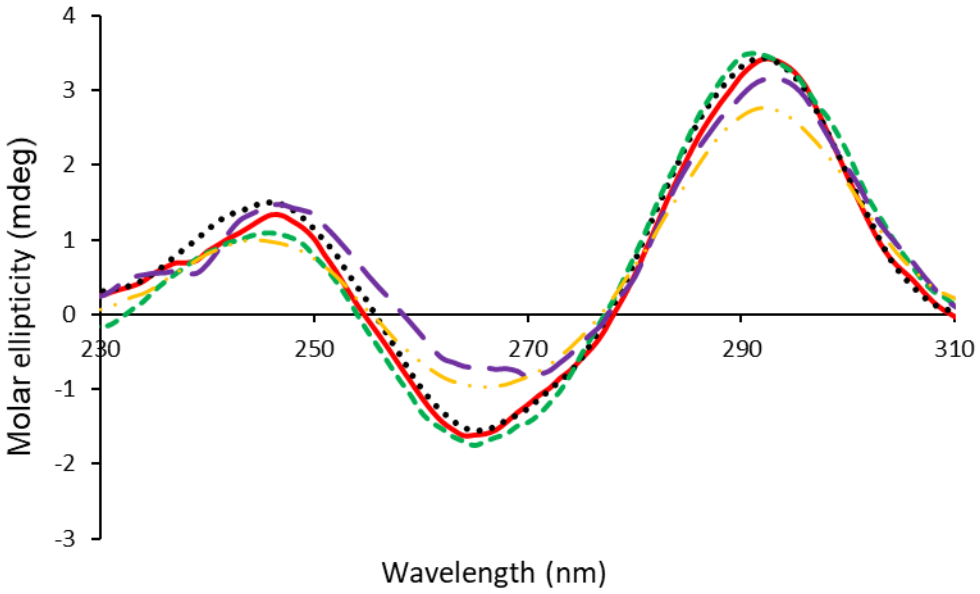
Supporting Information for:

Manipulation of a DNA Aptamer-Protein Binding Site through Arylation of Internal Guanine Residues.

Abigail J. Van Riesen, Kaila L. Fadock, Prashant S. Deore, Richard A. Manderville,* Shahin Sowlati-Hashjin, and Stacey D. Wetmore*

Table of Contents:

(a)



(b)

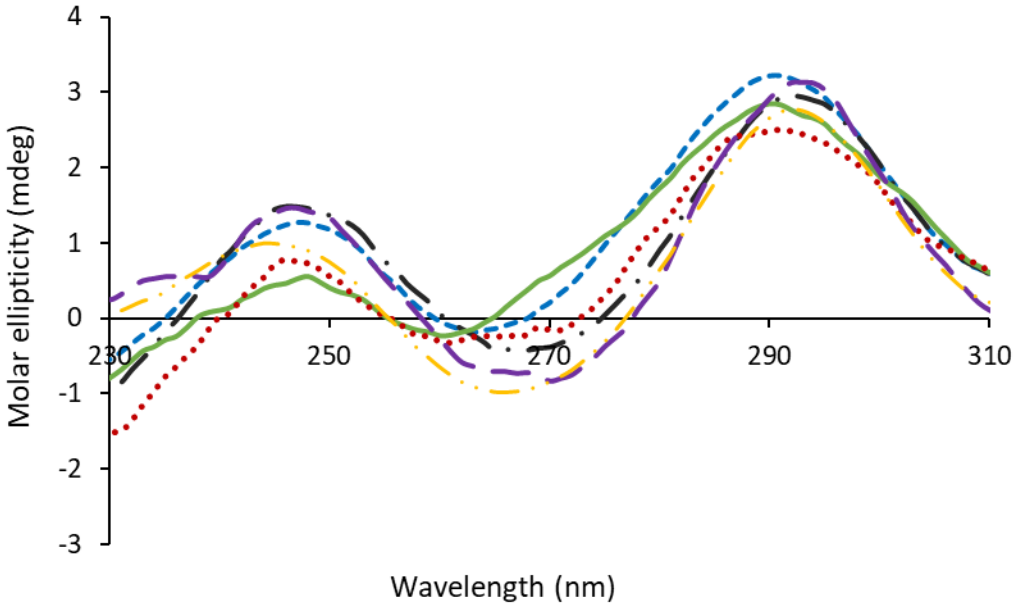


Figure S1. Circular Dichroism of mTBAs (3 µM) recorded in 100 mM phosphate buffer, pH 7 with 100 mM KCl at 10 °C for: (a) position G8 modification containing dG (purple long dash trace), 5' FAM label (yellow long dash dot dot), ^{CN}dG (green square dot trace), ^{Sty}dG (blue long dash dot trace), ^{BTh}dG (black round dot trace) and ThdG (red solid trace), and (b) position G5 and G6 modification containing dG (purple long dash trace), 5' FAM label (yellow long dash dot dot), ^{CN6}dG (green solid trace), ^{CN5}dG (red round dot trace), ^{Sty6}dG (blue square dot trace), ^{Sty5}dG (black long dash dot).

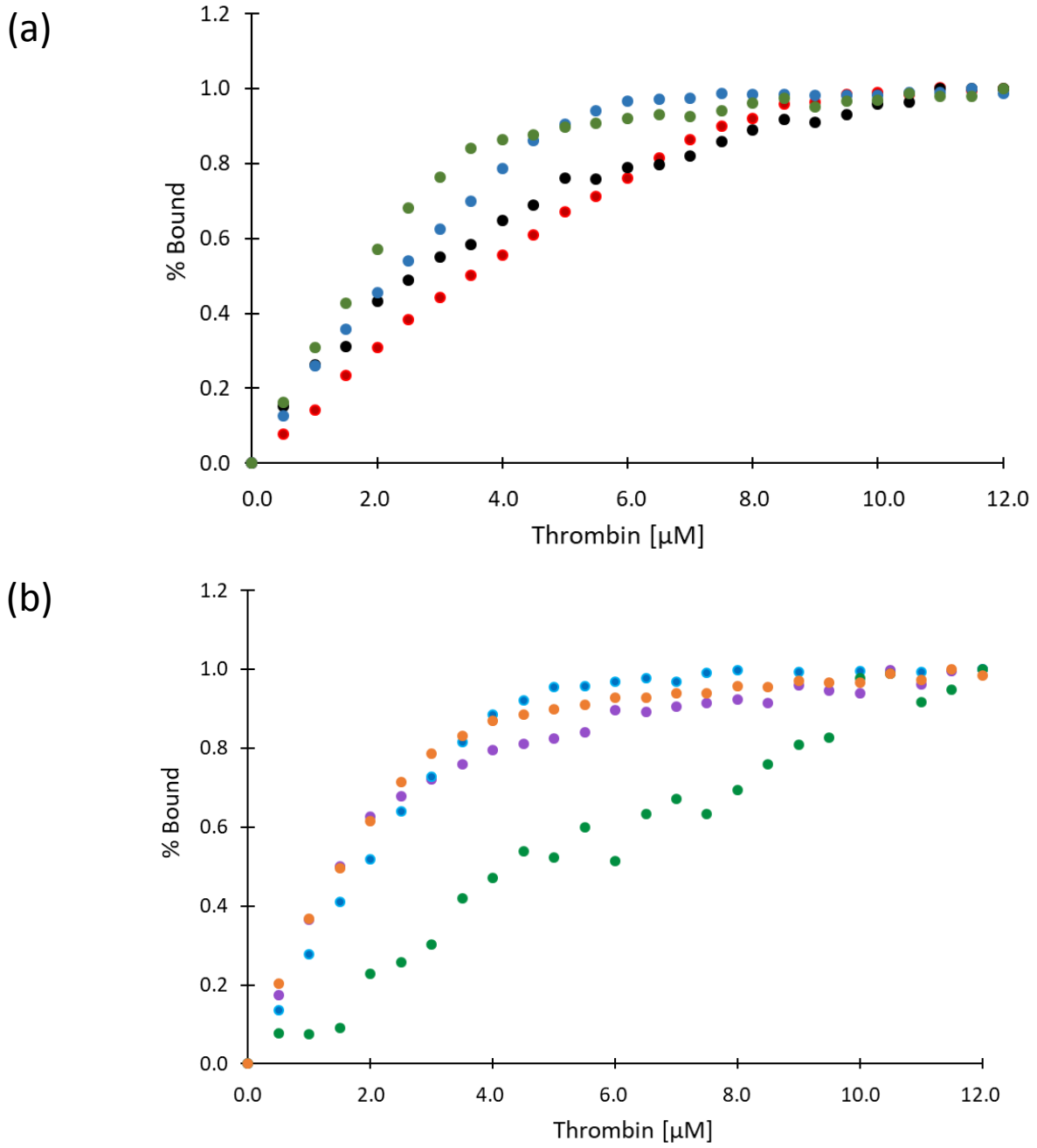
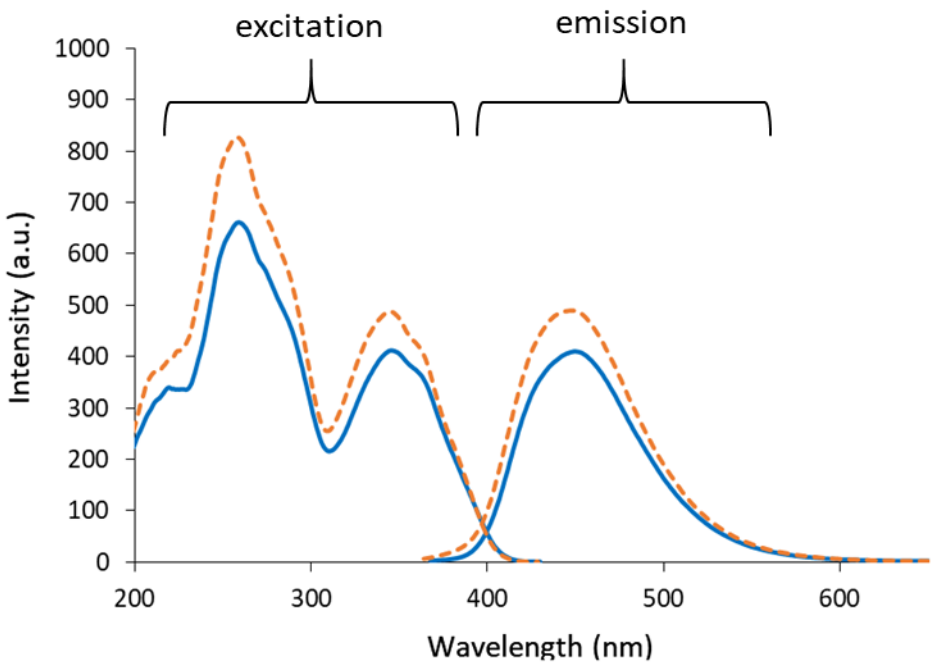


Figure S2. Polarization of mTBAs (3 µM) upon binding thrombin at 25 °C in 100 mM phosphate buffer, pH 7 with 100 mM KCl for: (a) position G8 modification containing ^{CN}dG (green trace), ^{Sty}dG (blue trace), ^{BTh}dG (black trace) and ThdG (red trace), and (b) positions G5 and G6 modification containing ^{Sty5}dG (blue trace), ^{Sty6}dG (orange trace), ^{CN5}dG (green trace), ^{CN6}dG (purple trace).

(a)



(b)

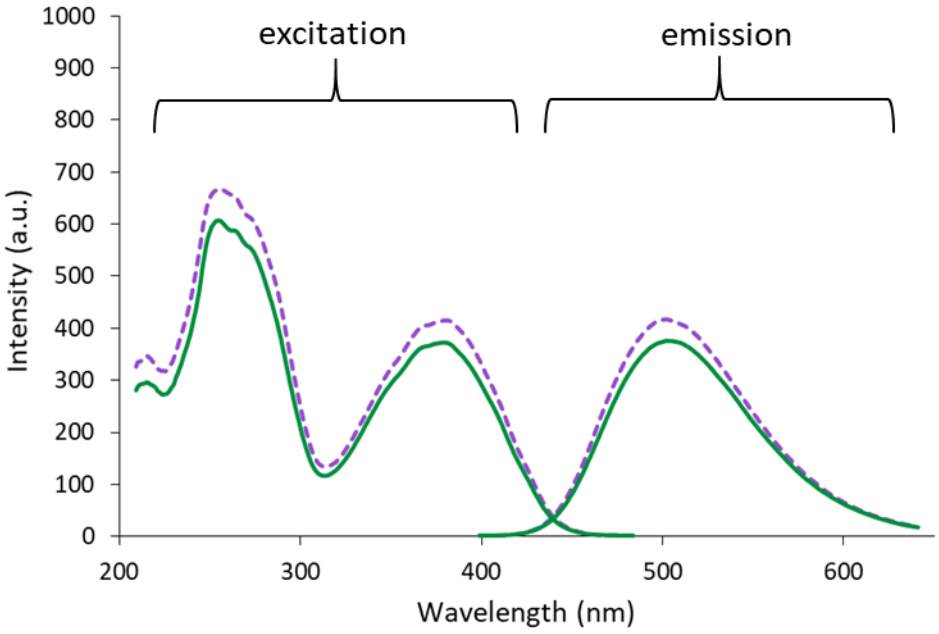


Figure S3. Fluorescence spectra of mTBAs (3 µM) at 25 °C in 100 mM phosphate buffer, pH 7 with 100 mM KCl for (a) ^{Sty}dG position G5 (solid blue trace) and G6 (dashed orange trace), and (b) ^{CN}dG for G5 (solid green trace) and G6 (dashed purple trace).

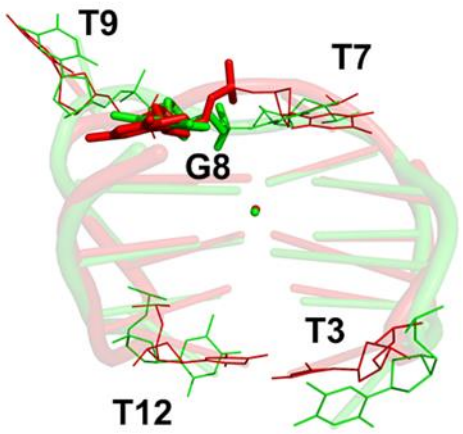


Figure S4. Overlay of the representative MD structure (red) and X-ray crystal structure (PDB ID: 4DII; green) of the native TBA (RMSD = 1.249 Å).

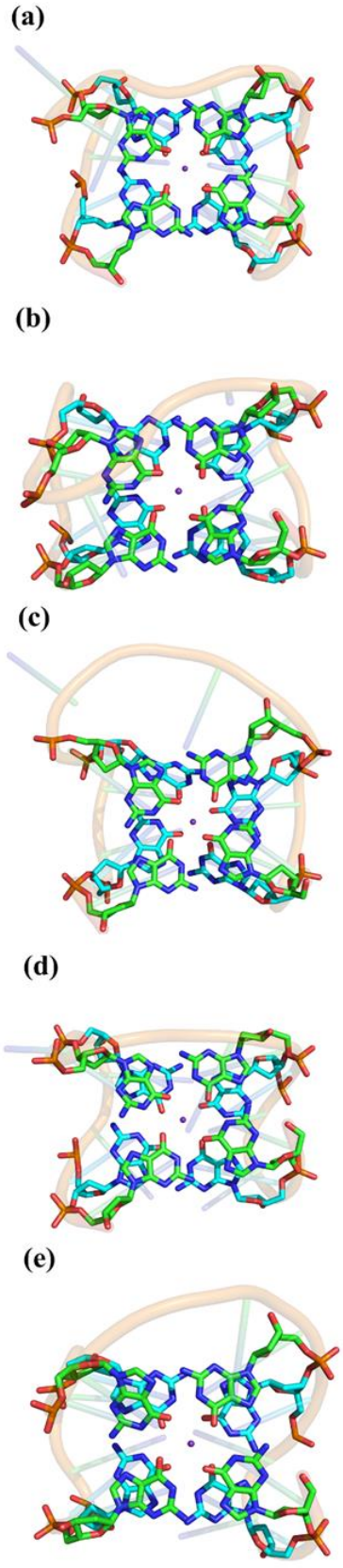


Figure S5. Structure of the TBA highlighting the G-quadruplexes for (a) dG, (b) *anti* ThdG, (c) *syn* ThdG, (d) *anti* ^{CN}dG or (e) *syn* ^{CN}dG at G8. G1G6G10G15 are shown on top and G2G5G11G14 on bottom.

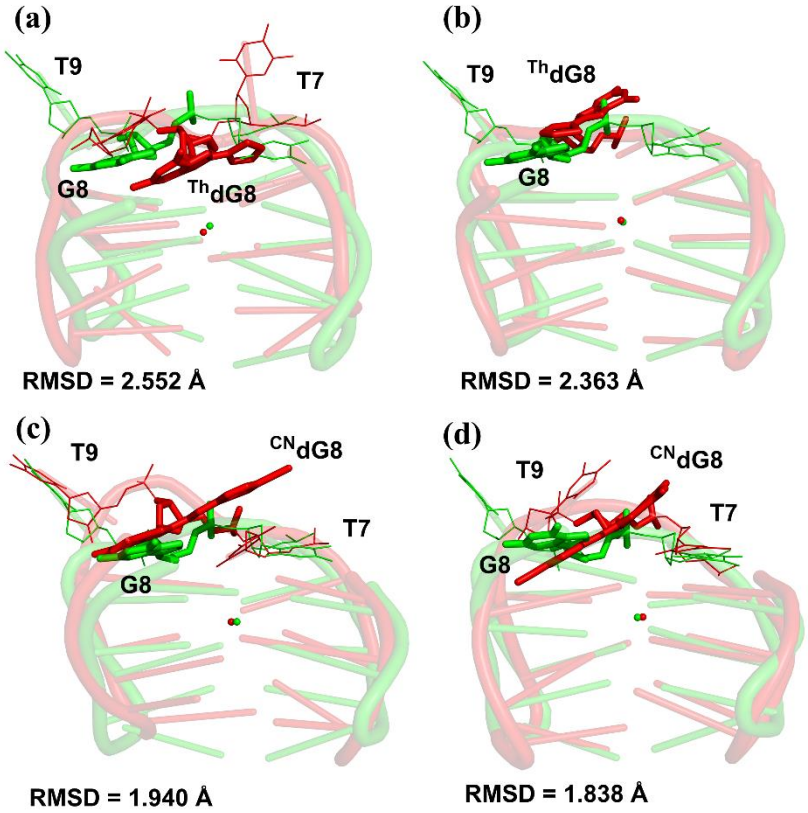


Figure S6. Overlay of representative MD structures of native (green) and modified TBA (red) containing (a) *anti* ThdG, (b) *syn* ThdG, (c) *anti* ^{CN}dG or (d) *syn* ^{CN}dG at G8. RMSD calculated according to the position of all heavy atoms for representative MD structures of modified with respect to native TBA.

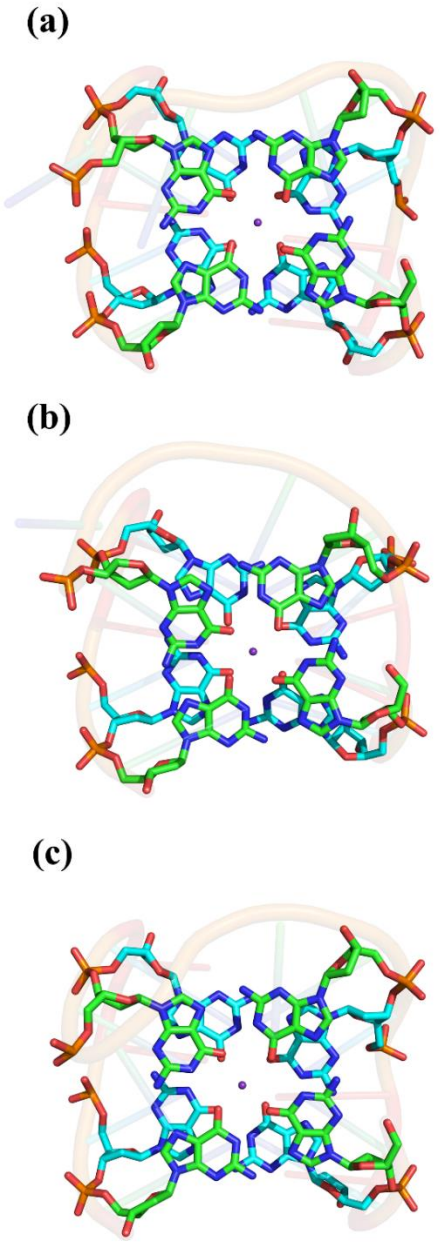


Figure S7. Structure of TBA in the TBA–thrombin complex highlighting the G-quadruplexes for TBA containing (a) native TBA, (b) *syn* ThdG or (c) *syn* ^{CN}dG at G8. G1G6G10G15 are shown on top and G2G5G11G14 on bottom.

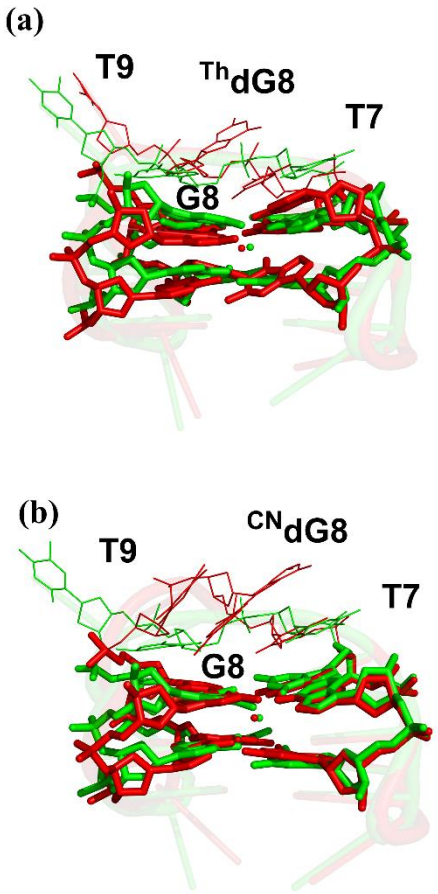


Figure S8. Overlay of representative MD structure of native (green) and modified (red) TBA bound to thrombin for TBA containing (a) *syn* ThdG or (b) *syn* ^{CN}dG at G8.

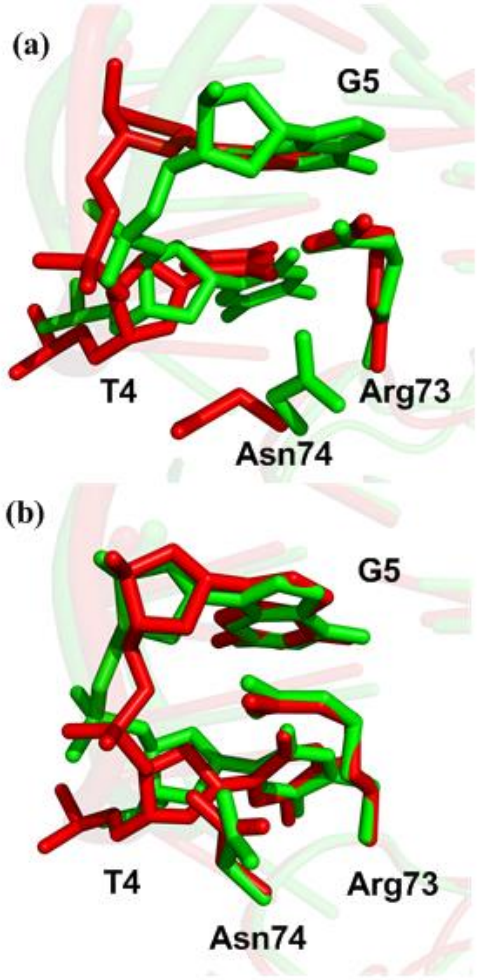


Figure S9. Overlay of TBA–protein interactions in the Arg73 region for representative MD structures of native (green) and modified (red) TBA with (a) *syn* ThdG or (b) *syn* ^{CN}dG at G8.

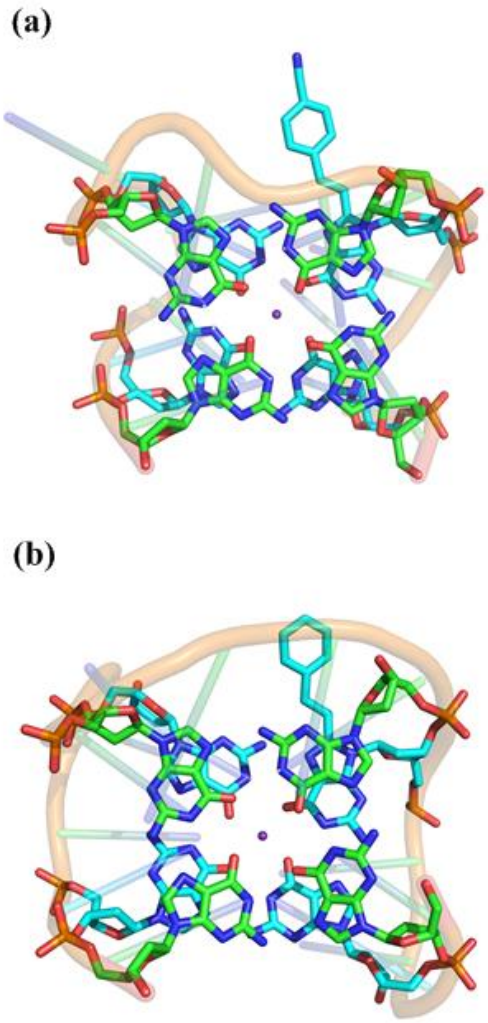


Figure S10. Structure of the TBA highlighting the G-quadruplexes for (a) ^{CN}dG and (b) ^{Sty}dG at G5. G1G6G10G15 are shown on top and G2G5G11G14 on bottom.

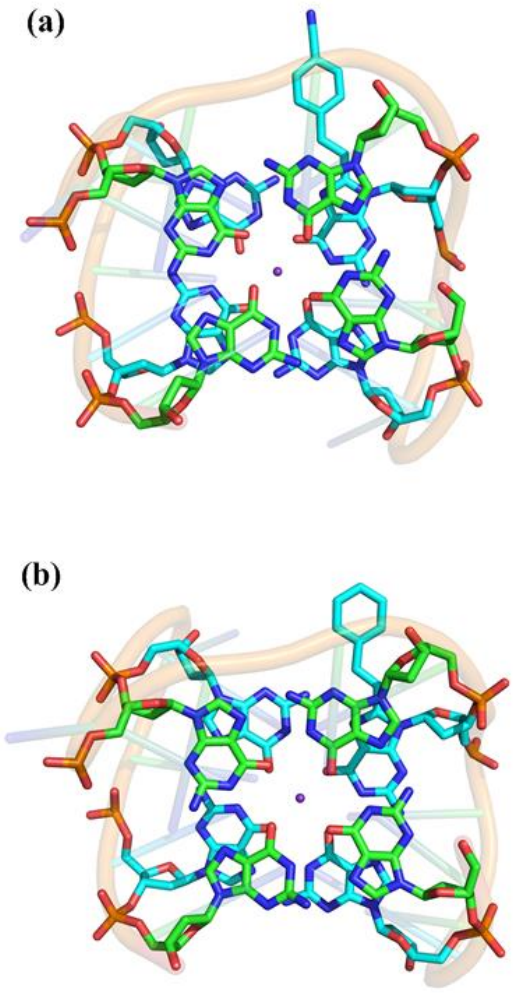


Figure S11. Structure of TBA in the TBA–thrombin complex highlighting the G-quadruplexes for TBA containing (a) ^{CN}dG and (b) ^{Sty}dG at G5. G1G6G10G15 are shown on top and G2G5G11G14 on bottom.

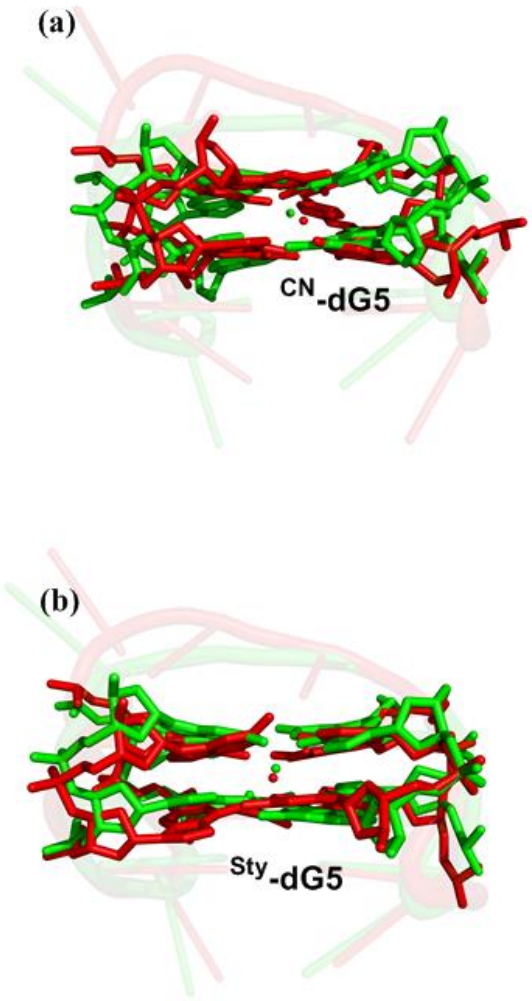


Figure S12. Overlay of representative MD structure of native (green) and modified (red) TBA bound to thrombin for TBA containing (a) ^{CN}dG and (b) ^{Sty}dG at G5.

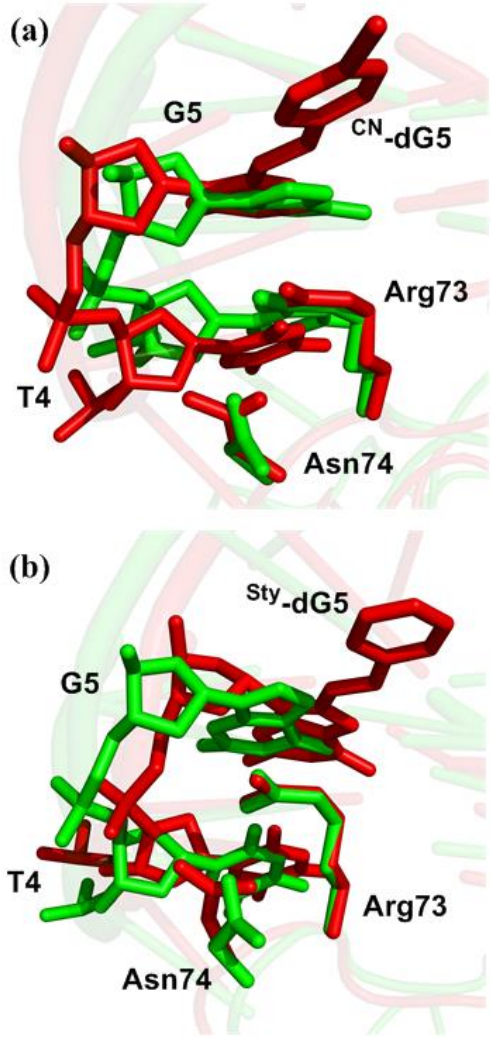


Figure S13. Overlay of TBA–protein interactions in the Arg73 region for representative MD structures of native (green) and modified (red) TBA with (a) ^{CN}dG and (b) ^{Sty}dG at G5.

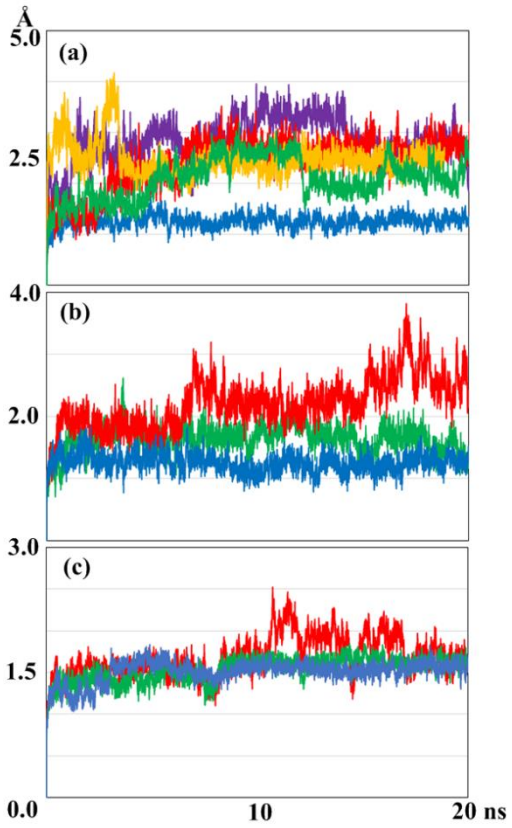


Figure S14. Root-mean-square-deviation (RMSD; Å) in the backbone of (a) unbound TBA, (b) TBA bound to thrombin, and (c) thrombin over the entire simulation trajectory (ns), with native TBA in blue and modified TBA containing *anti* ThdG at G8 in purple, *syn* ThdG in red, *anti* ^{CN}dG in orange, and *syn* ^{CN}dG in green.

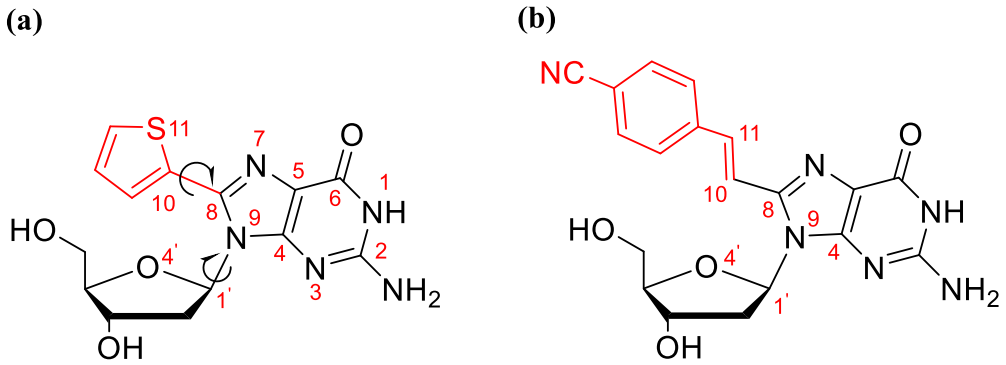


Figure S15. Structures of *anti*-ThdG and *anti*-^{CN}dG nucleosides. The dihedral angle χ (∠(O4′–C1′–N9–C4)) defines the glycosidic bond orientation to be *anti* (χ = 180 ± 90°) or *syn* (χ = 0 ± 90°), and θ (∠(N9–C8–C10–X11)) defines the degree of twist between the nucleobase and the C^8-substituent.

Table S1. Hydrogen-bonding occupancies (%) of G-quadruplexes in unbound TBA.[a]

		G8					G5	
Pair	**Interaction**	**native**	***anti* ThG**	***syn* ThG**	***anti* ^{CN}G**	***syn* ^{CN}G**	***syn* ^{CN}G**	***syn* StG**
G1-G6	N1–H···O6	100	99	99	99	99	99	100
	N2–H···N7	99	84	92	97	98	97	98
G1-G15	N7···H–N2	99	100	81	96	98	99	98
	O6···H–N1	99	84	89	95	94	98	98
G2-G5	N7···H–N2	100	100	98	99	99	98	99
	O6···H–N1	98	49	89	99	99	98	99
G2-G14	N1–H···O6	99	98	97	96	96	98	98
	N2–H···N7	97	47	89	90	96	96	97
G5-G11	N7···H–N2	99	98	97	99	99	98	98
	O6···H–N1	98	84	91	94	95	98	97
G6-G10	N1–H···O6	100	100	64	99	99	98	99
	N2–H···N7	99	99	62	98	96	97	98
G10-G15	N1–H···O6	100	94	100	94	100	97	99
	N2–H···N7	100	61	95	95	100	98	99
G11-G14	N7···H–N2	100	100	99	99	100	99	99
	O6···H–N1	99	63	100	98	99	98	98

[a] Average values are calculated over the entire production MD simulation.

Table S2. Average distances (Å; standard deviation in parentheses) for K^+ coordination in G-quadruplexes in unbound TBA.[a]

	G8					G5	
	native	***anti* ThG**	***syn* ThG**	***anti* ^{CN}G**	***syn* ^{CN}G**	***syn* ^{CN}G**	***syn* StG**
K···G1(O6)	2.795 (0.180)	2.848 (0.215)	2.821 (0.205)	2.777 (0.169)	2.778 (0.175)	2.896 (0.244)	2.729 (0.184)
K···G2(O6)	2.721 (0.133)	2.761 (0.180)	2.745 (0.136)	2.749 (0.141)	2.787 (0.168)	2.743 (0.138)	2.788 (0.188)
K···G5(O6)	2.804 (0.174)	2.852 (0.222)	2.738 (0.147)	2.913 (0.362)	2.812 (0.234)	2.807 (0.184)	2.842 (0.179)
K···G6(O6)	2.747 (0.140)	2.733 (0.138)	2.802 (0.172)	2.791 (0.166)	2.838 (0.246)	2.780 (0.159)	2.821 (0.156)
K···G10(O6)	2.740 (0.142)	2.769 (0.157)	2.763 (0.151)	2.763 (0.170)	2.888 (0.290)	2.935 (0.272)	2.748 (0.178)
K···G11(O6)	2.776 (0.156)	2.766 (0.154)	2.758 (0.143)	2.858 (0.344)	2.730 (0.133)	2.731 (0.136)	2.839 (0.184)
K···G14(O6)	2.751 (0.141)	2.857 (0.277)	2.769 (0.175)	2.903 (0.342)	2.723 (0.167)	2.802 (0.177)	2.654 (0.168)
K···G15(O6)	2.824 (0.185)	2.685 (0.124)	2.773 (0.170)	2.834 (0.246)	2.837 (0.180)	2.859 (0.209)	2.809 (0.201)

[a] Average values are calculated over the entire production MD simulation.

Table S3. Key dihedral angles (°) and relative energies (kJ/mol) for adducted nucleotides.[a]

Adduct	χ[b]	θ[b]	ΔE[c]	ΔG[d]
anti Th**G**	–116.1	156.1	0.0	0.0
syn Th**G**	56.4	147.7	–5.0	3.0
anti ^{CN}**G**	–131.9	179.1	0.0	0.0
syn ^{CN}**G**	68.3	177.4	–11.3	–12.6

[a] Na^+ neutralized 5'-PO_4 methyl-capped nucleotide.
[b] Obtained at IEF-PCM-B3LYP/6-31G(d) level of theory (solvent: water).
[c] Relative energies calculated according to single-point energies obtained with IEF-PCM-B3LYP/6-311+G(2df,p), which include scaled (0.9806) zero-point vibrational energy corrections obtained from the optimization level of theory.
[d] Gibbs energies obtained at IEF-PCM-B3LYP/6-311+G(2df,p) including unscaled zero-point corrections obtained at optimization level.

Table S4. Average dihedral angles (°) of native and adducted dG (standard deviation in parentheses) over the entire simulation for free and bound TBA.[a]

		Free TBA		Bound TBA	
site	nucleobase	χ	θ	χ	θ
G8	**G**	196.6 (15.0)	–	200.6 (14.6)	–
	anti Th**G**	190.3 (8.4)	304.8 (117.6)	180.5 (18.6)	249.1 (157.1)
	syn Th**G**	33.6 (8.5)	298.7 (125.1)	31.6 (9.2)	304.8 (118.8)
	anti ^{CN}**G**	181.6 (13.7)	180.8 (6.6)	188.7 (9.5)	179.8 (6.5)
	syn ^{CN}**G**	52.7 (19.7)	182.1 (6.9)	51.7 (8.6)	183.0 (6.3)
G5	**G**	49.0 (8.9)	–	46.7 (7.1)	–
	syn ^{CN}**G**	63.4 (18.5)	177.6 (3.5)	66.3 (7.7)	165.7 (6.2)
	syn St**G**	76.8 (12.6)	170.7 (6.9)	107.1 (9.9)	173.2 (4.5)

[a] Average values are calculated over the entire production MD simulation.

Table S5. Hydrogen-bonding occupancies (%) of G-quadruplexes in TBA–protein complexes.[a]

		G8					G5	
Pair	**Interaction**	**native**	***anti* ThG**	***syn* ThG**	***anti* ^{CN}G**	***syn* ^{CN}G**	***syn* ^{CN}G**	***syn* StG**
G1-G6	N1–H···O6	99	99	98	100	99	98	99
	N2–H···N7	99	96	95	95	93	95	99
G1-G15	N7···H–N2	100	100	96	99	98	99	100
	O6···H–N1	98	99	96	97	97	92	98
G2-G5	N7···H–N2	100	100	99	100	99	99	99
	O6···H–N1	100	98	96	99	88	97	98
G2-G14	N1–H···O6	100	98	98	100	99	98	100
	N2–H···N7	98	98	96	97	97	99	98
G5-G11	N7···H–N2	100	100	98	100	100	99	99
	O6···H–N1	99	99	97	99	99	98	99
G6-G10	N1–H···O6	100	100	97	100	100	100	100
	N2–H···N7	100	71	92	77	99	92	97
G10-G15	N1–H···O6	100	100	100	100	99	100	100
	N2–H···N7	100	77	97	98	99	95	98
G11-G14	N7···H–N2	100	99	95	99	99	100	98
	O6···H–N1	99	92	75	92	94	99	99

[a] Average values are calculated over the entire production MD simulation.

Table S6. Average distances for K^+ (Å; standard deviation in parentheses) coordination in G-quadruplexes in TBA–protein complexes.[a]

	G8					G5	
	native	***anti* ThG**	***syn* ThG**	***anti* ^{CN}G**	***syn* ^{CN}G**	***syn* ^{CN}G**	***syn* StG**
K···G1(O6)	2.689 (0.113)	2.785 (0.163)	2.838 (0.228)	2.762 (0.142)	2.687 (0.112)	2.716 (0.128)	2.766 (0.147)
K···G2(O6)	2.874 (0.199)	2.749 (0.137)	2.713 (0.140)	2.743 (0.134)	2.864 (0.187)	2.833 (0.175)	2.838 (0.210)
K···G5(O6)	2.803 (0.160)	2.882 (0.204)	2.857 (0.206)	2.995 (0.254)	2.797 (0.156)	2.824 (0.159)	2.792 (0.167)
K···G6(O6)	2.747 (0.138)	2.729 (0.130)	2.748 (0.141)	2.737 (0.130)	2.819 (0.170)	2.764 (0.149)	2.747 (0.184)
K···G10(O6)	2.732 (0.133)	2.745 (0.137)	2.768 (0.164)	2.765 (0.144)	2.746 (0.137)	2.766 (0.148)	2.728 (0.164)
K···G11(O6)	2.768 (0.144)	2.798 (0.156)	2.721 (0.131)	2.721 (0.121)	2.756 (0.140)	2.748 (0.136)	2.659 (0.226)
K···G14(O6)	2.771 (0.146)	2.900 (0.217)	2.835 (0.192)	2.794 (0.165)	2.767 (0.149)	2.858 (0.205)	2.769 (0.218)
K···G15(O6)	2.754 (0.143)	2.742 (0.140)	2.772 (0.200)	2.803 (0.171)	2.752 (0.143)	2.768 (0.149)	2.750 (0.225)

[a] Average values are calculated over the entire production MD simulation.

Table S7. Linear interaction energies (LIE) between TBA and protein (kJ/mol; standard deviation in parentheses).[a]

	G8			G5	
Interaction[b]	**native**	***syn* ThG**	***syn* ^{CN}G**	***syn* ^{CN}G**	***syn* StG**
T3···Ile9	−12.1 (2.4)	−1.4 (2.0)	−5.8 (2.6)	−3.6 (3.4)	−2.5 (3.7)
T3···His66	−8.2 (6.6)	−2.2 (1.1)	−16.2 (8.3)	−16.3 (8.0)	−4.6 (2.4)
T3···Ile75	−5.8 (2.5)	−8.4 (2.7)	−7.4 (2.3)	−5.9 (2.4)	−3.9 (1.4)
T3···Tyr114	−5.6 (1.5)	−7.4 (2.7)	−11.3 (2.1)	−8.8 (4.7)	−13.0 (6.7)
T4···Arg70	−101.0 (9.8)	−31.5 (18.5)	−95.4 (10.3)	−92.7 (9.9)	−88.7 (8.5)
T4···Arg73	−115.2 (26.8)	−102.1 (12.8)	−129.2 (29.5)	−101.9 (10.6)	−108.4 (24.6)
G5···Asn74	−13.2 (11.4)	−0.6 (7.3)	−15.3 (23.9)	−1.1 (4.2)	−36.8 (18.2)
G5···Arg73	−240.8 (95.7)	−155.2 (33.9)	−271.3 (59.9)	−141.4 (20.9)	−276.0 (36.5)
T12···Tyr71	−13.7 (6.5)	−27.6 (4.3)	−25.7 (4.6)	−20.8 (8.1)	−45.8 (8.7)
T13···Tyr71	−10.3 (2.4)	−23.3 (2.1)	−24.7 (2.2)	−31.4 (5.2)	−32.1 (4.2)
T13···Arg62	−23.5 (22.6)	−7.6 (17.2)	−36.7 (16.0)	−15.9 (10.4)	−22.9 (5.4)
T13···Arg70	−94.4 (12.3)	−72.4 (15.8)	−88.3 (13.9)	−83.3 (12.9)	−92.0 (14.5)
T13···Arg73	−37.7 (4.9)	−36.2 (5.1)	−34.8 (4.9)	−33.9 (3.7)	−34.3 (5.1)
G14···Arg70	−109.2 (12.4)	−59.5 (18.7)	−102.7 (17.6)	−100.3 (15.0)	−100.4 (16.8)

[a] Average values are calculated over the entire production MD simulation.
[b] For hydrogen-bonding interactions the electrostatic component and for stacking interactions van der Waals component of LIE is reported.

Table S8. Linear interaction energies (LIE) between T residues.

Interaction	**Native**	***syn* ^{CN}G**	***syn* StG**
T3···T4	−2.3 (2.0)	−3.9 (2.9)	−4.8 (1.0)
T3···T12	−1.1 (0.8)	−5.7 (0.9)	−0.9 (1.2)
T4···T12	−0.8 (1.2)	−2.8 (1.3)	−2.5 (1.1)
T4···T13	−0.3 (3.7)	−5.9 (1.9)	−3.0 (0.6)
T12···T13	−1.3 (1.3)	−2.5 (1.7)	−3.6 (1.2)
T7···G8	−1.0 (0.4)	−6.2 (1.8)	−10.5 (0.9)

Table S9. The average root-mean-square deviation (rmsd, Å; standard deviations in parentheses) for considered systems over the entire MD simulation with respect to the first frame of production phase.[a]

		TBA–Thrombin Complex		
System	**Free TBA**	**Bound TBA**	**Enzyme**	**Overall**
native	1.598 (0.210)	1.254 (0.147)	1.507 (0.134)	1.849 (0.163)
anti ThG	2.826 (0.494)	–	–	–
syn ThG	2.479 (0.467)	2.202 (0.410)	1.534 (0.145)	2.005 (0.263)
anti ^{CN}G	2.451 (0.645)	–	–	–
syn ^{CN}G	2.101 (0.384)	1.634 (0.222)	1.684 (0.235)	1.918 (0.225)
5 ^{CN}G	2.919 (0.618)	1.697 (0.315)	1.806 (0.254)	1.784 (0.334)
5 StG	2.516 (0.809)	1.332 (0.178)	1.709 (0.202)	1.656 (0.295)

[a] RMSD calculated according to the position of the heavy atoms in the backbone of TBA (P, O, C) and protein (O, N, C).

Atom types and partial charges for the ^{CN}dG residue

1	P	4.7160	4.0030	0.3930	P	1 CNA 1.209100
2	O5'	3.2940	3.2820	0.0000	OS	1 CNA -0.486800
3	O1P	5.8000	3.0220	0.2340	O2	1 CNA -0.786000
4	O2P	4.6890	5.3340	-0.2320	O2	1 CNA -0.795800
5	O3'	0.0000	0.0000	0.0000	OS	1 CNA -0.549300
6	C5'	3.1140	1.9140	0.3320	CT	1 CNA 0.055100
7	H5'	2.9640	1.7850	1.4190	H1	1 CNA 0.057500
8	H5"	3.9720	1.3010	0.0210	H1	1 CNA 0.057500
9	C4'	1.8680	1.4240	-0.3950	CT	1 CNA 0.101600
10	H4'	1.0470	2.1210	-0.1950	H1	1 CNA 0.123700
11	O4'	2.1150	1.3970	-1.8120	OS	1 CNA -0.383900
12	C1'	1.8740	0.0950	-2.3250	CT	1 CNA 0.129300
13	H1'	0.8570	0.0370	-2.7250	H2	1 CNA 0.115600
14	C2'	2.0040	-0.8580	-1.1300	CT	1 CNA -0.071000
15	H2'	3.0550	-1.0880	-0.9450	HC	1 CNA 0.045600
16	H2"	1.4550	-1.7930	-1.2790	HC	1 CNA 0.045600
17	C3'	1.4270	0.0000	0.0000	CT	1 CNA 0.188800
18	H3'	1.8140	-0.2910	0.9870	H1	1 CNA 0.049600
19	N	6.6890	-0.7800	-3.7460	NA	1 CNA -0.461200
20	H	7.6680	-1.0140	-3.8660	H	1 CNA 0.348200
21	N1	7.0530	-0.7280	-1.4340	N2	1 CNA -0.872600
22	H6	6.6590	-0.4160	-0.5560	H	1 CNA 0.394300
23	H7	8.0020	-0.4100	-1.5850	H	1 CNA 0.394300
24	N2	4.9030	-0.3790	-2.2490	NC	1 CNA -0.519000
25	N3	3.5040	-0.4080	-5.5640	NB	1 CNA -0.488500
26	N4	2.7740	-0.1640	-3.4370	N*	1 CNA -0.051500
27	C	6.1730	-0.6040	-2.4890	CA	1 CNA 0.617300
28	C6	4.1380	-0.3540	-3.3730	CB	1 CNA 0.193500
29	C7	4.5550	-0.5110	-4.6970	CB	1 CNA 0.093300
30	C8	5.9490	-0.7560	-4.9770	C	1 CNA 0.504000
31	C9	2.4450	-0.1860	-4.8070	CK	1 CNA 0.306700
32	O5	6.5330	-0.9400	-6.0300	O	1 CNA -0.539100
33	C10	1.1040	0.0070	-5.3190	CD	1 CNA -0.268500
34	H8	0.3280	0.2800	-4.6110	HA	1 CNA 0.185600
35	C11	0.8110	-0.1240	-6.6330	CD	1 CNA -0.063500
36	H9	1.6290	-0.4010	-7.2950	HA	1 CNA 0.131800
37	C12	-0.4910	0.0620	-7.2630	CA	1 CNA 0.066600
38	C13	-0.6130	-0.1940	-8.6440	CA	1 CNA -0.138800
39	H10	0.2650	-0.5150	-9.1990	HA	1 CNA 0.147200
40	C14	-1.8230	-0.0470	-9.3070	CA	1 CNA -0.165300
41	H11	-1.8970	-0.2510	-10.3710	HA	1 CNA 0.170800
42	C15	-1.6460	0.4840	-6.5690	CA	1 CNA -0.131200
43	H12	-1.5920	0.7040	-5.5070	HA	1 CNA 0.132300
44	C16	-2.8600	0.6340	-7.2190	CA	1 CNA -0.168200
45	H13	-3.7380	0.9590	-6.6710	HA	1 CNA 0.176000
46	C17	-2.9630	0.3680	-8.5980	CA	1 CNA 0.028300
47	C18	-4.2180	0.5220	-9.2690	C1	1 CNA 0.332200
48	N5	-5.2400	0.6450	-9.8130	N1	1 CNA -0.461000

Atom types and partial charges for the ThdG residue

1	P	4.1840	2.9240	-1.6150	P	1 THA 1.209200
2	O5'	3.8980	2.8480	-0.0000	OS	1 THA -0.479400
3	O1P	3.0690	2.2770	-2.3210	O2	1 THA -0.784800
4	O2P	4.6710	4.2850	-1.8830	O2	1 THA -0.797200

5	O3'	0.0000	0.0000	0.0000	OS	1 THA	-0.555800
6	C5'	3.2610	1.6970	0.5190	CT	1 THA	0.046000
7	H5"	3.1150	1.8530	1.5940	H1	1 THA	0.055700
8	H5"	3.8990	0.8090	0.3900	H1	1 THA	0.055700
9	C4'	1.8920	1.4530	-0.1460	CT	1 THA	0.125100
10	H4'	1.1430	2.1440	0.2510	H1	1 THA	0.119300
11	O4'	1.9740	1.6980	-1.5680	OS	1 THA	-0.397200
12	C1'	1.8310	0.4910	-2.2990	CT	1 THA	0.120100
13	H1'	0.8380	0.4670	-2.7570	H2	1 THA	0.108500
14	C2'	2.0000	-0.6410	-1.2710	CT	1 THA	-0.070600
15	H2'	3.0630	-0.8720	-1.1460	HC	1 THA	0.040400
16	H2"	1.4820	-1.5610	-1.5580	HC	1 THA	0.040400
17	C3'	1.4240	0.0000	0.0000	CT	1 THA	0.195200
18	H3'	1.8180	-0.4620	0.9170	H1	1 THA	0.043900
19	N	6.5550	1.6950	-3.7680	NA	1 THA	-0.406500
20	H	7.5260	1.9530	-3.8990	H	1 THA	0.344500
21	C	5.9680	1.8150	-2.5390	CA	1 THA	0.506900
22	N1	6.7110	2.3200	-1.5020	N2	1 THA	-0.796800
23	H6	6.1420	2.5660	-0.6940	H	1 THA	0.371900
24	H7	7.4320	2.9940	-1.7260	H	1 THA	0.371900
25	N2	4.7320	1.4320	-2.2970	NC	1 THA	-0.455200
26	C6	4.0880	0.9220	-3.3770	CB	1 THA	0.109800
27	C7	4.5750	0.7860	-4.6740	CB	1 THA	0.160400
28	C8	5.9310	1.1830	-4.9630	C	1 THA	0.464000
29	O5	6.5730	1.1450	-5.9970	O	1 THA	-0.536400
30	N3	3.6190	0.2530	-5.5000	NB	1 THA	-0.465900
31	C9	2.5600	0.0650	-4.7360	CK	1 THA	0.200700
32	N4	2.7890	0.4610	-3.4060	N*	1 THA	0.052500
33	C10	1.3150	-0.4780	-5.2520	C*	1 THA	-0.049700
34	C11	0.2910	-1.1820	-4.6520	C*	1 THA	0.001200
35	H8	0.2840	-1.4630	-3.6060	HA	1 THA	0.118400
36	C12	-0.7410	-1.5600	-5.5580	C*	1 THA	-0.300500
37	H9	-1.6210	-2.1220	-5.2650	HA	1 THA	0.184500
38	S	1.0060	-0.2920	-6.9670	S	1 THA	0.008800
39	C13	-0.4920	-1.1470	-6.8390	CA	1 THA	-0.174800
40	H10	-1.1010	-1.3050	-7.7190	H4	1 THA	0.215800

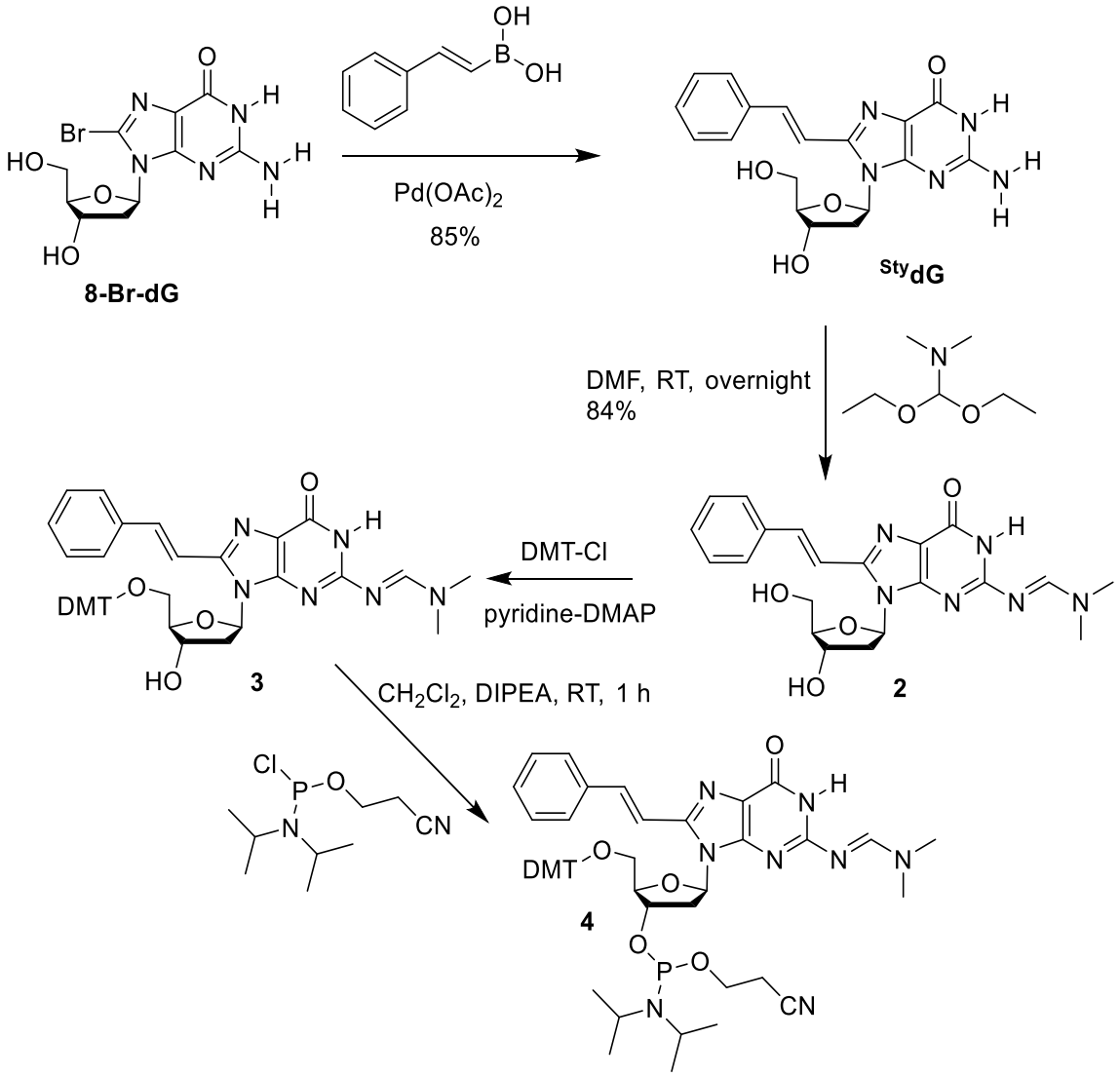
Atom types and partial charges for the ^{Sty}dG residue

1	P	5.5960	1.6510	0.1630	P	STY	1.2095
2	O	4.0690	1.0700	0.0000	OS	STY	-0.4968
3	O1	5.5200	3.0560	0.5900	O2	STY	-0.7864
4	O2	6.3650	0.6040	0.8530	O2	STY	-0.7947
5	O3	0.0000	0.0000	0.0000	OS	STY	-0.551
6	C5*	3.0580	1.9430	-0.4880	CT	STY	-0.0098
7	H5*1	3.4180	2.5450	-1.3350	H1	STY	0.0741
8	H5*2	2.7160	2.6330	0.3000	H1	STY	0.0741
9	C4*	1.8680	1.1240	-0.9650	CT	STY	0.1929
10	H4*	1.0170	1.8000	-1.1190	H1	STY	0.0951
11	O4*	2.1980	0.4860	-2.2110	OS	STY	-0.4113
12	C1*	1.8610	-0.9000	-2.1340	CT	STY	0.1456
13	H1*	0.8130	-1.0320	-2.4290	H2	STY	0.068
14	C3*	1.4300	0.0000	0.0000	CT	STY	0.1433
15	H3*	1.8210	0.1570	1.0130	H1	STY	0.0737
16	C2*	2.0200	-1.2530	-0.6550	CT	STY	-0.0399
17	H2*1	3.0750	-1.3510	-0.4030	HC	STY	0.0392

18	H2*2	1.4950	-2.1730	-0.3770	HC	STY	0.0392
19	N	5.7990	-4.0730	-3.1170	NA	STY	-0.4271
20	H	6.5420	-4.7620	-3.1590	H	STY	0.3381
21	N1	6.0920	-3.9910	-0.7940	N2	STY	-0.9012
22	H6	5.8670	-3.4450	0.0290	H	STY	0.4017
23	H7	7.0860	-4.1480	-0.9060	H	STY	0.4017
24	N2	4.3950	-2.7350	-1.7680	NC	STY	-0.441
25	N3	3.2280	-2.3900	-5.1570	NB	STY	-0.4846
26	N4	2.6550	-1.6330	-3.1030	N*	STY	0.0239
27	C	5.4030	-3.5590	-1.9110	CA	STY	0.5751
28	C6	3.7600	-2.4440	-2.9350	CB	STY	0.1236
29	C7	4.0790	-2.9030	-4.2160	CB	STY	0.0964
30	C8	5.1880	-3.8080	-4.3890	C	STY	0.4908
31	C9	2.3860	-1.6290	-4.4850	CK	STY	0.2815
32	O5	5.6340	-4.3440	-5.3900	O	STY	-0.548
33	C10	1.3090	-0.8500	-5.0660	CD	STY	-0.2831
34	H8	0.8360	-0.1140	-4.4250	HA	STY	0.1843
35	C11	0.9350	-0.9910	-6.3560	CD	STY	-0.0507
36	H9	1.4660	-1.7360	-6.9450	HA	STY	0.1226
37	C12	-0.1100	-0.2480	-7.0590	CA	STY	0.0335
38	C13	-0.3670	-0.5600	-8.4080	CA	STY	-0.097
39	H10	0.2210	-1.3390	-8.8880	HA	STY	0.1329
40	C14	-1.3530	0.1080	-9.1300	CA	STY	-0.1971
41	H11	-1.5300	-0.1530	-10.1700	HA	STY	0.1483
42	C15	-0.8830	0.7670	-6.4600	CA	STY	-0.1291
43	H12	-0.7110	1.0370	-5.4220	HA	STY	0.133
44	C16	-1.8690	1.4330	-7.1800	CA	STY	-0.1657
45	H13	-2.4530	2.2130	-6.6970	HA	STY	0.1392
46	C17	-2.1100	1.1090	-8.5190	CA	STY	-0.0988
47	H14	-2.8800	1.6330	-9.0790	HA	STY	0.132

Synthesis of ^{Sty}dG-Phosphoramidite

Scheme S1. Synthesis of ^{Sty}dG-phosphoramidite (**4**).



1) ^{Sty}dG

A mixture of 8-Br-dG (1 g, 2.88 mmol), (*E*)-2-phenylvinylboronic acid (0.55 g, 3.75 mmol), Na_2CO_3 (0.61 g, 5.75 mmol), and $Pd(OAc)_2$ (19 mg, 0.084 mmol) were added into a dry 50- ML RBF along with a Teflon-coated stir bar. The reaction flask was evacuated for 5 minutes then refilled with dry nitrogen gas. A degassed mixture (30 mL) of H_2O/acetonitrile (2/1) was added to the flask via a syringe. The mixture was heated overnight at 80°C under nitrogen atmosphere, then cooled to room temperature, poured in to 100 mL of H_2O and neutralized with 10% HCl. The precipitate formed was filtrated and dried *in vacuo* to yield ^{Sty}dG as a yellow powder (0.9 g, 85% yield). ^{1}H NMR (300 MHz, DMSO-d_6) δ : 10.65 (s, 1H), 7.69 (d, J = 7.3, 2H) 7.56- 7..23 (m, 5H) 6.47 (bs, 2H), 6.37 (dd, *J* = 8.6, 6.4, 1H), 5.27 (d, *J* = 4.2, 1H), 5.17 (t, *J* = 5.1,

1H), 4.45 (m, 1H), 3.82 (m, 1H), 3.74- 3.6 (m, 2H), 2.60 (m, 1H), 2.09 (m, 1H). HRMS calculated for $C_{18}H_{20}N_5O_4^+$ [M+H$^+$] 370.1515; found 370.1500.

2) *N*^{2}-(dimethylformamidyl)-(*E*)-8-styryl-2′-deoxyguanosine (2)

^{Sty}dG (0.85 g, 2.3 mmol) was placed in a 50-mL RBF and reverse filled with argon. Dry DMF (15 mL) was added, followed by dimethylformamide diethyl acetal (1.3 mL, 9.2 mmol). The mixture was stirred overnight at room temperature. The solvent was then evaporated under reduced pressure. The residue was triturated with a mixture of hexane/methanol; 1:1. The isolated solid was filtered, washed with ethyl acetate to remove the residual DMF. The solid product was dried under vacuum to yield compound **2** (0.75 g, 82.4 %) which was pure enough to be used for the next step. ^{1}HNMR (300 MHz, DMSO-*d6*) δ = 11.36 (bs, 1H), 8.55 (s, 1H), 7.7 (d, *J* = 7.3, 2H), 7.58 (d, *J*= 15.8 Hz, 1H), 7.48 (d, *J*= 15.8 Hz, 1H), 7.4- 7.28 (m, 3H), 6.49 (t, *J* = 7.4 Hz, 1H), 5.33 (d, *J* = 4.3 Hz, 1H), 5.11 (t, *J* = 5.4 Hz, 1H), 4.50 (m, 1H), 3.81 (q, *J* = 4.4 Hz,1H), 3.75- 3.61 (m, 2H), 3.15 (s, 3H), 3.03 (s, 3H), 2.76- 2.66 (m, 1H), 2.18-2.12 (m, 1H). HRMS calculated for $C_{21}H_{25}N_6O_4^+$ [M+H$^+$] 425.1937; found 425.1926

3) 5′-*O*-(4,4′-dimethoxytrityl)-*N*^{2}-(dimethylformamidyl)-(*E*)-8-styryl-2′-deoxyguanosine (3)

Compound **2** (0.41 g, 0.96 mmol) was coevaporated from dry pyridine (3 × 5 mL) in a 25-mL RBF. DMAP (0.023 g, 0.188 mmol) was added to the reaction flask. The RBF was then fitted with a dropping funnel and reverse filled with nitrogen gas. Dry pyridine (10 mL) was added to the RBF. DMT-Cl (0.4 g, 1.18 mmol) was placed in the dropping funnel and dissolved in 10 mL of dry pyridine. The DMT-Cl/pyridine solution was then allowed to add dropwise over 30 min. The reaction mixture was allowed to stir at room temperature under nitrogen atmosphere overnight. Upon completion, the mixture was diluted with DCM (25 mL) and washed with water (2 × 10 mL). The organic layer was dried over anhydrous sodium sulfate. The crude product was then triturated with a mixture of ethyl acetate/hexanes (1/2) to remove the DMAP and the unreacted DMT-Cl. The residue was dried under vacuum and was used for the next step without further purification. HRMS calculated for $C_{42}H_{43}N_6O_6^+$ [M + H$^+$] 727.3244, found 727.3226.

4) 5′-*O*-(4,4′-dimethoxytrityl)-3′-*O*-[2-cyanoethoxy-(*N*,N-diisopropylamino)phosphino]-*N*^{2}-(dimethylformamidyl)-8-styryl-2′-deoxyguanosine (4).

Compound **3** (0.29 g, 0.399 mmol) was introduced into a 25-mL RBF connected to a nitrogen line to remove any air from the flask. A dry 4 ^{o}A molecular sieves (around 2 g) was added under nitrogen

atmosphere followed by dry methylene chloride (10 mL) using a syringe. The reaction mixture was stirred at room temperature for 20 minutes. Diisopropyl ethyl amine (DIPEA), (0.2 mL, 1.19 mmol) and dimethylaminopyridine (0.2 equiv.) were added to the reaction mixture via a syringe then the reaction mixture was stirred for 5 minutes at room temperature. Using a 1-mL syringe, 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (0.14 mL, 0.6 mmol) was added to the reaction mixture. The reaction mixture was stirred at room temperature for 1.5 h, where TLC (EtOAc: MeOH: TEA; 95: 3: 2) indicated the disappearance of the starting materials and the formation of less-polar product (two spots near to each other). The reaction mixture was filtered to remove the molecular sieves. The solvent was removed under vacuum. The residue was purified by flash column chromatography using (EtOAc: MeOH: TEA; 95: 3: 2) as eluent. The product **4** was obtained as pale yellow oil (0.19 g, 52.3%). ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.47 (s, 1H), 7.88 (d, J = 15.6 Hz, 1H), 7.42-7.30 (m, 4H), 7.30-7.10 (m, 10H), 6.70 (d, J = 8.7 Hz, 4H), 6.48-6.41 (m, 1H), 6.41 (t, J = 6.6 Hz), 4.91-4.78 (m, 1H), 4.21-4.12 (m, 1H), 3.68 (s, 6H), 3.61-3.43 (m, 5H), 3.41-3.29 (m, 2H), 3.05 (s, 6H), 2.42-2.28 (m, 3H), 1.22-1.01 (m, 12H).^{31}PNMR (162 MHz, $CDCl_3$, referenced to H_3PO_4): δ 149.97. HRMS calculated for $C_{51}H_{61}N_8O_7P^+$ [M+H$^+$]: 927.4323; found 927.4251.

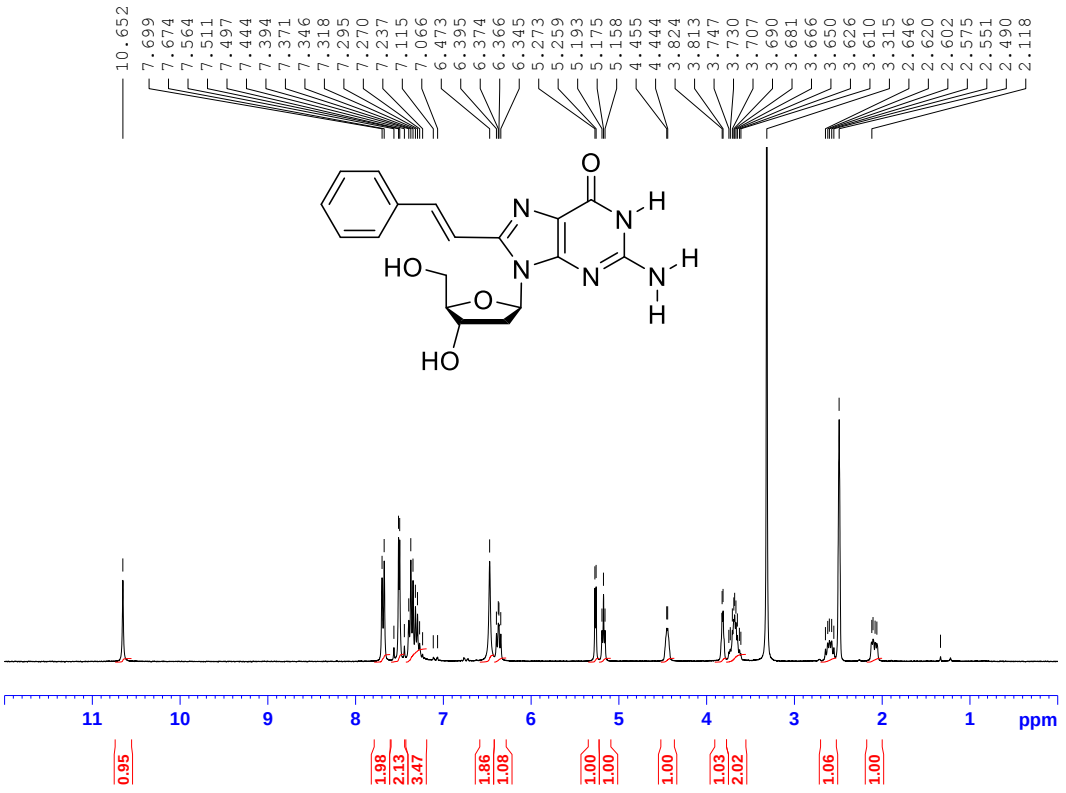


Figure S16. ^{1}HNMR (DMSO-*d6*) of ^{Sty}dG.

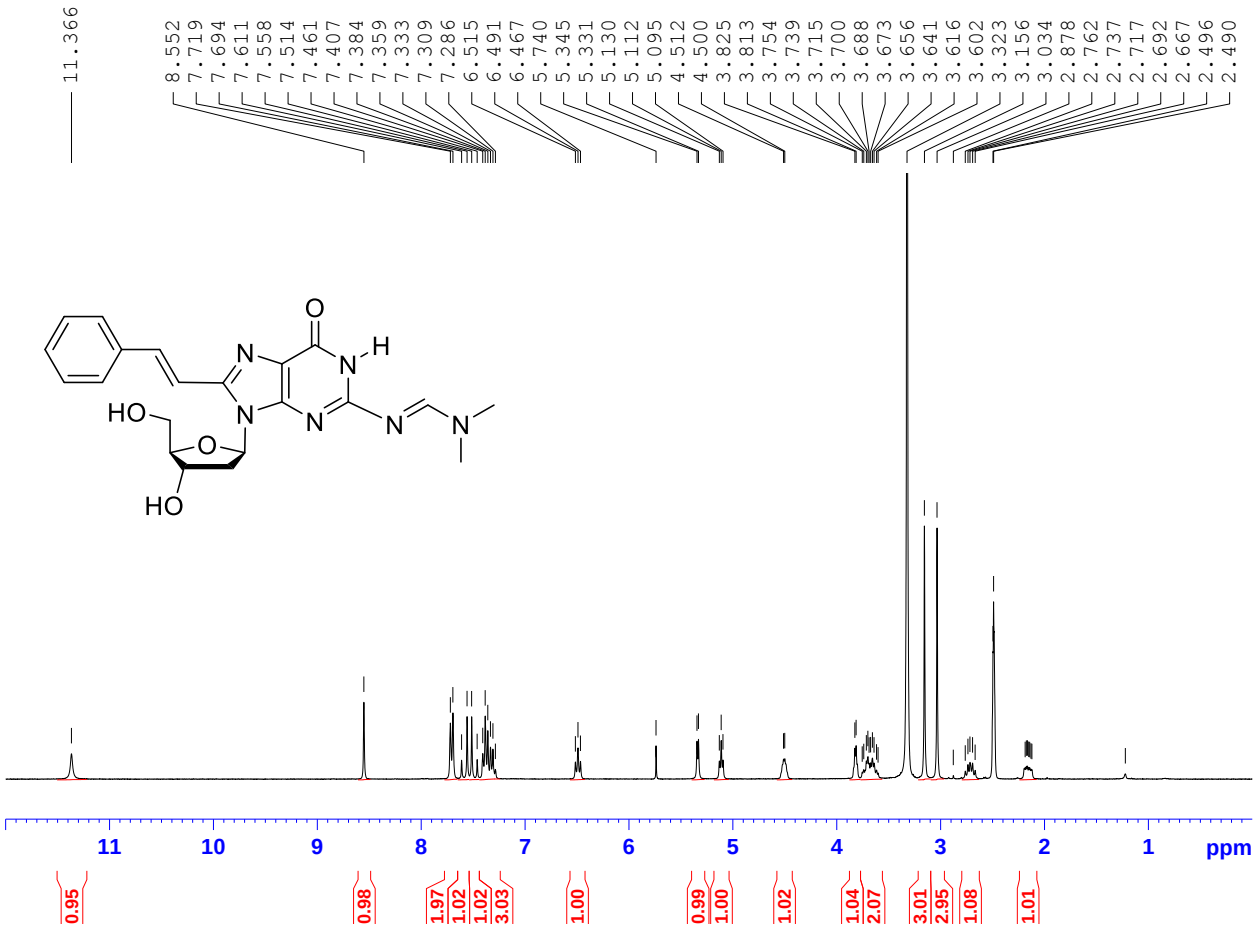


Figure S17. ^{1}HNMR (DMSO-*d6*) of **2**.

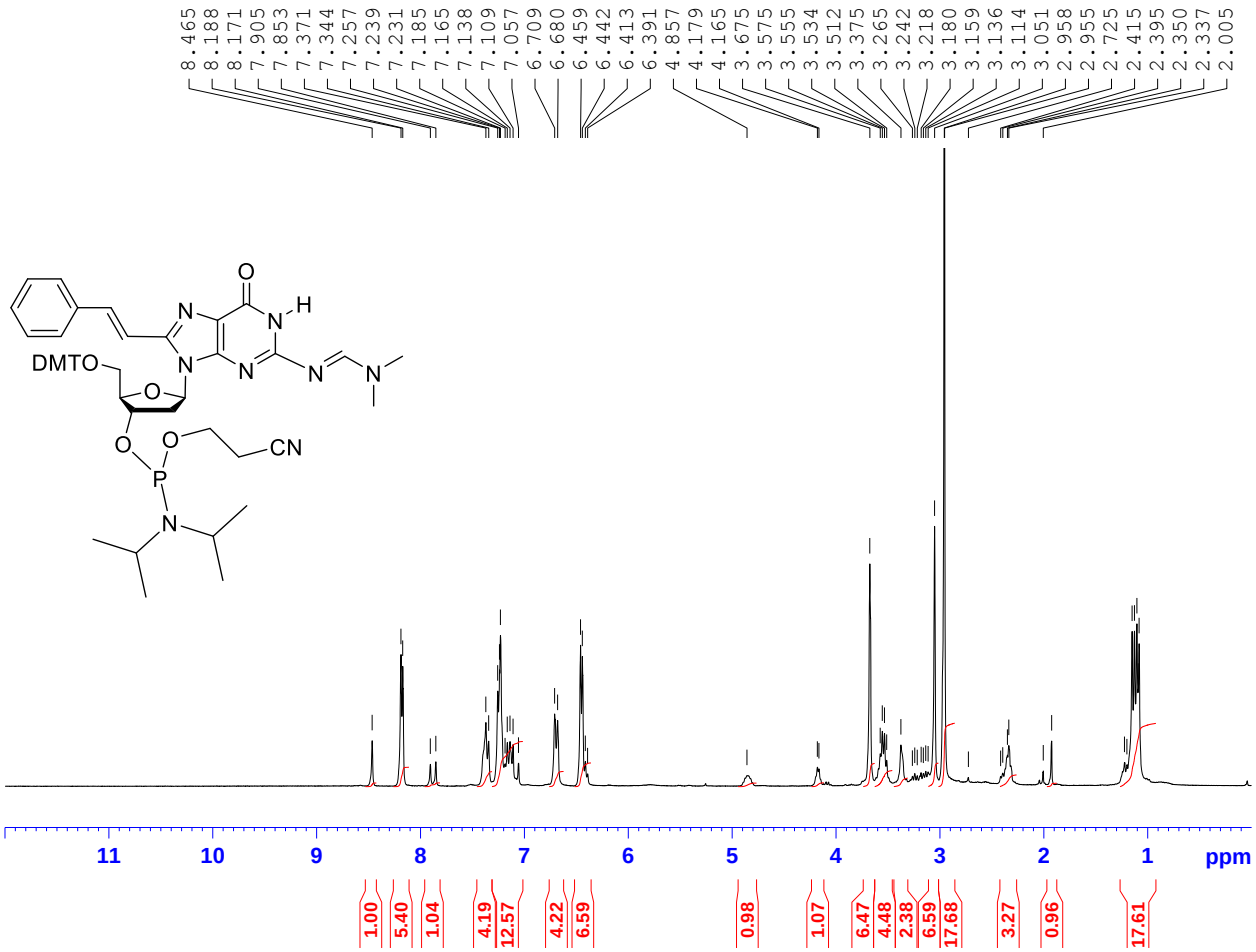


Figure S18. ^{1}HNMR ($CDCl_3$) of **4**.

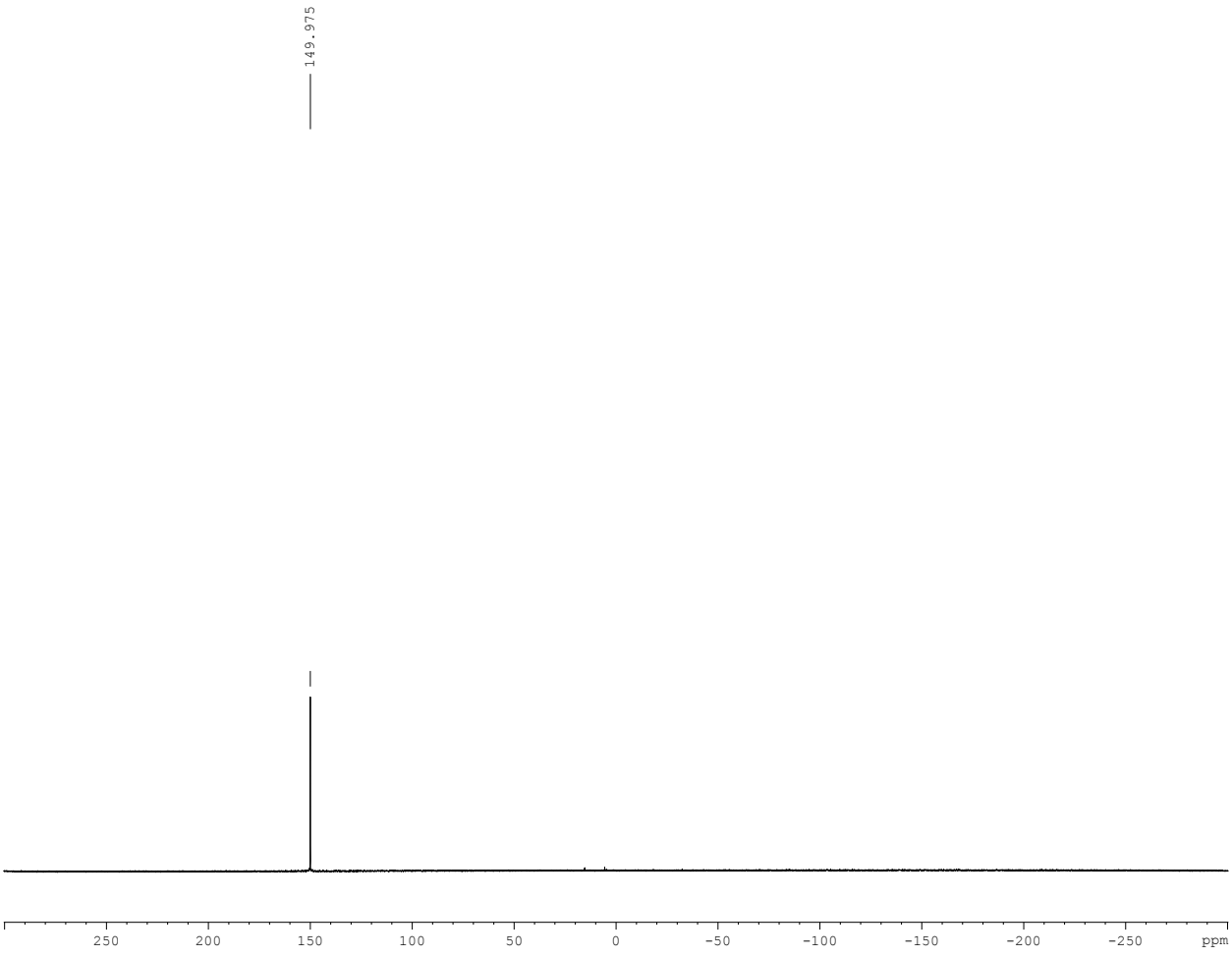
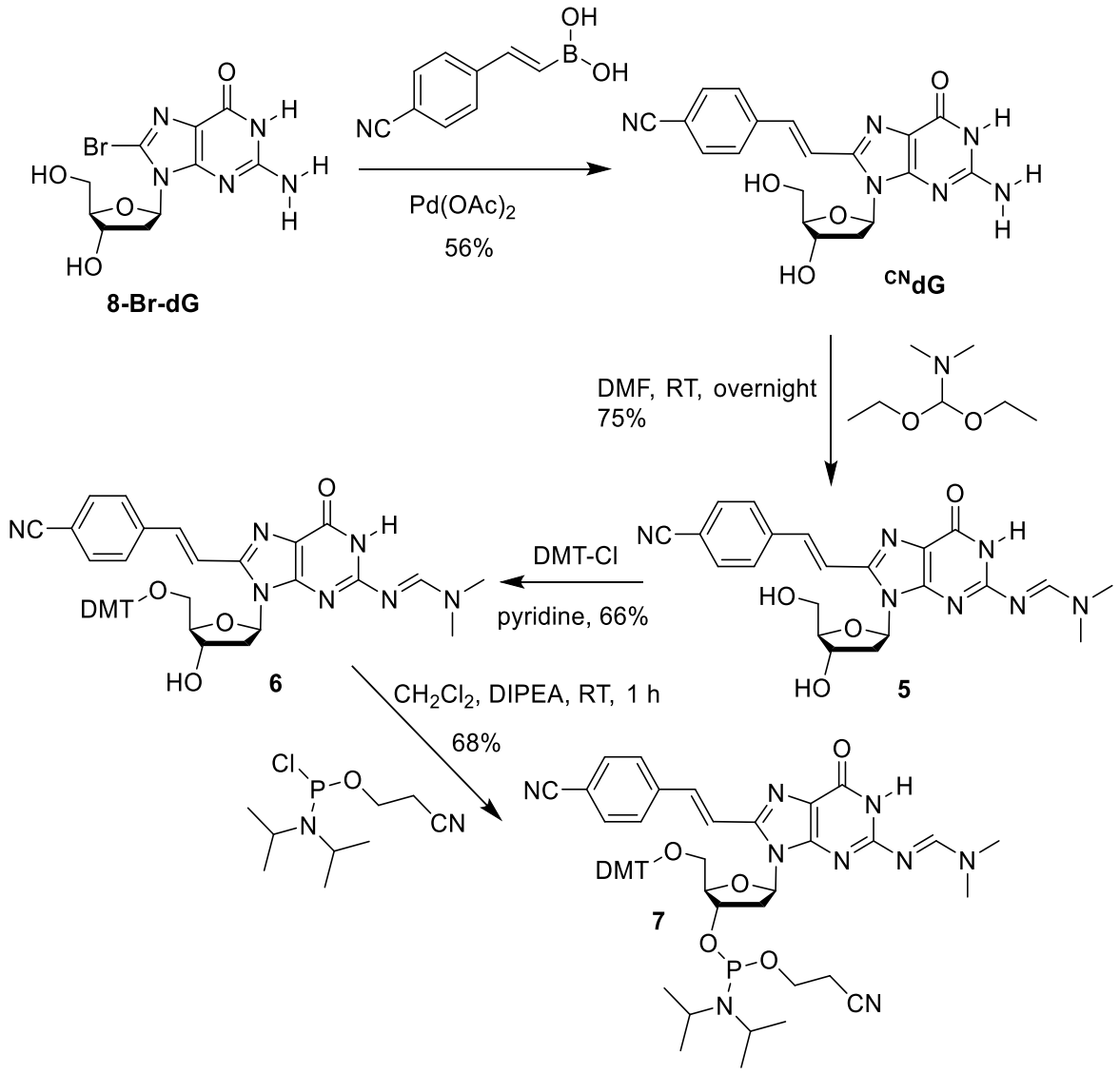


Figure S19. ^{31}P-NMR ($CDCl_3$) of compound (**4**)

Synthesis of ^{CN}dG-Phosphoramidite

Scheme S2. Synthesis of ^{CN}dG-phosphoramidite (**7**).



1) ^{CN}dG)

Synthesis was performed as described for ^{Sty}dG. (Yield 56%)^{1}H NMR (300 MHz, DMSO-d_6) δ: 10.74 (bs, 1H), 7.89 (d, J = 8.4, 2H), 7.8 -7.6 (m, 4H), 6.54 (bs, 2H), 6.45 (m, 1H), 5.31 (d, J = 4.1, 1H), 5.26 (t, J = 10.4, 1H), 4.60 (m, 1H), 3.83 (m, 1H), 3.73- 3.63 (m, 2H), 2.57 (m, 1H), 2.08 (m, 1H).; ^{13}C NMR (400 MHz, DMSO-d_6) δ = 156.9, 154.0, 152.3, 144.1, 141.5, 133.2, 133.1, 128.3, 119.9, 119.5, 117.4, 110.5, 87.8, 83.0, 71.0, 61.6. HRMS calculated for $C_{19}H_{19}N_6O_4^+$ [M+H$^+$] 395.1468; found 395.1438.

2) N^2-(dimethylformamidyl)-8-(4''-cyanostyrene)-2'-deoxyguanosine (5)

Synthesis was performed as described for **2**. (Yield 75%) ^{1}HNMR (300 MHz, DMSO-d_6) δ = 11.37 (bs, 1H), 8.55 (s, 1H), 7.7 (d, J = 7.3, 2H), 7.58 (d, J= 7.8 Hz, 1H), 7.60 - 7.28 (m, 2H), 6.47 (t, J = 8.5 Hz, 1H), 5.33 (d, J = 3.3 Hz, 1H), 5.14 (m, 1H), 4.50 (m, 1H), 3.79 (m,1H), 3.82- 3.61 (m, 2H), 3.15 (s, 3H), 3.03 (s, 3H), 2.73-2.61 (m, 1H), 2.19-2.11 (m, 1H).; ^{13}C NMR 158.60, 157.78, 157.6, 145.12, 141.6, 133.07, 132.06, 128.40, 120.49, 119.65, 119.47, 110.62, 87.72, 82.94, 70.48, 61.47, 41.24, 35.17. HRMS calculated for $C_{22}H_{24}N_7O_4^+$ [M+H$^+$] 450.1890; found 450.1852.

3) 5'-*O*-(4,4'-dimethoxytrityl)-N^2-(dimethylformamidyl)-8-(4''-cyanostyrene)-2'-dG (6)

Synthesis performed as outlined for [3]. (66% yield). ^{1}H NMR (300 MHz, DMSO-d_6) δ = 8.46 (s, 1H), 7.82 (d, J=8.5 Hz, 1H), 7.38 (d, J= 8.5 Hz, 6H), 7.39 (m, 4H), 7.28 (m, 4H), 7.18 (m, 4H), 6.63 (m, 4H), 6.43 (m, 1H), 4.65 (bs, 1H), 4.06 (m, 1H), 3.69 (s, 3H), 3.46 (s, 3H), 3.38 (m, 1H), 3.41 (m, 1H), 3.08 (m, 1H), 3.01 (s, 3H), 3.04 (s, 3H), 2.40 (m, 1H). ^{13}C NMR (400) MHz, $CDCl_3$) δ= 158.3, 158.2, 157.9, 157.1, 150.6, 145.4, 141.2, 136.0, 133.1, 131.9, 130.0, 128.2, 127.0, 121.0, 119.4, 119.0, 113.4, 110.6, 85.7, 83.2, 70.7, 63.0, 55.3, 52.4, 45.9, 41.3, 35.2. HRMS calculated for $C_{43}H_{42}N_7O_6^+$ [M + H$^+$] 752.3197, found 752.3132.

4) 5'-*O*-(4,4'-dimethoxytrityl)-3'-*O*-[(2-cyanoethoxy)(diisopropylamino)phosphino]-N^2-(dimethylformamidyl)-8-(4''-cyanostrene)-2'-dG (7)

Synthesis was performed as outlined for **4**. (68% yield based on 31P NMR, 32% P^V). ^{1}H NMR (300 MHz, $CDCl_3$) δ = δ: 8.73 (br s, 1H), 8.45 (s, 1H), 7.82 (d, J = 16 Hz, 1H), 7.49-7.30 (m, 4H), 7.24-7.09 (m, 9H), 6.69-6.63 (m, 5H), 6.46-6.41 (m, 1H), 5.59 (s, 1H), 4.82-4.63 (m, 1H), 4.17-4.09 (m, 1H), 3.66 (s, 6H), 3.60-3.42 (m, 4H), 3.40-3.26 (m, 2H), 3.05 (s, 6H), 2.43-2.27 (m, 3H), 1.31-1.01 (m, 12H). ^{13}C NMR (75

MHz, $CDCl_3$) δ= 158.4, 157.3, 156.1, 154.9, 150.7, 150.4, 145.5, 140.5, 144.5, 140.5, 138.3, 135.6, 132.3, 129.9, 128.2, 128.1, 128.0, 127.7, 127.4, 113.0, 86.2, 82.9, 78.5, 77.2, 63.4, 55.1, 53.0, 46.9, 45.2, 43.3, 41.3, 35.2, 8.9. ^{31}P NMR (162 MHz, $CDCl_3$) δ = 149.1, 14.3 (P^V). HRMS calculated for $C_{52}H_{59}N_9O_7P^+$ [M+H$^+$]: 952.4275; found 952.4186.

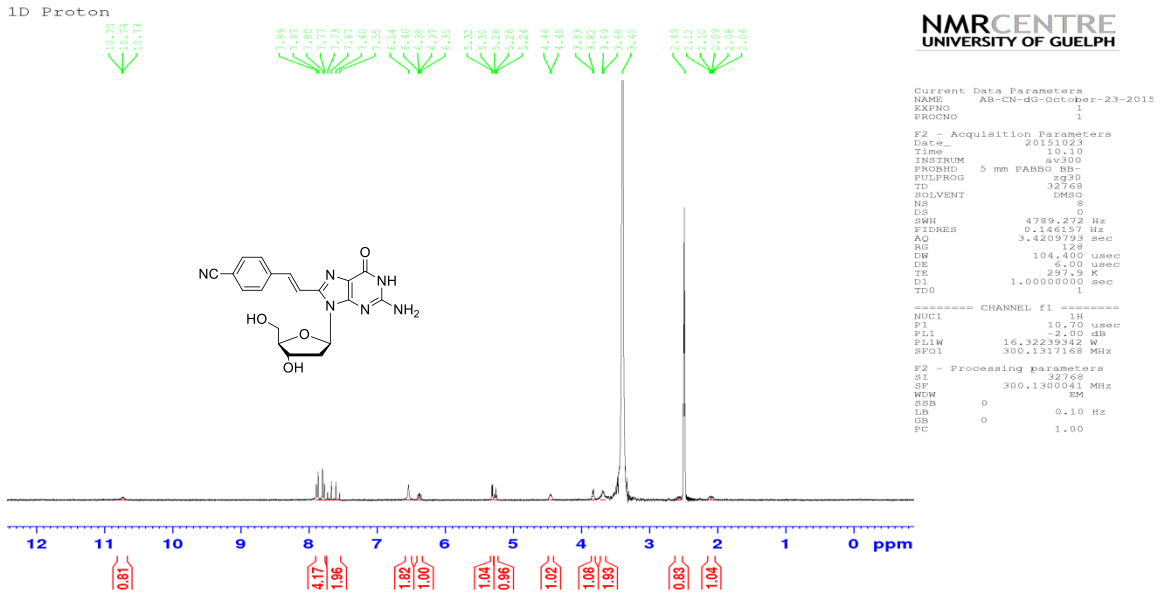


Figure S20. ^{1}HNMR (DMSO-*d6*) of ^{CN}dG.

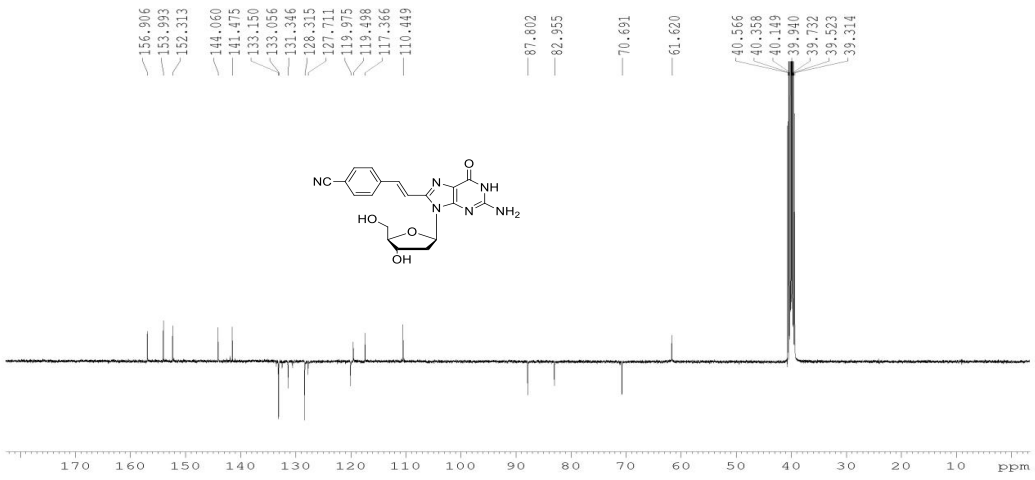


Figure S21. ^{13}CNMR (DMSO-*d6*) of ^{CN}dG.

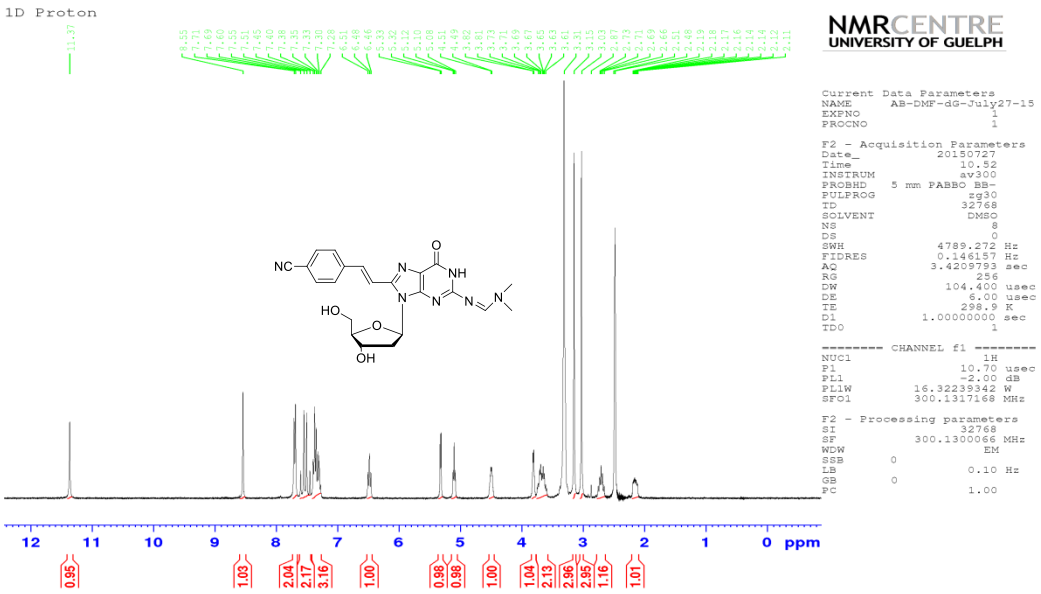


Figure S22. [1]HNMR (DMSO-*d6*) of **5**.

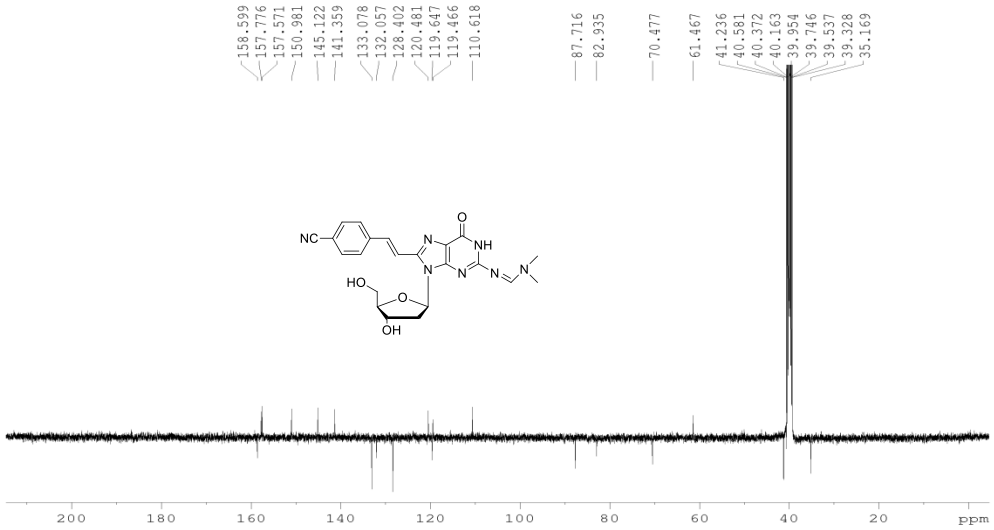


Figure S23. [13]CNMR (DMSO-*d6*) of **5**.

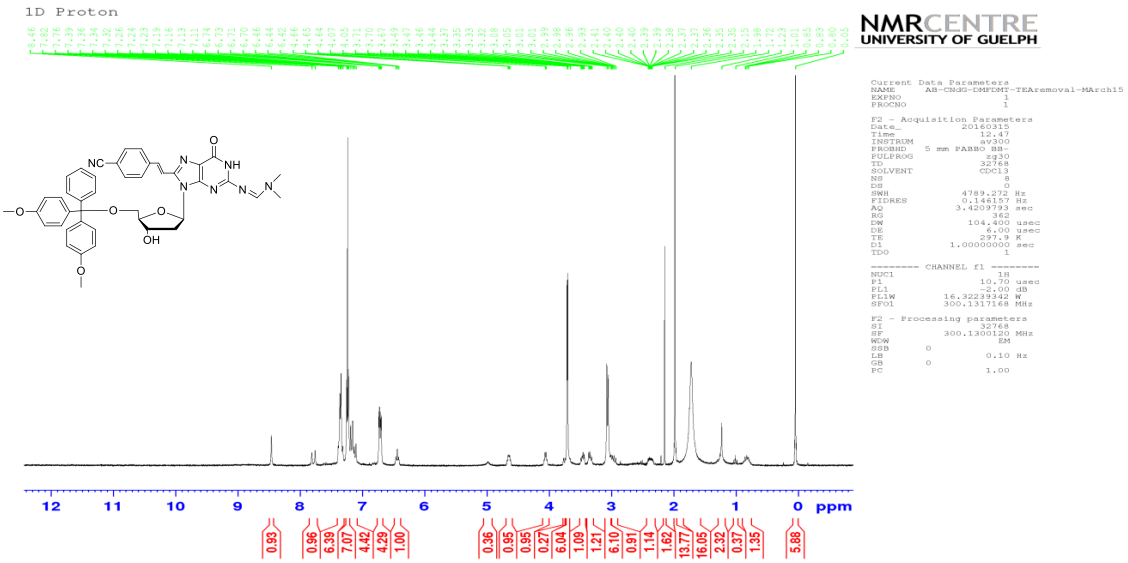


Figure S24. ^{1}HNMR (DMSO-*d6*) of **6**.

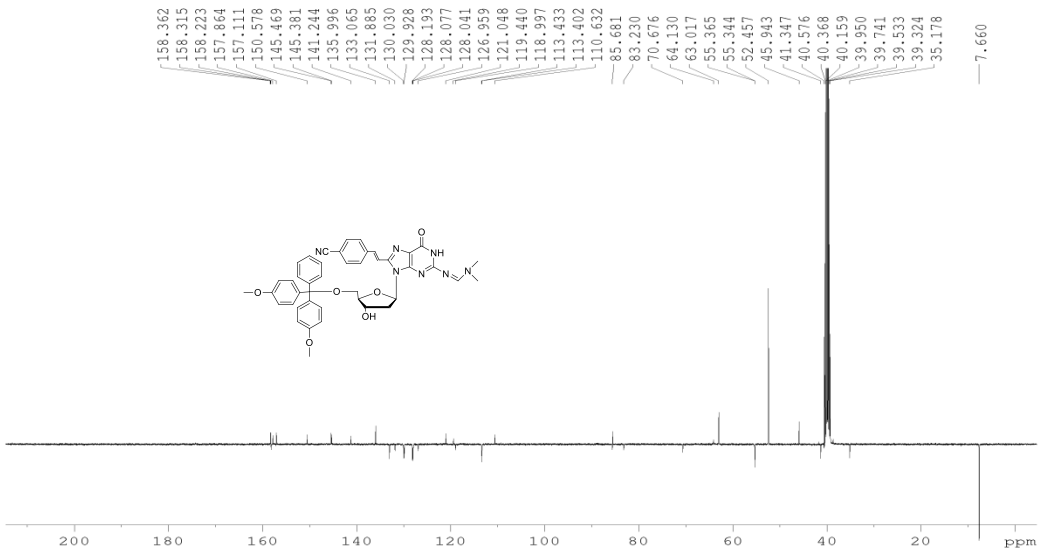


Figure S25. ^{13}CNMR ($CDCl_3$) of **6**.

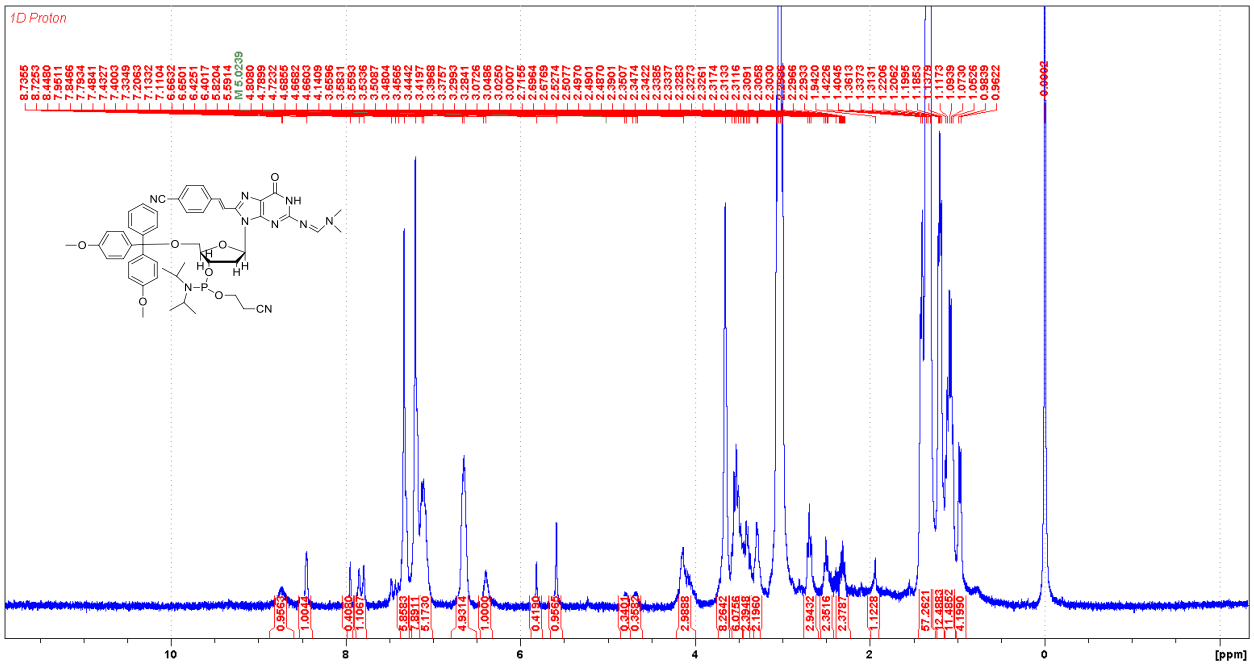


Figure S26. ^{1}HNMR ($CDCl_3$) of **7**.

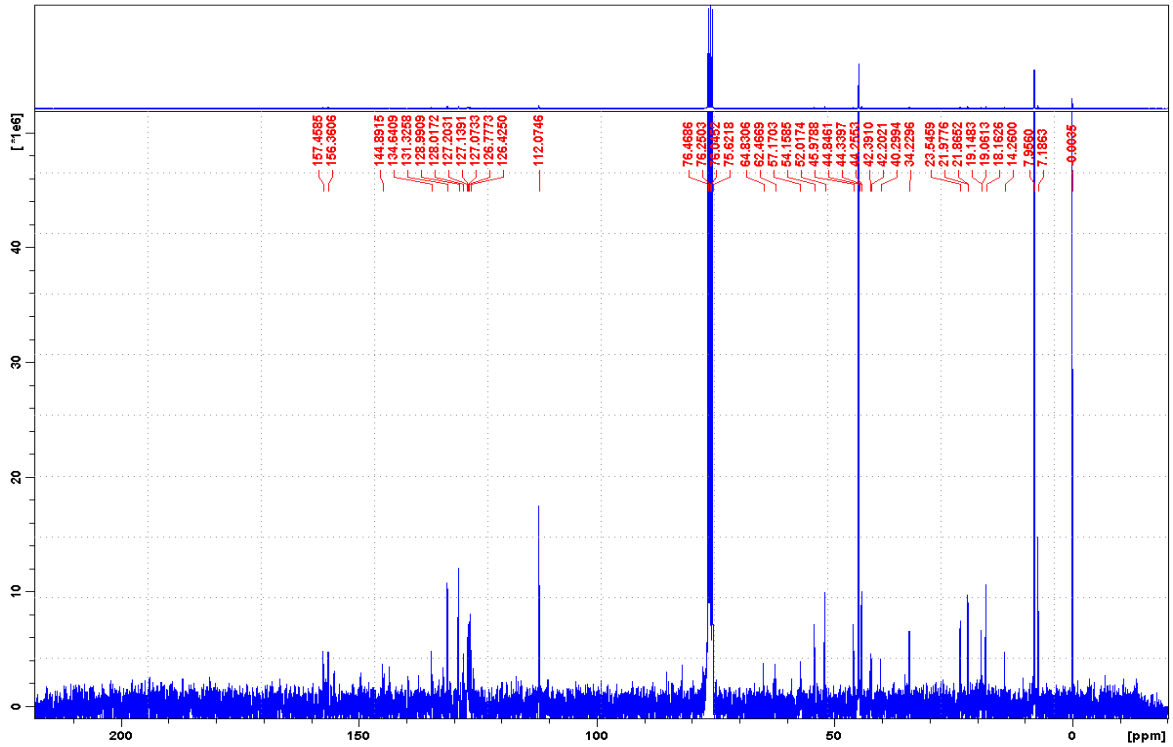


Figure S27. ^{13}CNMR ($CDCl_3$) of **7**.

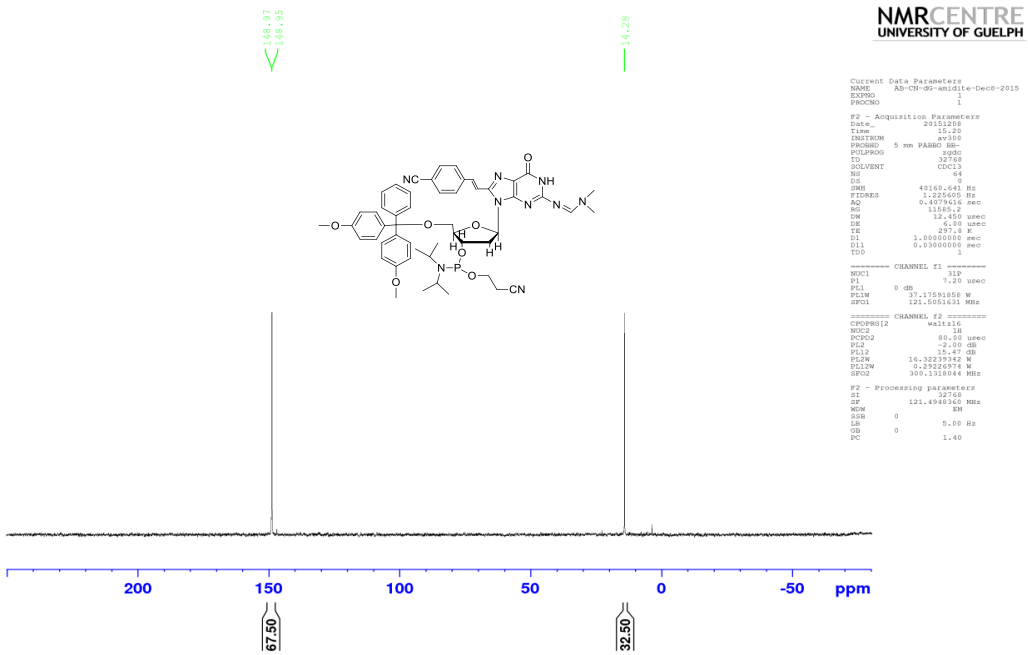


Figure S28. ^{31}PNMR ($CDCl_3$) of **7**.

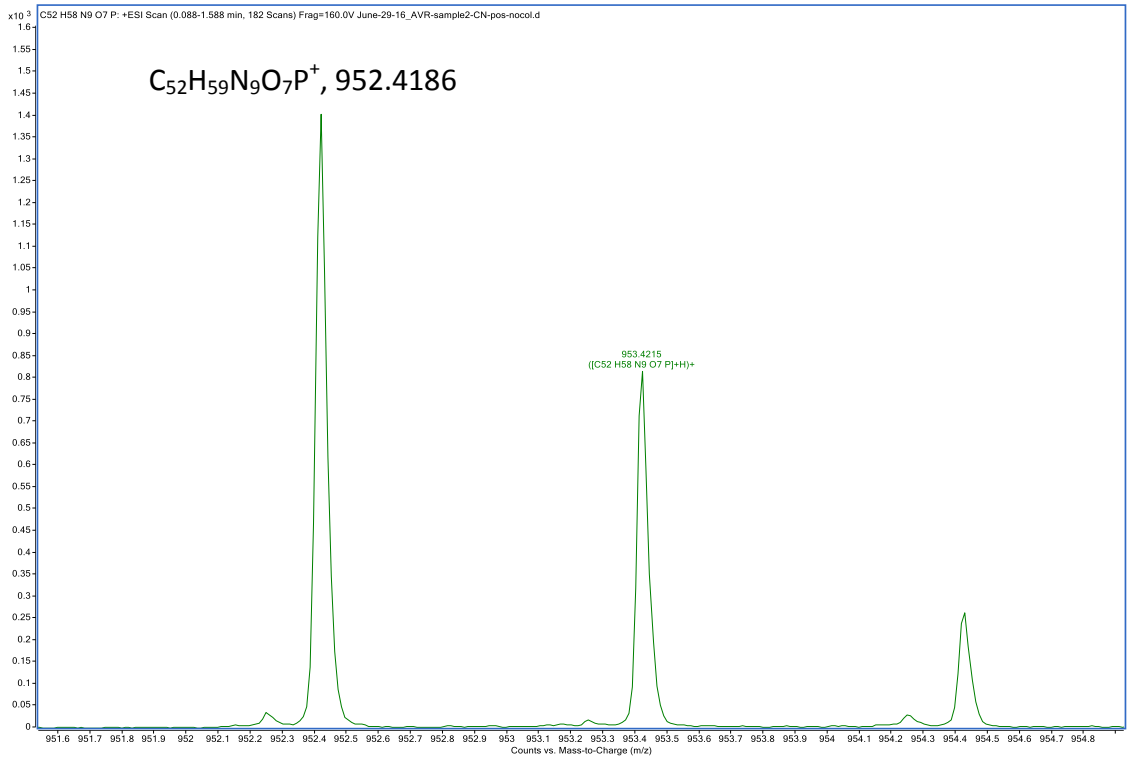


Figure S29. HRMS of **7**.

Table S10. Mass analysis of mTBA strands.

Oligo.	Product Formula	Calculated mass	Experimental m/z (ESI^-)	Experimental mass
ThdG−mTBA	$C_{154}H_{189}N_{57}O_{94}P_{14}S$	4808.2	$[M-6H]^{6-}$ = 800.30	4807.80
			$[M-7H]^{7-}$ = 685.80	4807.60
			$[M-8H]^{8-}$ = 600.00	4808.00
			$[M-9H]^{9-}$ = 533.20	4807.80
^{BTh}dG-mTBA	$C_{158}H_{191}N_{57}O_{94}P_{14}S$	4855.8	$[M-6H]^{6-}$ = 809.0	4860.0
			$[M-7H]^{7-}$= 693.0	4858.0
			$[M-8H]^{8-}$ = 606.1	4856.8
			$[M-9H]^{9-}$ = 538.7	4857.3
^{Sty}dG-mTBA	$C_{158}H_{193}N_{57}O_{94}P_{14}$	4828.2	$[M-5H]^{5-}$ = 964.5	4827.4
			$[M-6H]^{6-}$ = 803.6	4827.5
			$[M-7H]^{7-}$ = 688.7	4827.8
			$[M-8H]^{8-}$ = 602.5	4828.2
^{CN}dG-mTBA	$C_{159}H_{192}N_{58}O_{94}P_{14}$	4853.1	$[M-5H]^{5-}$ = 970.0	4855.0
			$[M-6H]^{6-}$ = 808.1	4854.6
			$[M-7H]^{7-}$ = 692.2	4852.4
			$[M-8H]^{8-}$ = 605.5	4852.0

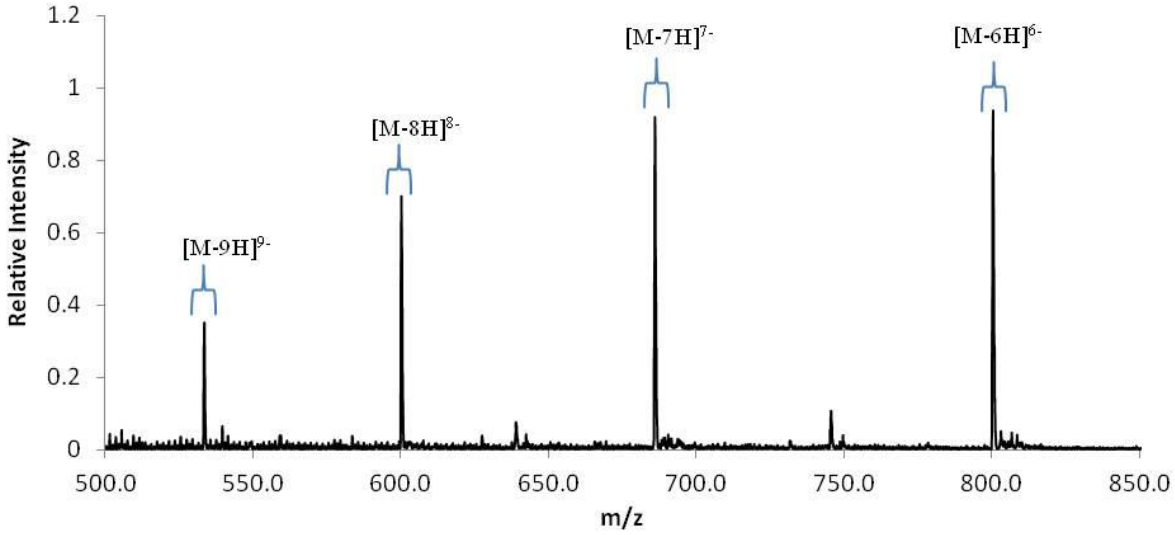


Figure S30. ESI-MS of ThdG-mTBA.

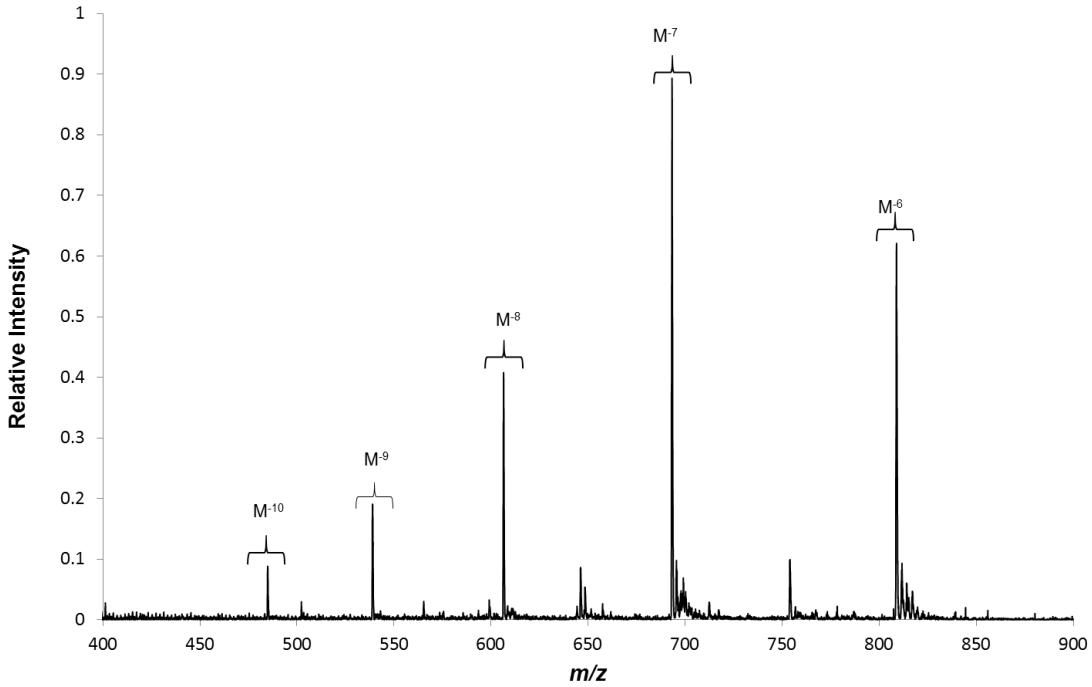


Figure S31. ESI-MS of ^{BTh}dG-mTBA.

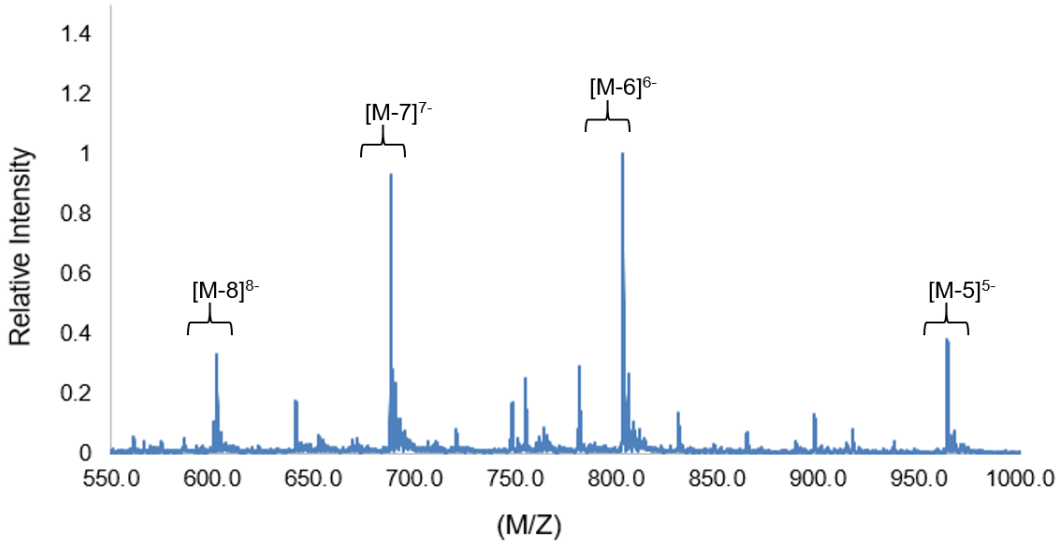


Figure S32. ESI-MS of ^{Sty}dG-mTBA.

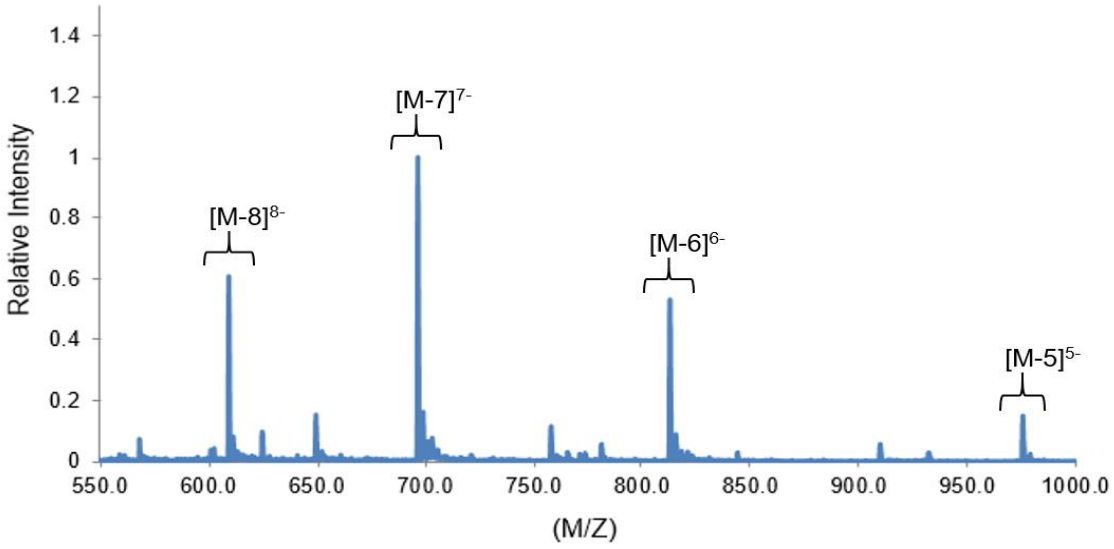


Figure S33. ESI-MS of ^{CN}dG-mTBA.