glass cells. Any water used for buffers or spectroscopic solutions was obtained from a filtration system (18.2 M Ω). Highresolution mass spectra were recorded on an Agilent Q-Tof instrument, operating in nanospray ionization at µL/min detecting positive ions.

Oligonucleotide Synthesis: Oligonucleotide synthesis for the 8-aryl-dG modified TBA oligonucleotide (5'-GGTTG₅G₆TG₈TGGTTGG) with modifications at G₅, G₆ or G₈ = 2Fur dG or CNPh dG was carried out on a 1 µmol scale on a BioAutomation Corp. MerMade 12 automatic DNA synthesizer using standard β-cyanoethylphosphoramidite chemistry, as outlined previously.² Following synthesis, oligonucleotides were cleaved from the solid support, deprotected using 2 mL of 30% ammonium hydroxide solution at 55 °C for 12 h and purified by reverse phased HPLC. Synthesis of the ^{Fur}dG-phosphoramidite (3'-O-[(2-cyanoethoxy) (diisopropylamino)phosphino]-5'-O-(4,4'-dimethoxytrityl)-N²-dimethylformamidyl-8-(2"-furyl)-2'-deoxyguanosine) for use on the DNA synthesizer was performed as previously published.²

The nucleoside ^{CNPh}dG was converted into a phosphoramidite using the synthetic strategy outlined in Scheme 1.

Scheme 1. Synthesis of ^{CNPh}dG Phosphoramidite.



8-(4''-Cyanophenyl)-2'-deoxyguanosine (^{CNPh}dG): Pd(OAc)₂ (22mg, 0.1mmol), TPPTS (148 mg, 25 mmol), Na₂CO₃ (800 mg, 7.5 mmol), 8-Bromo-dG (1.30g, 3.75 mmol) and 4cyanophenyl-boronic Acid (0.85 g, 5.63 mmol) were placed in a round bottomed flask fitted with a condenser and reverse filled with argon. Degassed 2:1 H₂O:CH₃CN (35mL) solution was added and the solution was heated to 100°C for 4 hours. Following completion the mixture was diluted with 200 mL of H₂O and the pH was adjusted to 7.5 with 1.0 M aqueous HCl. The mixture was then cooled to 0°C, filtered and dried to yield 1.03 g (80% yield) of a white-grey powder. ¹H NMR (300 MHz, DMSO-d₆) δ: 7.97 (d, *J* = 8.3 Hz, 2H), 7.87 (d, *J* = 8.3 Hz, 2H), 6.87 (bs, 2H), 6.08 (t, *J* = 7.4 Hz, 1H), 5.52 (bs, 1H), 5.18 (bs, 1H), 4.36 (m, 1H), 3.82 (m, 1H), 3.68 (dd, *J* = 4.3, 12 Hz, 1H), 3.53 (dd, *J* = 4.3, 12 Hz, 1H), 3.16 (m, 1H), 2.05 (m,1H). ¹³C NMR (300 MHz, DMSO-d₆) δ: 160.7, 156.0, 152.6, 144.4, 135.0, 133, 132.5, 129.6, 128.0, 118.6, 117.9, 111.4, 88.1, 84.8, 71.2, 62.2, 36.9. HRMS calcd for C₁₇H₁₇N₆O₄⁺ [M+H⁺] 369.1306, found 369.1321. *N*²-(**Dimethylformamidyl**)-8-(4''-cyanophenyl)-2'-deoxyguanosine (1): ^{CNPh}dG (1.21 g, 3.3 mmol) was placed in 15 mL dry DMF under argon. *N*,*N*-Dimethylformamide diethyl acetal (2.7 mL,13.5 mmol) was then added and the mixture was allowed to stir for 5 hours. The reaction mixture was then evaporated to dryness and the solid washed with MeOH and dried to yield 1.02g (91% yield) of the white coloured product. ¹H NMR (300 MHz, DMSO-d₆) δ: 8.48 (s, 1H), 8.00 (d, *J* = 8.3 Hz, 2H), 7.85 (d, *J* = 8.3 Hz, 2H), 6.11 (t, *J* = 7.7 Hz, 1H), 5.75 (bs, 1H), 5.30 (bs, 1H), 4.44 (m, 1H), 3.86 (m, 1H), 3.67 (dd, *J* = 3.5, 12 Hz, 1H), 3.55 (dd, *J* = 3.8, 11.9 Hz, 1H), 3.24 (m, 1H), 3.09 (s, 3H), 3.00 (s, 3H), 2.07 (m, 1H). ¹³C NMR (300 MHz, DMSO-d₆) δ: 157.7, 151.6, 145.0, 135.0, 133.0, 132.6, 129.7, 128.1, 121.3, 118.6, 111.4, 88.3, 85.2, 71.3, 62.3, 40.5, 34.4. HRMS calcd for C₂₀H₂₂N₇O₄⁺ [M+H⁺] 424.1733, found 424.1714.

5'-O-(4,4'-Dimethoxytrityl)-N²-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-

deoxyguanosine (2): Compound **1** (1.14, 2.7 mmol) was co-evaporated from dry pyridine (3 x 5 mL) in a round bottomed flask. The round bottomed flask was then fitted with a constant pressure dropping funnel and reverse filled with argon. Pyridine (15 mL) was added to the round bottomed flask while DMT-Cl (1.28 g, 3.78 mmol) was placed in the dropping funnel, and dissolved in 15 mL dry pyridine. The addition of the DMT-Cl solution was added over 1 hour and the reaction was allowed to stir at room temperature for 3h under argon and was monitored by TLC. Upon completion, the mixture was diluted with methylene chloride (10mL) and washed with water (2 x 10 mL). Triethylamine (TEA, 1 mL) was added and the mixture was evaporated to dryness. The solid was then dissolved in dichloromethane (3 mL) and hexanes (10 mL) was added and the reaction was stirred overnight. The resulting suspension was filtered, and the solid was loaded onto a silica column and eluted with MeOH:CH₂Cl₂:TEA (5:90:5) to afford product **2** as a white powder (73% yield). ¹H NMR (300 MHz, DMSO-d₆) & 8.36 (s, 1H), 7.99 (d, *J* = 8.6

Hz, 2H), 7.71 (d, J = 8.6 Hz, 2H), 7.42-7.39 (m, 2H), 7.31-7.19 (m, 7H), 6.79-6.74 (m, 4H), 6.24-6.19 (m, 1H), 4.69 (m, 1H), 4.07 (m, 1H), 3.76 (s, 3H), 3.75 (s, 3H), 3.50 (m, 1H), 3.36-3.27 (m, 2H), 3.05 (s, 3H), 2.94 (s, 3H), 2.50 (TEA), 3.32 (m, 1H), 1.00 (TEA). ¹³C NMR (300 MHz, DMSO-d₆) δ : 159.0, 158.4, 158.38, 156.8, 152.1, 147.9, 145.4, 136.3, 136.2, 135.1, 132.6, 130.4, 130.3, 128.5, 128.2, 127.2, 121.3, 118.8, 113.4, 113.3, 86.6, 86.2, 84.6, 72.6, 64.8, 55.6, 41.6, 38.2, 35.3. HRMS calcd for C₄₁H₄₀N₇O₆⁺ [M+H⁺] 726.3040, found 726.3060.

3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-(4,4'-dimethoxytrityl)- N^2 -(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-deoxyguanosine (3): Compound 2 (0.500 g, 0.706 mmol) was co-evaporated from dry THF (3 x 5 mL), reverse filled with argon and dissolved in 10 mL dry, degassed CH₂Cl₂. To this was added dry, degassed TEA (0.4 mL, 2.83 mmol) and 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (0.24 mL, 1.06 mmol). The reaction was monitored via TLC and, upon completion, washed successively with saturated NaHCO₃ solution. The organic phase was separated, dried and loaded onto flash chromatography column eluting with 92:5:3 CH_2Cl_2 :MeOH:TEA. The phosphoramidite 3 was isolated as an off-white foam in 72% yield. ¹H NMR (400 MHz, CDCl₃) δ: 9.05-8.89 (m, 1H), 8.43 (s, 1H), 8.05 (d, J = 8.4 Hz, 2H), 7.72-7.67 (m, 2H), 7.34 (m, 2H), 7.26-7.13 (m, 8H), 6.72-6.67 (m, 4H), 6.12 (m, 1H), 5.16 (m, 1H), 4.20 (m, 1H), 3.73 (s, 3H), 3.72 (s, 3H), 3.64 (m, 1H), 3.55-2.27 (m, 7H), 3.05 (s, 3H), 2.92 (s, 3H), 2.50 (t, J = 6.1 Hz, 1H), 2.40-2.22 (m, 2H), 1.39 (d, J = 6.4 Hz, 1.13 (t, J = 7.4 Hz, 6H), 1.02 (t, J = 6.1 Hz, 6H). ¹³C NMR (400 MHz, CDCl₃) δ : 158.4, 157.6, 156.1, 151.2, 144.6, 135.8, 134.4, 132.2, 130.2, 130.0, 128.1, 127.7, 126.8, 121.3, 118.5, 117.5, 113.0 112.9, 86.1, 84.2, 83.8, 63.4, 62.7, 58.4, 55.2, 45.3, 43.3 24.6, 20.1. HRMS calcd for $C_{50}H_{57}N_9O_7P^+$ [M+H⁺] 926.4113, found 926.4130.

Oligonucleotide Purification and Sample Preparation: The 8-aryl-G-modified TBA oligonucleotide solutions were first filtered using syringe filters (PVDF 0.20 µm), and concentrated under diminished pressure. Purification was performed using an Agilent 1200 series HPLC instrument equipped with an autosampler, a diode array detector (monitored at 258 nm and λ_{Abs} of the incorporated modified nucleoside), fluorescence detector (monitored at λ_{ex} and λ_{em} of the incorporated modified nucleoside), and autocollector. Separation was carried out at 50 °C using a Phenomenex clarity 5 μ m reversed-phase (RP) semipreparative C18 column (100 \times 10 mm) with a flow rate of 3.5 mL/min, and various gradients of buffer B in buffer A (buffer A = 95:5 aqueous 50 mM TEAA, pH 7.2/acetonitrile; buffer B = 30:70 aqueous 50 mM TEAA, pH 7.2/ acetonitrile. Collected DNA samples were lyophilized to dryness and redissolved in 18.2 M Ω water for quantification by UV-vis measurement using ε_{260} . Extinction coefficients were obtained from the following website: http://www.idtdna.com/analyzer/applications/ oligoanalyzer. The 8-aryl-G modified TBA oligonucleotides were assumed to have the same extinction coefficient as the natural TBA oligonucleotide. In all cases of hybridization, oligonucleotides were annealed by heating at 90 °C for 10 min, cooling to room temperature, and refrigerating until analysis.

MS Analysis of Oligonucleotides: MS experiments for DNA identification were conducted on a Bruker AmaZon quadrupole ion trap SL spectrometer. Masses were acquired in the negative ionization mode with an electrospray ionization source. Oligonucleotide samples were prepared in 90% Milli-Q filtered water/10% methanol containing 0.1 mM ammonium acetate. Full scan MS spectra were obtained by direct infusion at a rate of 5–10 μ L/min. The capillary and cone voltages were optimized for each analyte and varied from 2.5 to 3.5 kV and

25-35 V, respectively. A source offset of 60 V was used for all samples. The desolvation temperature was between 250 and 350 °C.

UV Melting: All thermal melting temperatures (T_m) of oligonucleotides were measured using a Cary 300-Bio UV–vis spectrophotometer equipped with a 6 x 6 multicell Peltier blockheating unit using Hellma 114-QS 10 mm light path cells. Oligonucleotide samples were prepared in 100 mM M⁺-Phosphate, pH 7.0 with 100 mM MCl (M⁺ = Na⁺ or K⁺), using equivalent amounts (6 µM) of the unmodified or 8-aryl-dG modified TBA oligonucleotide and its complementary strand. The UV absorption (at 260 nm for duplexes and 295 nm for Gquadruplexes) was monitored as a function of temperature and consisted of forward-reverse scans from 10 to 90 °C at a heating rate of 1 °C/min, which was repeated five times. The T_m values were determined using hyperchromicity calculations provided in the Varian Thermal software.

Circular Dichroism (CD): Spectra were recorded on a Jasco J-815 CD spectropolarimeter equipped with a 1×6 Multicell block thermal controller and a water circulator unit. Measurements were carried out in 100 mM M⁺-phosphate, pH 7.0 with 100 mM MCl (M = Na⁺ or K⁺), using 6 μ M of the unmodified or 8-aryl-G modified TBA oligonucleotide and its complementary strand in equivalent amounts. Quartz glass cells (110-QS) with a light path of 1 mm were used for measurements. Spectra were collected at 10 °C between 200 and 400 nm, with a bandwidth of 1 nm and scanning speed at 100 nm/min. The spectra were the averages of five accumulations that were smoothed using the Jasco software.

Photophysical Properties of ^{CNPh}**dG:** Stock solutions of ^{CNPh}**dG** (4 mM) were prepared in DMSO due to limited solubility in other solvents. Experimental spectroscopic solutions contained \leq 1% DMSO and were prepared in reagent grade solvents (DMSO, acetonitrile, acetone, chloroform, and tetrahydrofuran) to a final volume of 2 mL containing 40 μ M ^{CNPh}dG. UV-Vis spectral measurements were observed from 220 to 400 nm with background correction to remove spurious solvent absorbance. Fluorescence emission spectra from 10 nm above the excitation wavelength to 600 nm were recorded with background correction using the absorbance maxima of ^{CNPh}dG as the excitation wavelength. All UV-Vis and fluorescence emission spectra were recorded with baseline correction and stirring using 10 mm light path quartz glass cells. Quantum yields for ^{CNPh}dG were determined, as outlined previously in detail,⁵ using the comparative method,⁶ with quinine bisulfate ($\Phi_{\eta} = 0.546$ in 0.5 M H₂SO₄) serving as the fluorescence quantum yield standard.⁷

Fluorescence of Modified TBA: All fluorescence spectra were recorded in 100 mM M⁺phosphate, pH 7.0 with 100 mM MCl (M = Na⁺ or K⁺) buffer . In each case, both excitation and emission spectra were recorded for the 8-aryl-G modified TBA oligonucleotide. All oligonucleotide samples were prepared to a final concentration of 6 μ M, and duplex samples were prepared using equivalent amounts (6 μ M) of the 8-aryl-G modified TBA oligonucleotide and its complementary strand. All measurements were made using quartz cells (108.002F-QS) with a light path of 10 × 2 mm; excitation and emission slit-widths were kept constant at 5 nm. All fluorescence excitation spectra were recorded at the emission wavelength (maximum) of the 8-aryl–dG probe, from 200 to 10 nm below the emission wavelength, while fluorescence emission spectra were recorded at the excitation wavelength (maximum) of the probe, from 10 nm above the excitation wavelength to 600 nm. Spectra were recorded at 10 °C.

DFT calculations at the nucleoside level: The preferred conformations of the ^{Fur}dG and ^{CNPh}dG adducts were analyzed based on our previous studies of the C8-pyrrole-dG (^{Pyr}dG)⁵ and *para*-phenoxyl-C8-dG ($^{p-PhOH}dG$)⁸ adducts. Specifically, our previous studies analyzed the

preferred conformations of ^{Pyr}dG and ^{*p*-PhOH}dG using B3LYP/6-31G(d) potential energy surface (PES) scans with respect to rotation about θ (the dihedral angle controlling the relative orientation of the 8-moiety and the nucleobase ring) and χ (the dihedral angle controlling the relative orientation of the nucleobase and the 2'-deoxyribose moiety). Full optimizations were subsequently performed at the same level to characterize the identified minima and transition states on the PES, and scaled zero-point vibrational energy (ZPVE) corrected relative energies were obtained at the B3LYP/6-311+G(2df,p) level.

In the present work, all minima and transition states previously identified for the ^{Pyr}dG and ^{*p*-PhOH}dG adducts were used to generate the initial conformations of the ^{Fur}dG and ^{CNPh}dG probes. Specifically, the pyrrole ring of ^{Pyr}dG was replaced with a furan ring to generate initial structures for ^{Fur}dG, whereas the hydroxyl group of ^{*p*-PhOH}dG was replaced with a cyano group to generate ^{CNPh}dG. For both adducts considered in the present work, full optimizations and frequency calculations were performed on minima and transition states at the B3LYP/6-31G(d) level, while higher-level (B3LYP/6-311+G(2df,p)) single-point calculations were performed to obtain relative energies, which include scaled ZPVE corrections. Although *syn* minima and transition states were previously reported for ^{Fur}dG,² these structures are presented in this paper along with the *anti* minima and transition states to provide complete information about the PES of these adducts, which includes *anti/syn* energy differences and rotational barriers.

The global minima identified for ^{CNPh}dG was used to calculate the B3LYP/6-31G(d) orbital energies and TD-B3LYP/6-31G(d) vertical excitation energies. Geometry optimization of the first excited singlet state of the ^{CNPh}dG adduct was carried out with CIS/6-31G(d) using the ground state geometries as a starting point. Analogous excited state calculations for the ^{Fur}dG

were reported in a previous study.² These calculations were performed using Gaussian 09 (Revision A.02).⁹

MD Simulations on TBA and Post-MD Analysis: The NMR-based structure of the non-modified 15-mer thrombin binding aptamer (PDB code: 148D, eighth frame) was used as a starting structure for MD simulations since a previous study identified this model as the best candidate.¹⁰ Structures of TBA containing modified (adducted) nucleosides at the G₅ or G₈ position were built by carrying out the appropriate 8 modification using Gaussview (Version 5.0).¹¹ Initial structures were built with the *syn*-conformations of the adducted nucleosides at the G₅ position (since only the *syn*, not *anti*, conformation at G₅ stabilizes the G-quadruplex), whereas both the *syn*- and *anti*-conformations of the adducted nucleosides were analyzed at the G₈ position (since the G₈ position is in a loop and can therefore adopt both conformations about the glycosidic bond). Two possible orientations of the 8-moiety with respect to the nucleobase, which roughly correspond to $\theta \sim 0^\circ$ or 180°, were considered for the *syn*- and *anti*-conformations of CNPhdG is symmetrical, only one possible θ value ($\theta \sim 0^\circ$) was considered. For comparison, simulations were also carried out on natural (unmodified) TBA starting from the NMR structure.

The parmbsc0¹² modification to the parm99¹³ force field was used for simulating the natural nucleosides (dA, dC, dG and dG), whereas GAFF¹⁴ parameters were used for the adducted nucleosides. Partial charges for the ^{Fur}dG and ^{CNPh}dG adducted nucleosides were calculated using the RED v.III.8¹⁵ program interfaced with Gaussian 09. Simulations were carried out by immersing DNA in a 15 Å octahedral box of TIP4P water molecules using the TLEAP module of AMBER 11.¹⁶ Since previous studies have suggested that a single potassium ion bonds the 15-mer TBA,¹⁷⁻¹⁹ the folded TBA was stabilized by placing a potassium cation in

the geometrical center of the G-quadruplex stem. In synchrony with previous MD simulations in TBA,¹⁰ the ion coordinates were determined by averaging the Cartesian coordinates of the O6 atoms of the guanines forming the quadruplex stem (i.e., G1, G2, G5, G6, G10, G11 and G15). The system was further neutralized by placing 13 sodium ions at random positions inside the water box. Minimization of solvent and ions was performed for 500 steps using the steepest descent algorithm followed by another 500 steps using the conjugate gradient minimization algorithm, with the DNA held fixed using a force constant of 500 kcal mol⁻¹ Å⁻². Subsequently, 2500 steps of minimization (1000 steps of steepest descent followed by 1500 steps of conjugate gradient) were performed on the entire system with constraints removed on DNA. The system was subsequently heated from 0 to 300 K with the DNA restrained using a force constant of 10 kcal mol⁻¹ Å⁻² for 20 ps under constant volume. Finally, a total of 20 ns of unrestrained MD simulations were carried out for each system at constant temperature (300 K) and pressure (1 atm) using AMBER 12.²⁰⁻²² The bonds involving hydrogen atoms were constrained using SHAKE and a 2 fs time step was used throughout the simulation. The stability of the MD simulations was confirmed through analysis of root mean square derivations (RMSD) of the backbone residues over the course of the simulation (Figures S18–S27).

A free energy analysis was carried out using snapshots from the MD trajectories after removing the water and ions. The molecular mechanics energy (which is the sum of the energy terms stemming from deviations of the bond lengths, bond angles and dihedral angles from their equilibrium values, the van der Waals interaction potential and the electrostatic energy) and the solvation free energy (which is the sum of the electrostatic solvation energy and the nonploar contribution to the solvation free energy) were calculated using the Poisson-Boltzmann method, whereas the entropy term was estimated from normal mode calculations using the MMPBSA²³

method implemented in AMBER 11 and 12. The molecular mechanics and solvation free energy terms were calculated using 1000 snapshots (1 snapshot every 20 ps), while the entropy term was calculated using 100 snapshots (1 snapshot every 200 ps) from simulation trajectories.

For each modified TBA aptamer, the conformation with the lowest free energy was used for further analysis. Specifically, a representative structure was chosen by clustering the entire associated simulation based on the location of the atoms forming the θ and χ dihedral angles of the adducted nucleoside in the G_5 or G_8 position. For the unmodified TBA, a representative structure was chosen for further analysis by clustering the entire simulation based on the atoms comprising the χ dihedral angle in the guanine residues at the G₅ and G₈ positions. DFT calculations were carried out on the representative structures to estimate the hydrogen-bonding and stacking interactions between the adduct and neighboring bases using the Q-Chem software.²⁴ The ω B97X-D/aug-cc-pVTZ level was used for this analysis since this method has been shown to provide reasonable estimates of both hydrogen bonding and stacking energies.²⁵ The default SCF convergence threshold of 10^{-5} was used along with the 10^{-10} cutoff for neglecting two-electron integrals (using the THRESH keyword). For these calculations, the sugar-phosphate backbone was removed and the dangling bonds were saturated with hydrogen atoms, where the hydrogen atom orientation corresponds to the heavy atom position and the C-Nbond distance is replaced by a typical C-H bond distance according to Gaussview.

REFERENCES

- 1. L. Gillet, O. D. Schärer, Org. Lett. 2002, 4, 4205-4208.
- 2. K. M. Rankin, M. Sproviero, K. Rankin, P. Sharma, S. D. Wetmore, R. A. Manderville, J. Org. Chem. 2012, 77, 10498-10508.
- 3. K. M. Schlitt, K. M. Sun, R. J. Paugh, A. L. Millen, L. Navarro-Whyte, S. D. Wetmore, R. A. Manderville, *J. Org. Chem.* 2009, **74**, 5793-5802.
- 4. E. C. Western, J. R. Daft, E. M. Johnson, II, P. M. Gannett, K. H. Shaughnessy, *J. Org. Chem.* 2003, **68**, 6767-6774.
- 5. K. M. Schlitt, A. L. Millen, S. D. Wetmore, R. A. Manderville, *Org. Biomol. Chem.*, 2011, 9,1565–1571.
- 6. S. Fery-Forgues, D. J. Lavabre, J. Chem. Educ. 1999, 76, 1260-1264.
- 7. J. N. Demas, G. A. Crosby, J. Phys. Chem. 1971, 75, 991-1024.
- 8. A. L. Millen, C. K. McLaughlin, K. M. Sun, R. A. Manderville, S. D. Wetmore, J. Phys. Chem. A, 2008, **112**, 3742–3753.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. M. Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09*, Revision A.02 ed.; Gaussian, Inc.: Wallingford CT, 2009.
- 10. R. Reshetnikov, A. Golovin, V. Spiridonova, A. Kopylov, A.; Šponer, J., J. Chem. Theory Comput., 2010, 6, 3003–3014.
- 11. R. Dennington, T. Keith, J. Millam, GaussView, Version 5, 2009 Semichem Inc., Shawnee Mission, KS.
- 12. A. Perez, I. Marchán, D. Svozil, J. Šponer, T. E. Cheatham III, C. A. Laughton, M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
- 13. T. E. Cheatham III, P. Cieplak, P. A. Kollman, J. Biomol. Struct. Dyn., 1999, 16, 845-862.
- J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, J. Comput. Chem., 2004, 25, 1157–1174.
- F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821–7839.
- 16. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. E., R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, P. A. Kollman, *AMBER 11*, 2010, University of California, San Francisco, CA.
- 17. N. V. Hud, F. W Smith, F. A. Anet, J. Feigon, Biochemistry, 1996, 35, 15383–15390.
- 18. M.Vairamani, M. L. Gross, J. Am. Chem. Soc., 2003, 125, 42-43.

- 19. P. R. Majhi, J. Qi, C. Tang, R. H. Shafer, *Biopolymers*, 2008, 89, 302–309.
- D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F.Wong, F. Paesani, J. Vanicek, R.M.Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, *AMBER 12*, 2012, University of California, San Francisco.
- 21. A. W. Goetz, M. J. Williamson, D. Xu, D. Poole, S. L. Grand, R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
- 22. R. Salomon-Ferrer, A. W. Goetz, D. Poole, S. L. Grand, R. C. Walker, J. Chem. Theory Comput. 2012, DOI: 10.1021/ct400314y.
- 23. B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, J. Chem. *Theory Comput.*, 2012, **8**, 3314–3321.
- 24. A. Aspuru-Guzik, C. Chang, R. G. Edgar, E. Sundstrom, J. Parkhill, K. Lawler, M. Gordon, M. Schmidt, N. Shenvi, D. Lambrecht, M. Goldey, R. Olivares-Amaya, Y. Bernard, L. Vogt, M. Watson, J. Liu, S. Yeganeh, B. Kaduk, O. Vydrov, X. Xu, I. Kaliman, K. Khistyaev, N. Russ, I.Y. Zhang, W.A. Goddard III, F. Liu, R. King, A. Landau, M. Wormit, A. Dreuw, M. Diedenhofen, A. Klamt, A.W. Lange, D. Ghosh, D. Kosenkov, T. Kus, A. Landau, D. Zuev, J. Deng, S.P. Mao, Y.C. Su, D. Small, L. D. Jacobson. *Q-Chem*, Version 4.0.1, Q-Chem, Inc., Pittsburgh, PA (2007).
- 25. L. A. Burns, Á. Vázquez-Mayagoitia, B. G. Sumpter, C. D. Sherrill, *J. Chem. Phys.*, 2011, **134**, 084107.



Figure S1. B3LYP/6-31G(d) Twist Angle (θ , deg.) for Minima and Transition States of ^{Fur}dG (ZPVE Corrected B3LYP/6-311+G(2df,p) Relative Energies from Single-Point Calculations Provided in Parenthesis, kJ mol⁻¹).

Electronic Supplementary Material (ESI) for Chemical Science This journal is C The Royal Society of Chemistry 2013



Figure S2. B3LYP/6-31G(d) Twist Angle (θ , deg.) for Minima and Transition States of ^{CNPh}dG (ZPVE Corrected B3LYP/6-311+G(2df,p) Relative Energies from Single-Point Calculations Provided in Parenthesis, kJ mol⁻¹).



Figure S3. Circular Dichroism spectra of 6 uM modified TBA duplexes with modifications at G_5 , G_6 and G_8 . Modified TBA = (5'-GGTTG₅G₆TG₈TGGTTGG) complimentary strand = (5'-CCAACCAACCACCAACC). ^{Fur}dG (Red) and ^{CNPh}dG (Blue) modifications at G₅ (solid traces), G₆ (dotted traces) and G₈ (dashed traces).



Figure S4. ¹H NMR of ^{CNPh}dG in DMSO-*d*₆.



Figure S5. ¹³C NMR of ^{CNPh}dG in DMSO-*d*₆.



Figure S6. ¹H NMR of **1** in DMSO- $d_{6.}$



Figure S7. ¹³C NMR of **1** in DMSO- $d_{6.}$



Figure S8. ¹H NMR of 2 in DMSO- $d_{6.}$



Figure S9. ¹³C NMR of **2** in DMSO- d_{6} .



Figure S10. ¹H NMR of 3 in CDCl_{3.}



Figure S11. ¹³C NMR of 3 in CDCl_{3.}



Figure S12. ³¹P NMR of 3 in CDCl_{3.}

Table S1. ESI-MS Analysis of 8-Aryl-G Modified TBA Oligonucleotides

Oligonucleotide	Product Formula	calcd	exptl m/z (ESI) ^b	exptl mass
		mass ^a		
$TBA(X = {}^{CNPh}G)$	$C_{157}H_{190}N_{58}O_{94}P_{14}$	4824.81	$[M - 7H]^{7-} = 688.7$	4824.4
			$[M - 8H]^{8-} = 602.1$	4824.8
			$[M - 9H]^{9-} = 535.5$	4824.5
			$[M - 10H]^{10-} = 481.5$	4825.0
TBA ($X = FurG$)	$C_{154}H_{189}N_{57}O_{95}P_{14}$	4789.90	$[M - 7H]^{7} = 683.3$	4790.1
			$[M - 8H]^{8} = 597.7$	4789.6
			$[M - 9H]^{9} = 531.3$	4790.0
_			$[M - 10H]^{10} = 478.0$	4790.0
TBA (X = ^{Fur}G ,	$C_{161}H_{192}N_{58}O_{95}P_{14}$	4890.83	$[M - 6H]^{6-} = 814.1$	4890.6
^{CNPh} G)			$[M - 7H]^{7-} = 697.8$	4891.6
			$[M - 8H]^{8-} = 610.4$	4891.2
			$[M - 9H]^{9-} = 542.5$	4891.5

^{*a*} Monoisotopic mass of most abundant isotopologue; assumes one ¹³C isotope. ^{*b*} Measured m/z from mass spectrum.



Figure S13. ESI-MS spectrum of TBA ($X = {}^{Fur}dG$ at G_5).



Figure S14. ESI-MS spectrum of TBA ($X = {}^{CNPh}dG$ at G_5).



Figure S15. ESI-MS spectrum of TBA ($X = {}^{Fur}dG$ at G_5 , ${}^{CNPh}dG$ at G_8).

S16. Mol2 file of the ^{CNPh}dG nucleoside adduct.

@ <tripos< td=""><td>>MOLECU</td><td>JLE</td><td></td><td></td><td></td></tripos<>	>MOLECU	JLE			
LIG					
44 47	1 0 0				
SMALL					
No Charge o	r Current C	Charge			
@ <tripos< td=""><td>>ATOM</td><td></td><td></td><td></td><td></td></tripos<>	>ATOM				
1 P	4.6450	0.0580 -	1.3020 P	1 LIG	1.166100
2 O	4.1900	1.0240	0.0000 OS	1 LIG	-0.498000
3 01	4.5440	0.8590	-2.5640 O2	1 LIG	-0.768000
4 O2	5.8540	-0.6970	-0.8410 O2	1 LIG	-0.768000
5 O3	0.0000	0.0000	0.0000 OS	1 LIG	-0.541400
6 C5*	3.1300	1.9330	-0.2340 CI	1 LIG	-0.025000
7 H5*1	3.4200	2.7280	-0.9400 H1	1 LIG	0.086400
8 H5*2	2.8980	2.4100	0.7260 H1	1 LIG	0.086400
9 C4*	1.8670	1.2590	-0.7730 CT	1 LIG	0.155400
10 H4*	1.0400	1.9790	-0.7790 H1	1 LIG	0.108600
11 O4*	2.1100	0.8310	-2.1350 OS	1 LIG	-0.411200
12 C1*	1.8770	-0.5650	-2.2640 CT	1 LIG	0.189500
13 H1*	0.8700	-0.7500	-2.6430 H2	1 LIG	0.074500

14 C3*	1.4280 -0.0000 -0.0000 CT	1 LIG 0.132000
15 H3*	1.8260 -0.0030 1.0230 H1	1 LIG 0.078800
16 C2*	2.0030 -1.1440 -0.8460 CT	1 LIG -0.058300
17 H2*1	3.0470 -1.3160 -0.5830 HC	1 LIG 0.045200
18 H2*2	1.4450 -2.0790 -0.7330 HC	1 LIG 0.045200
19 N	2.7980 -1.1020 -3.2600 N*	1 LIG -0.080800
20 C	2.4890 -1.7250 -4.4810 CK	1 LIG 0.244200
21 C6	4.1630 -1.2100 -3.1260 CB	1 LIG 0.165300
22 N1	3.5530 -2.2390 -5.0610 NB	1 LIG -0.497800
23 C7	4.6050 -1.9240 -4.2360 CB	1 LIG 0.165500
24 N2	4.9100 -0.6900 -2.1180 NC	1 LIG -0.472000
25 C8	6.0130 -2.2030 -4.3790 C	1 LIG 0.444000
26 C9	6.2010 -0.9240 -2.2280 CA	1 LIG 0.536200
27 O5	6.6250 -2.8130 -5.2350 O	1 LIG -0.522700
28 N3	6.7370 -1.6290 -3.2700 NA	1 LIG -0.398300
29 H	7.7290 -1.8360 -3.2890 H	1 LIG 0.341500
30 N4	7.0410 -0.4860 -1.2400 N2	1 LIG -0.827100
31 H6	7.9980 -0.2700 -1.4830 H	1 LIG 0.387000
32 H7	6.6120 0.1600 -0.5850 H	1 LIG 0.387000
33 C10	1.1540 -1.7580 -5.0960 CA	1 LIG 0.037700
34 C11	0.7830 -2.9000 -5.8270 CA	1 LIG -0.086300
35 H8	1.4780 -3.7300 -5.8870 HA	1 LIG 0.141200
36 C12	0.2720 -0.6630 -5.0470 CA	1 LIG -0.106900
37 H9	0.5660 0.2520 -4.5420 HA	1 LIG 0.135300
38 C13	-0.4470 -2.9630 -6.4670 CA	1 LIG -0.182800
39 H10	-0.7310 -3.8500 -7.0250 HA	1 LIG 0.177400
40 C14	-0.9630 -0.7220 -5.6820 CA	1 LIG -0.151700
41 H11	-1.6380 0.1270 -5.6450 HA	1 LIG 0.173500
42 C15	-1.3360 -1.8760 -6.3910 CA	1 LIG 0.004100
43 C16	-2.6120 -1.9410 -7.0400 CZ	1 LIG 0.340400
44 N5	-3.6500 -1.9950 -7.5630 NZ	1 LIG -0.452100

S17. Mol2 file of the ^{Fur}dG nucleoside adduct. @<TRIPOS>MOLECUILF

@ <tripc< th=""><th>DS>MOLEC</th><th>ULE</th><th></th><th></th><th></th></tripc<>	DS>MOLEC	ULE			
LIG					
40 43	1 0 0				
SMALL					
No Charge	e or Current	Charge			
@ <tripc< td=""><td>OS>ATOM</td><td></td><td></td><td></td><td></td></tripc<>	OS>ATOM				
1 P	4.6680	0.0340	-1.1840 P	1 LIG	1.166200
2 O	4.1610	1.1180	0.0000 OS	1 LIG	-0.490300
3 01	4.5670	0.6930	-2.5260 O2	1 LIG	-0.768000
4 O2	5.8900	-0.6300	-0.6270 O2	1 LIG	-0.768000
5 O3	0.0000	0.0000	0.0000 OS	1 LIG	-0.549700
6 C5*	3.0790	1.9630	-0.3460 CI	1 LIG	-0.007000
7 H5*	1 3.3670	2.7080	-1.1050 H1	1 LIG	0.080300

8 H5*2	2.7870 2.5080 0.5600 H1	1 LIG 0.080300
9 C4*	1.8710 1.1920 -0.8760 CT	1 LIG 0.132000
10 H4*	1.0190 1.8740 -0.9920 H1	1 LIG 0.104700
11 O4*	2.2050 0.6470 -2.1720 OS	1 LIG -0.394800
12 C1*	1.9390 -0.7490 -2.2040 CT	1 LIG 0.134700
13 H1*	0.9510 -0.9370 -2.6240 H2	1 LIG 0.090700
14 C3*	1.4290 -0.0000 -0.0000 CT	1 LIG 0.186900
15 H3*	1.8230 0.0850 1.0200 H1	1 LIG 0.056600
16 C2*	2.0060 -1.2120 -0.7420 CT	1 LIG -0.065800
17 H2*1	3.0450 -1.3780 -0.4440 HC	1 LIG 0.042400
18 H2*2	1.4360 -2.1300 -0.5710 HC	1 LIG 0.042400
19 N	2.9060 -1.3870 -3.0990 N*	1 LIG -0.005900
20 C	2.6710 -2.1590 -4.2480 CK	1 LIG 0.222900
21 C6	4.2720 -1.2120 -3.0540 CB	1 LIG 0.111300
22 N1	3.7820 -2.4640 -4.8880 NB	1 LIG -0.566500
23 C7	4.7860 -1.8740 -4.1650 CB	1 LIG 0.246000
24 N2	4.9610 -0.5170 -2.1120 NC	1 LIG -0.513300
25 C8	6.2100 -1.8680 -4.3940 C	1 LIG 0.449500
26 C9	6.2620 -0.4890 -2.3010 CA	1 LIG 0.614600
27 O5	6.8810 -2.3760 -5.2730 O	1 LIG -0.527600
28 N3	6.8670 -1.1120 -3.3570 NA	1 LIG -0.461900
29 H	7.8780 -1.1130 -3.4380 H	1 LIG 0.356000
30 N4	7.0530 0.1500 -1.3800 N2	1 LIG -0.846700
31 H6	7.9250 0.5500 -1.7000 H	1 LIG 0.384200
32 H7	6.5340 0.7320 -0.7300 H	1 LIG 0.384200
33 C10	1.3800 -2.6130 -4.7270 C*	1 LIG 0.138700
34 C11	1.0320 -3.1460 -5.9390 C*	1 LIG -0.153100
35 H8	1.7010 -3.2860 -6.7750 HA	1 LIG 0.172700
36 C12	-0.3570 -3.4670 -5.8560 C*	1 LIG -0.215700
37 H9	-0.9740 -3.9040 -6.6290 HA	1 LIG 0.186300
38 C13	-0.7610 -3.1100 -4.6060 CA	1 LIG -0.131400
39 H10	-1.7030 -3.1480 -4.0810 H4	1 LIG 0.199400
40 O6	0.2850 -2.5900 -3.8980 OS	1 LIG -0.117100



Figure S18. Radar Plots for the Percent Distribution of the χ (a and c, deg.) and θ (b and d, deg.) Dihedral Angles Throughout the 20 ns MD Trajectories Corresponding to the Energetically Most Stable Conformations of the ^{CNPh}dG (blue) and ^{Fur}dG (red). Adducts at the (a and b) G₅ or (c and d) G₈ Position of the Modified Thrombin Binding Aptamer.

G5 position	bond	% occupancy
G	N2-H(G5)N7(G2)	99.7
	N1-H(G5)O6(G2)	99.1
	N2-H(G11)N7(G5)	99.3
	N1-H(G11)O6(G5)	97.9
^{Fur} G	$N2-H(^{Fur}G5)N7(G2)$	99.4
	$N1-H(^{Fur}G5)O6(G2)$	99.2
	$N2-H(G11)N7(^{Fur}_{T}G5)$	97.1
	$N1-H(G11)O6(^{Fur}G5)$	98.8
CINPNG	N2-H(CNPnG5)N7(G2)	98.9
	N1-H(CNPhG5)O6(G2)	99.1
	N2-	91.0
	$H(G11)N7(^{CNPh}G5)$	99.5
	NÌ-	
	$H(G11)O6(^{CNPh}G5)$	

Table S2. Occupancies for the Hydrogen Bonds Between the G_5 -Adduct and the Paired Bases in G-quartet over the Duration of the MD Simulations.^a

^aH-bond distance cut-off implemented was a 3.40 Å heavy atom separation and a 120° X-H-X angle. Only H-bonds with occupancy greater than 10% are reported.



Figure S19. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG adduct Placed at the G5 position in *syn* Orientation and Initial $\theta \sim 0^\circ$.



Figure S20. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG Adduct Placed at the G5 position in *syn* Orientation and Initial θ ~180°.



Figure S21. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG Adduct Placed at the G8 position in *anti* Orientation and Initial $\theta \sim 0^\circ$.



Figure S22. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG Adduct Placed at the G8 position in *anti* Orientation and Initial θ ~180°.



Figure S23. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG Adduct Placed at the G8 position in *syn* Orientation and Initial $\theta \sim 0^\circ$.



Figure S24. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{Fur}dG Adduct Placed at the G8 position in *syn* Orientation and Initial θ ~180°.



Figure S25. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{CNPh}dG Adduct Placed at the G5 position in *syn* Orientation.



Figure S26. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{CNPh}dG Adduct Placed at the G8 position in *anti* Orientation.



Figure S27. Backbone RMSD from MD Simulations on Thrombin Binding Aptamer with the ^{CNPh}dG Adduct Placed at the G8 position in *syn* Orientation.



Figure S28. Backbone RMSD from MD Simulations on the Natural (non-modified) Thrombin Binding Aptamer.