Supplementary Information

N7 Methylation Alters Hydrogen Bonding Patterns of Guanine in Duplex DNA

Yi Kou, Myong-Chul Koag and Seongmin Lee*

Division of Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, TX 78712

Synthesis of N7mG phosphoramidite. The N7mG phosphoramidite was prepared starting from a ribose derivative according to the synthetic procedures described previously (1).

DNA Sequences—Oligonucleotides were obtained from Integrated DNA Technologies (IDT) or Midland Certified Reagent Co. (Midland, TX). They were purified by the HPLC, and their sequences were verified by MALDI-TOF or ESI mass spectrometry. DNA substrates used for this crystallographic study consisted of a 16-mer template, a complementary 10-mer primer, and 5-mer downstream oligonucleotides (2). The templates DNA sequence used for crystallization were 5'-CCGACGTCGCAT[X]AGC-3', X= A, T, G, and C. The upstream primer sequences were 5'-GCT[G or N7mG]ATGCGC-3'. The downstream oligonucleotide sequence was 5'-GTCGG-3', and the 5' terminus was phosphorylated. The DNA sequences for this study were designed to be almost identical to a published binary complex structure (PDB ID 1BPX) to minimize sequence-dependent structural differences. The oligonucleotides were mixed and annealed to give a 1 mM mixture of gapped DNA as described (2).

Protein Expression and Purification—Pol β was induced and purified from *Escherichia coli* with minor modifications of the method described previously (3).

Pol β-HGC DNA crystallization—The pol β-HGC DNA complexes containing guanine or N7mG base paired with four different bases in a single-nucleotide gapped DNA were prepared under conditions similar to those described previously (2). Polβ was complexed with a single-nucleotide gapped DNA containing a 16-mer template (5'-CCGACGTCGCAT[X]AGC-3', X= A, T, G, and C), a complementary 10-mer primer (5'-GCT[G or N7mG]ATGCGC-3'), and a 5-mer downstream oligonucleotide (5'-pGTCGG-3'). The resulting eight different polβ-DNA complexes were used to obtain each of HGC complex crystals, respectively. All of the polβ-HGC DNA complex crystals were grown in the same condition containing 50mM imidazole, pH 7.5, 14–23% PEG3400, and 350mM sodium acetate as described previously (3). Crystals were cryo-protected in mother liquor supplemented with 12% ethylene glycol and were flash-frozen in liquid nitrogen.

Data Collection and Refinement—Diffraction data were collected at 100 K using either a Rigaku MicroMax-007 HF microfocus X-ray generator with R-Axis IV^{++} imaging plate area detector or the beamline 5.0.3 Advanced Light Source at Berkeley Center for Structural Biology. All diffraction data were processed using HKL 2000 (4). The structures of the pol β-HGC DNA complexes were solved by molecular replacement with pol β with a single-nucleotide gapped DNA (PDB code 1BPX) as the search model. The model was built using COOT and refined using PHENIX software (5,6) . Ramachandran plots were generated using MolProbity program (7).

Table S1. Data collection and refinement statistics.

PDB code	N7mG:dC (5DB6)	N7mG:dT (5DB7)	N7mG:dA (5DB8)	N7mG:dG (5DB9)	dG:dT (5DBA)	dG:dA (5DBB)	dG:dG (5DBC)
Data Collection							
Space Group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell constants							
a (Å)	54.138	54.396	54.149	54.201	54.344	54.389	54.458
b	78.828	79.480	78.904	79.278	79.345	79.372	79.352
c	54.764	54.912	54.751	54.720	54.584	54.938	54.864
α (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
β	105.92	105.36	105.39	105.63	106.14	105.45	105.45
γ	90.00	90.00	90.00	90.00	90.00	90.00	90.00
Resolution (Å)[a]	20-2.83 (2.88-2.83)	20-2.22 (2.26-2.22)	20-2.54 (2.58-2.54)	20-2.45 (2.49-2.45)	20-1.97 (2.00-1.97)	20-2.25 (2.29-2.25)	20-2.40 (2.44-2.40)
R_{merge} [b] (%)	0.143 (0.395)	0.084 (0.584)	0.137 (0.505)	0.117 (0.446)	0.072 (0.210)	0.130 (0.378)	0.109 (0.349)
$\langle I/\sigma \rangle$	9.4 (2.10)	20.4 (3.54)	11.2 (2.00)	13.8 (2.00)	12.9 (3.07)	17.9 (2.89)	16.0 (2.60)
Completeness (%)	100 (99.8)	99.9 (99.7)	100 (99.7)	99.4 (96.0)	99.6 (99.1)	100 (99.5)	94.7 (94.6)
Redundancy	4.1 (3.7)	4.8 (4.5)	5.4 (4.5)	4.8 (4.1)	4.1 (3.7)	4.2 (3.5)	3.2 (3.0)
Refinement							
R_{work} [c]/R_{free} [d] (%)	19.9/26.6	22.1/27.3	20.0/26.2	20.2/25.3	20.2/24.5	20.5/26.5	21.4/27.9
Unique reflections	10562	22542	14558	16411	31502	21462	16613
Mean B factor ($Å^2$)							
Protein	23.5	35.4	29.7	33.3	37.6	34.3	33.1
Ligand	27.2	38.5	31.1	36.5	36.6	35.1	32.0
Solvent	19.3	34.2	29.5	33.3	35.2	36.6	30.5
Ramachandran plot							
Most favored (%)	96.0	96.6	97.2	97.5	96.0	96.6	96.0
Additional allowed (%)	3.7	3.1	2.5	2.2	3.7	3.4	4.0
RMSD							
Bond lengths (Å)	0.003	0.003	0.003	0.003	0.008	0.003	0.003
Bond angles (°)	0.834	0.855	0.876	0.820	1.143	0.642	0.739

[a] Values in parentheses are for the highest resolution shell.

[b] $R_{merge} = \Sigma|I-\langle I\rangle|/\Sigma I$ where I is the integrated intensity of a given reflection.

[c] $R_{work} = \Sigma|F(obs)-F(calc)|/\Sigma F(obs)$.

[d] $R_{free} = \Sigma|F(obs)-F(calc)|/\Sigma F(obs)$, calculated using 5% of the data.

2. Melting assays for N7mG:dN- and dG:dN-containing DNA

The DNA duplex was first prepared by annealing the template strand containing N7mG (Template strand: 5'-CCGAC/N7mG/TCGCATCAGC-3') and the complementary strand containing pairing dN (Complementary strand: 5'-GCTGATGCGA/dN/GTCGG-3') in annealing buffer (pH 7.4) containing 50 mM $MgCl_2$ and 20mM NaCl. For acquisition of fluorescence signal, SYBR Green I was used (10,000X concentrate in DMSO, Life Technologies). A 20 µl reaction mixture for each assay was then prepared by mixing the 10 µl of DNA duplex (25 ng/µl), with 1 µl (500X) SYBR green I dye in 9µl annealing buffer. The reaction mixtures were then loaded into 96 well PCR Microplate (Axygen Scientific). The melting assay was performed by LightCycler 480 (Roche Applied Science), which is programmed to start

with the initial stabilization at 20 ˚C for 30 min, then to increase with linear step of 0.04 ˚C/min from 20 ˚C to the 99 ˚C, and finally cool down to 20 ˚C. Three independent repeats were performed to calculate the average and standard deviation of the melting temperature.

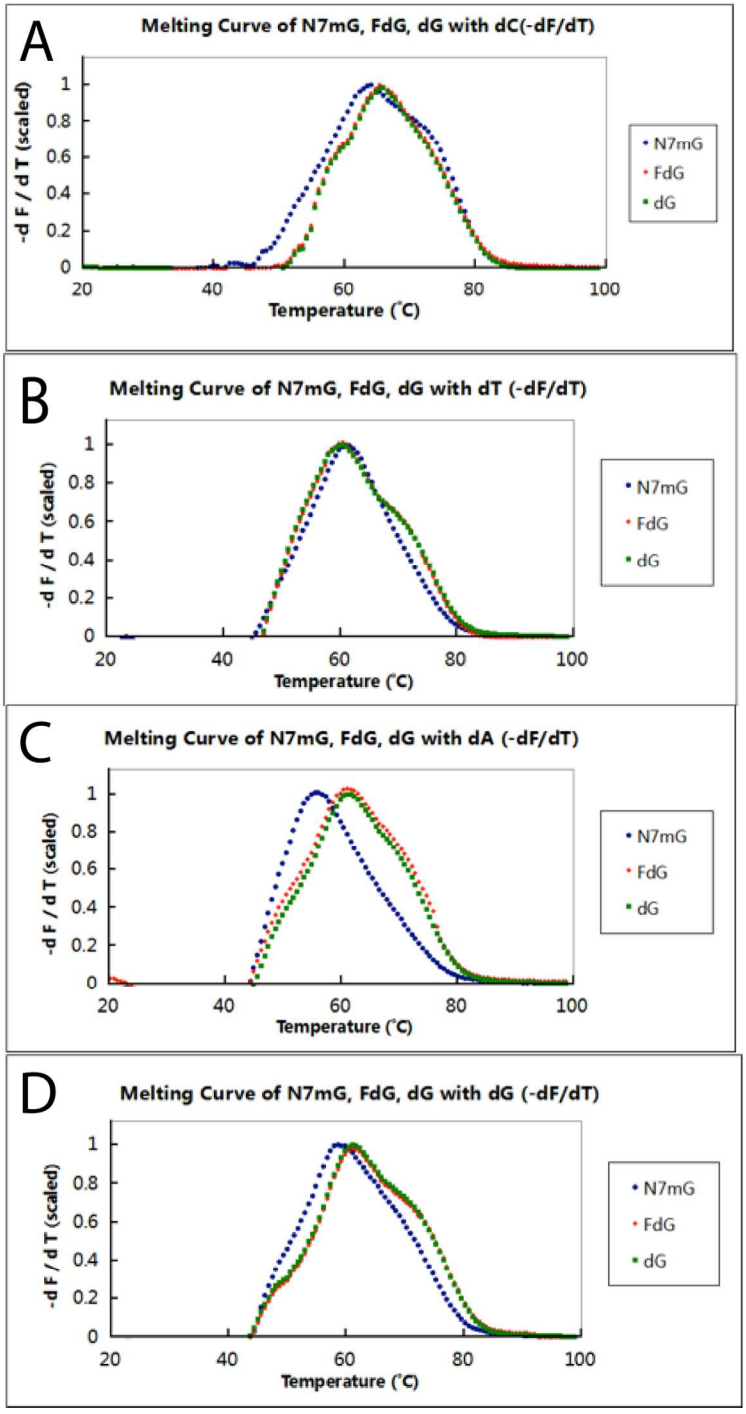


Figure S1. Melting curves of N7mG, FdG and dG with dC, dT, dG and dA.

3. Distortion parameters for the N7mG:dN HGC complexes

All distortion parameters were calculated by 3DNA v2.1 (8) based on the HGC structures described in the paper (Figure S2). In order to compare without bias, normal dG:dC base pair is also introduced at the same position in the HGC. The units for Shear, Stretch and Stagger distortions are distances (Å). Units for Buckle, Propeller and Opening distortions are angles (˚). Local base pair parameters describe only the distortions within the N7mG:dN and its neighbouring dA:dT base pairs. Local base pair step parameters describe the spatial distortions between the base pairs involving N7mG. The units for Shift, Slide and Rise distortions are distances (Å); for Tilt, Roll and Twist distortions are angles (°).

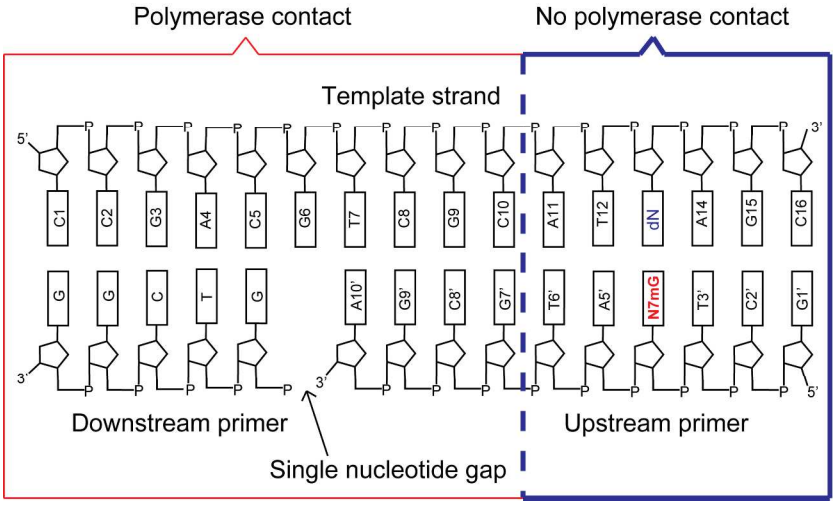


Figure S2. Schematic diagram of polβ Host-Guest-Complex system.

3.1 Distortion parameters of N7mG:dC

Table S2.1. Distortion parameters of N7mG:dC

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger(Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
N7mG:dC	-0.14	-0.02	-0.33	10.68	-15.59	14.08
dG:dC	0.11	-0.07	0.17	-4.60	-13.41	-1.19

Comparison of the distortions parameters of N7mG:dC in the HGC complex to the normal dG:dC pair in the same position (in our own HGC structure, the same as follows). The distortions are limited to the indicated base pairs only.

Table S2.2. Distortion parameters of neighbouring base pairs of N7mG:dC.

Local base pair parameters	Shear(Å)	Stretch(Å)	Stagger(Å)	Buckle(˚)	Propeller(˚)	Opening (˚)
A5':T12 (N7mG)	0.21	-0.16	0.28	2.46	-17.39	-1.01
T3':A14 (N7mG)	-0.17	-0.12	0.23	7.49	-16.72	1.78
A5':T12 (dG)	0.15	-0.16	0.03	5.75	-11.47	-4.13
T3':A14 (dG)	-0.08	-0.06	0.09	-4.84	-8.30	3.05

Table S2.3. Distortion parameters of neighbouring base pairs of N7mG:dC.

Local base pair step parameters	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (˚)	Roll (˚)	Twist (˚)
dT12dC13/N7mG4'dA5'	1.61	0.67	2.91	10.05	0.08	39.60
dC13dA14/dT3'N7mG4'	-1.63	0.36	3.64	-8.23	4.03	32.99
dT12dC13/dG4'dA5'	0.19	-0.24	3.36	1.31	0.31	41.71
dC13dA14/dT3'dG4'	-0.31	-0.030	3.24	0.10	5.76	34.13

Table S2.4 Distortion parameters comparison of N7mG:dC in different HGC systems

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
N7mG:dC (Polβ)	-0.14	-0.02	-0.33	10.68	-15.59	14.08
N7mG:dC (AlkA)	0.01	0.10	0.21	-3.95	-12.81	5.60
dG:dC	0.11	-0.07	0.17	-4.60	-13.41	-1.19

Figure S2.1. Shear distortion of base pairs in oligo containing N7mG:dC vs. dG:dC

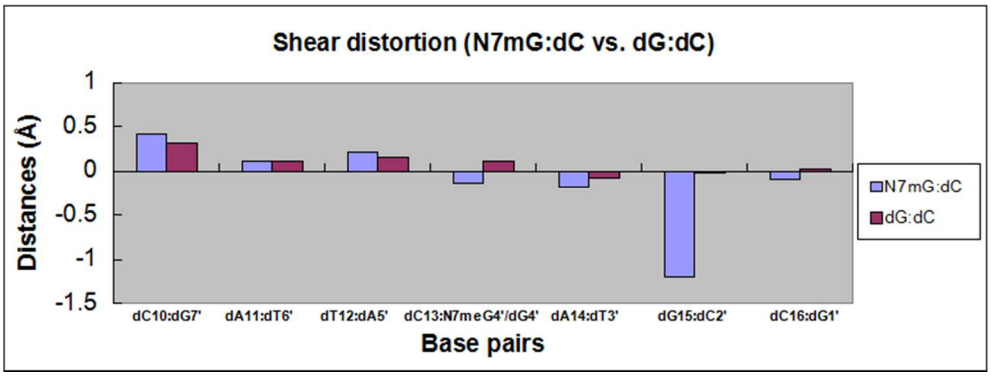


Figure S2.2. Stagger distortion of base pairs in oligo containing N7mG:dC vs. dG:dC

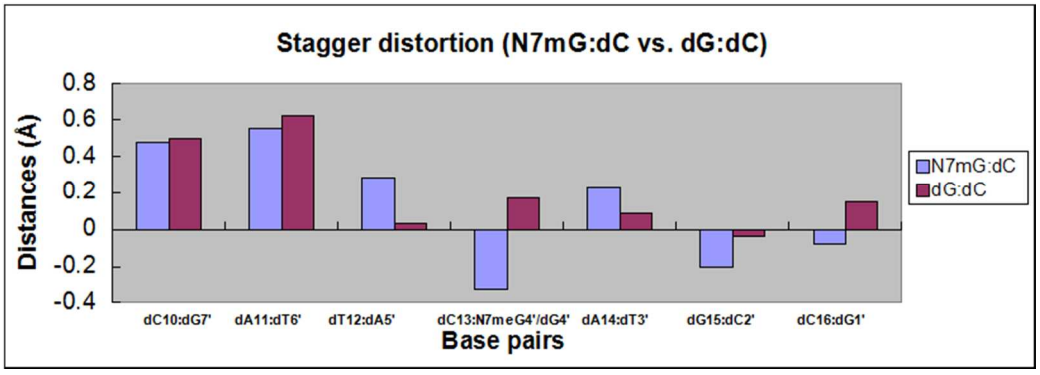


Figure S2.3. Buckle distortion of base pairs in oligo containing N7mG:dC vs. dG:dC

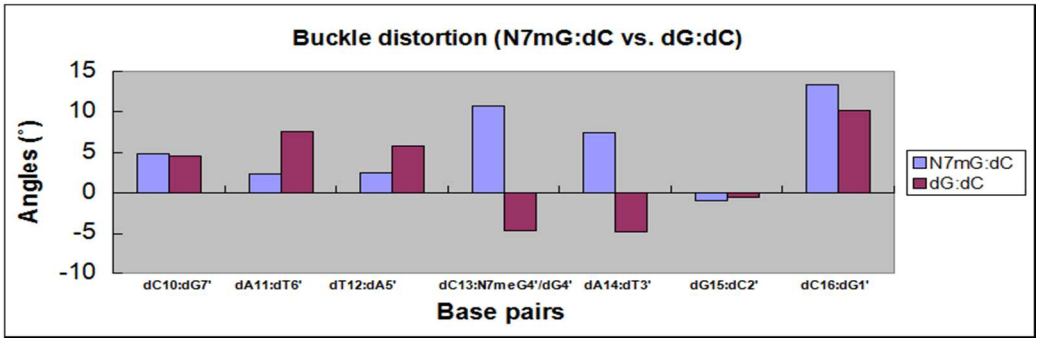


Figure S2.4. Opening distortion of base pairs in oligo containing N7mG:dC vs. dG:dC

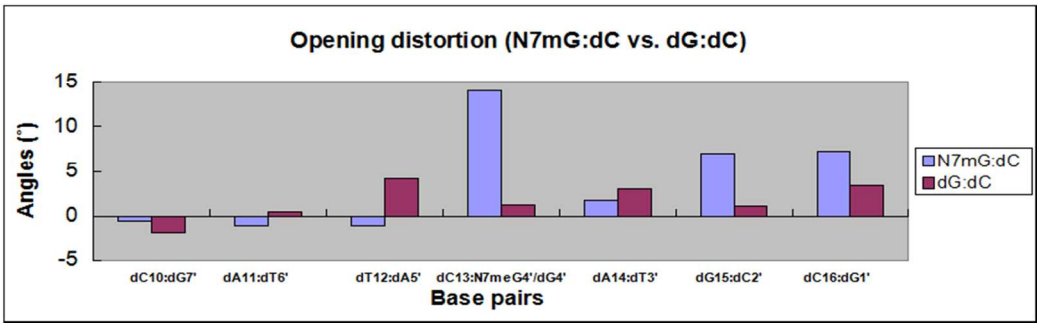


Figure S2.5. Slide distortion of base pairs in oligo containing N7mG:dC vs. dG:dC

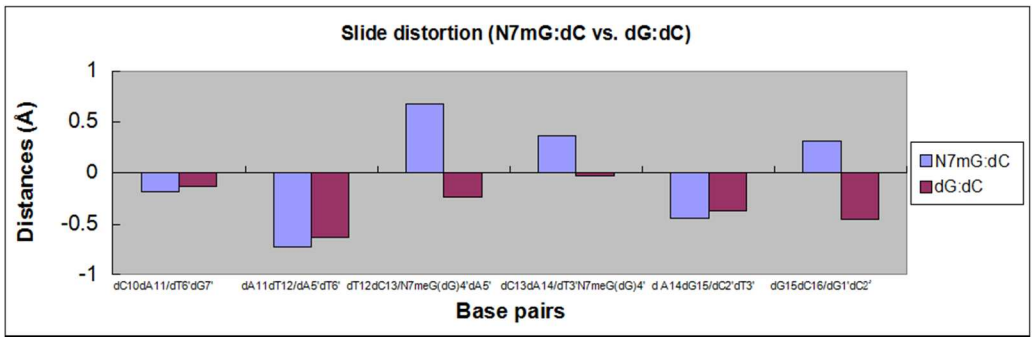
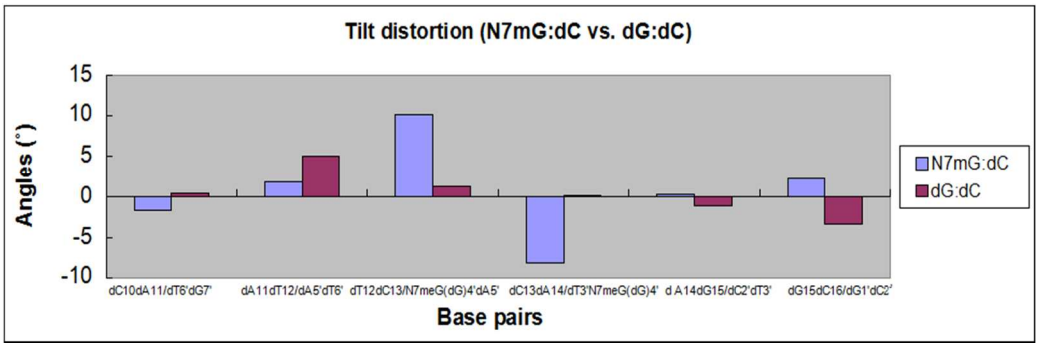


Figure S2.6. Tilt distortion of base pairs in oligo containing N7mG:dC vs. dG:dC



3.2 Distortion parameters of N7mG:dA

Table S2.6. Local base pair distortions of N7mG:dA vs. dG:dA

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
N7mG:dA	-1.99	-3.13	-0.70	11.25	19.77	86.11
dG:dA	0.31	-4.77	-0.020	11.44	10.53	85.78

Figure S2.7. Shear distortion of base pairs in oligo containing N7mG:dA vs. dG:dA

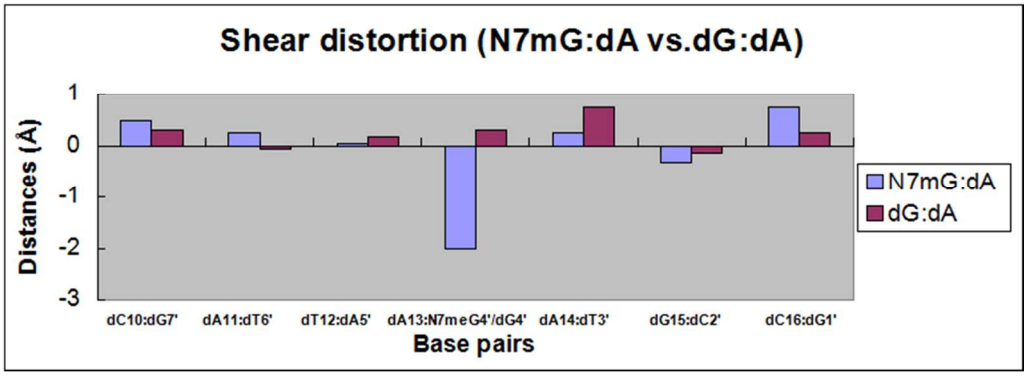


Figure S2.8. Stagger distortion of base pairs in oligo containing N7mG:dA vs. dG:dA

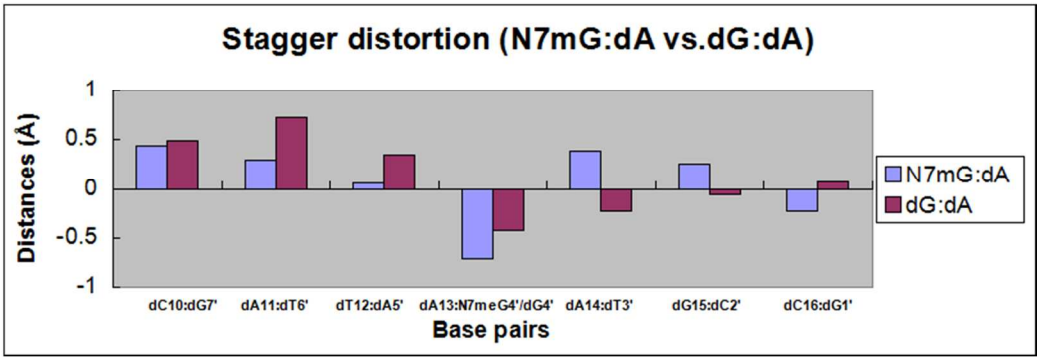


Figure S2.9. Propeller distortion of base pairs in oligo containing N7mG:dA vs. dG:dA

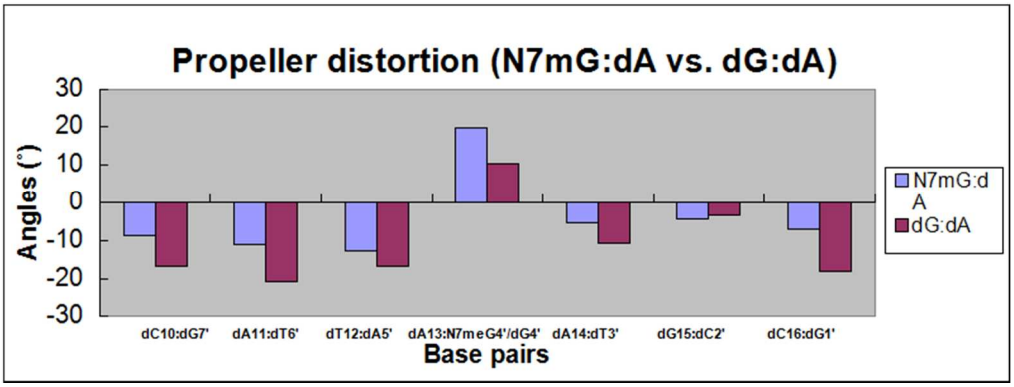


Table S2.7. Local base pair distortions of neighboring dA:dT base pairs of N7mG:dA vs. dG:dA

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
dA5':dT12 (N7mG:dA)	0.05	-0.14	-0.06	9.96	-12.47	-0.01
dT3':dA14 (N7mG:dA)	0.25	-0.20	0.39	7.45	-5.32	12.30
dA5':dT12 (dG:dA)	0.16	0.03	0.35	2.90	-16.46	7.06
dT3':dA14 (dG:dA)	0.75	-0.18	-0.23	-3.19	-10.66	5.24

Table S2.8. Local base pair step distortions of N7mG:dA

Local base pair step parameters	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (˚)	Roll (˚)	Twist (˚)
dT12dA13/N7mG4'A5'	0.60	-3.67	2.37	-173.25	12.23	0.08
dA13dA14/T3'N7mG4'	-3.01	-1.28	-0.77	133.39	-114.73	23.07
dT12dA13/dG4'dA5'	0.46	-3.42	2.13	-174.50	1.48	31.97
dA13dA14/dT3'dG4'	0.92	3.26	-0.61	-134.43	110.45	-154.01

Figure S2.10. Roll distortion of base pairs in oligo containing N7mG:dA vs. dG:dA

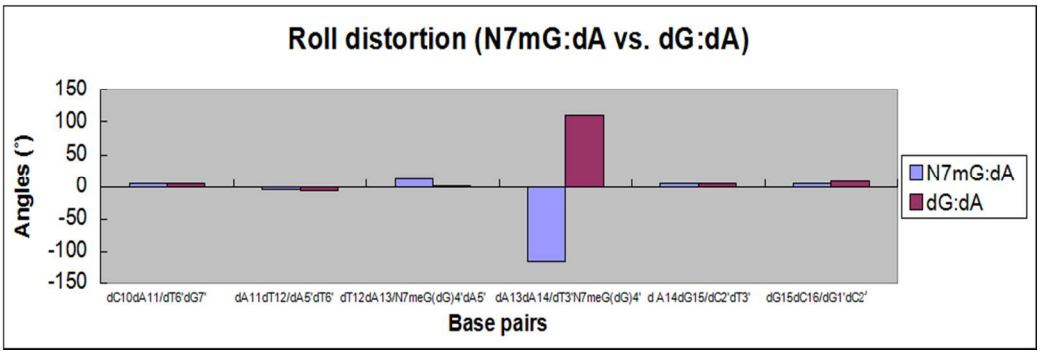
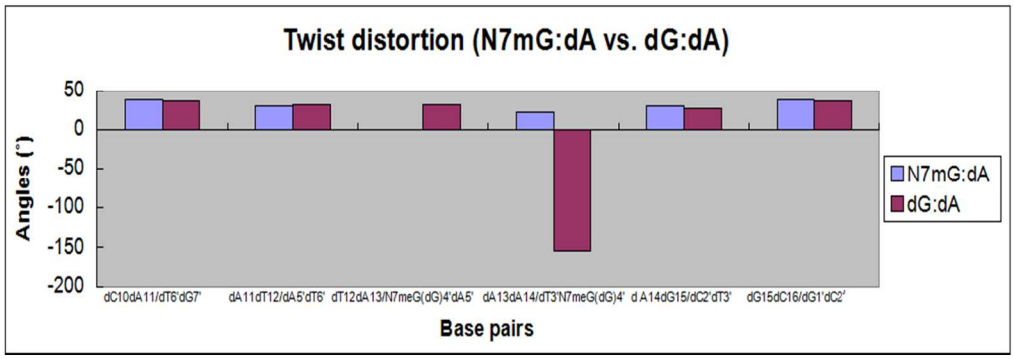


Figure S2.11. Twist distortion of base pairs in oligo containing N7mG:dA vs. dG:dA



3.3 Distortion parameters of N7mG:dT

Table S2.10. Local base pair distortions of N7mG:dT vs. dG:dT

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
N7mG:dT	0.75	-0.26	0.46	1.29	-13.72	3.12
dG:dT	2.32	-0.64	-0.19	3.86	-9.03	-0.49

Figure S2.12. Stretch distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

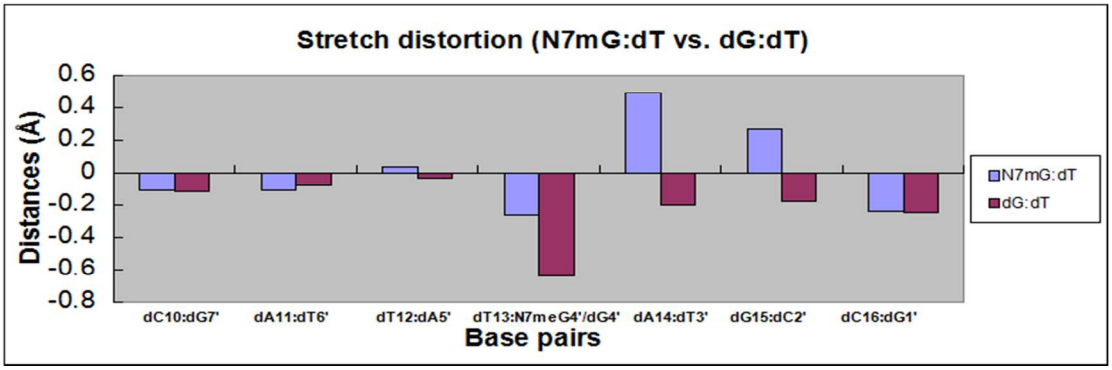


Figure S2.13. Stagger distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

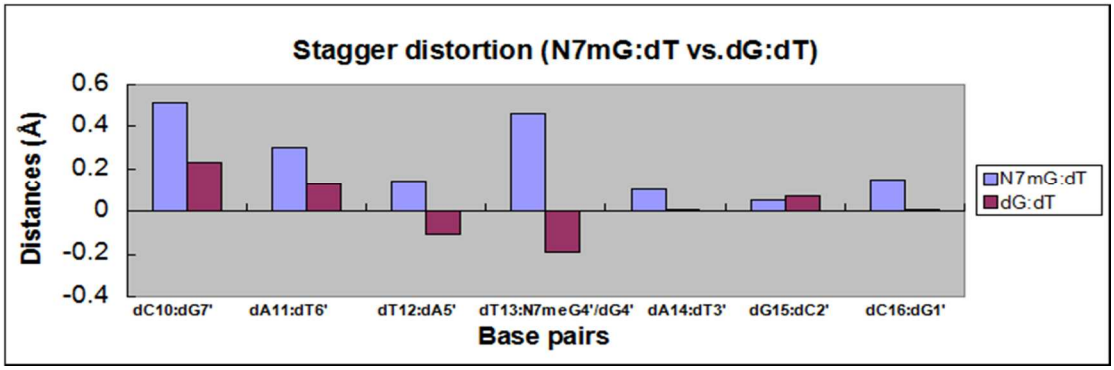


Figure S2.14. Buckle distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

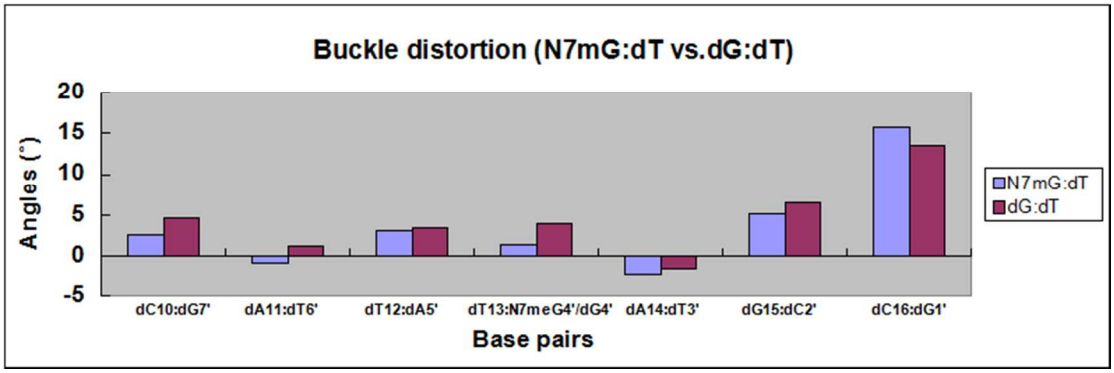


Figure S2.15. Opening distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

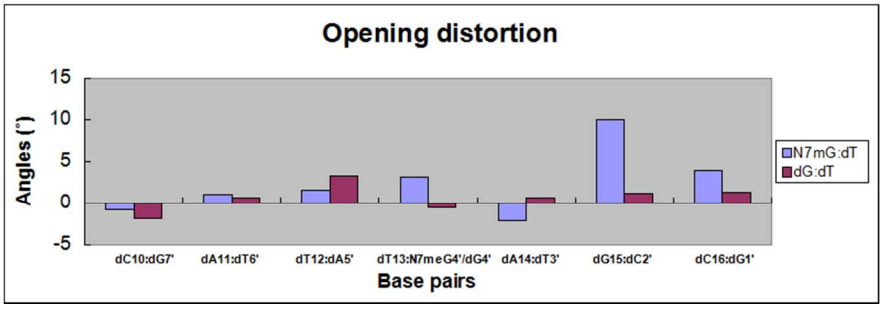


Table S2.11. Local base pair distortions of neighbouring bases of N7mG:dT vs. dG:dT

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (°)	Propeller (°)	Opening (°)
dA5':dT12 (N7mG:dT)	0.23	0.04	0.14	3.04	-18.05	-1.62
dT3':dA14 (N7mG:dT)	0.07	-0.49	0.10	-2.28	-11.52	-2.02
dA5':dT12 (dG:dT)	-0.12	-0.03	-0.11	3.40	-12.24	3.37
dT3':Ad14 (dG:dT)	-0.11	-0.20	0.01	-1.54	-6.12	0.65

Table S2.12. Local base pair step distortions of neighbouring bases N7mG:dT vs. dG:dT

Local base pair step parameters	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (°)	Roll (°)	Twist (°)
dT12dT13/N7mG4'dA5'	0.75	0.17	3.41	-0.73	-2.31	40.08
dT13dA14/dT3'N7mG4'	-1.28	-0.58	3.40	-3.19	2.17	31.17
dT12dT13/dG4'dA5'	0.49	0.10	3.22	3.12	1.46	47.12
dT13dA14/dT3'dG4'	-0.51	-0.16	3.29	-0.14	7.26	28.24

Figure S2.16. Shift distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

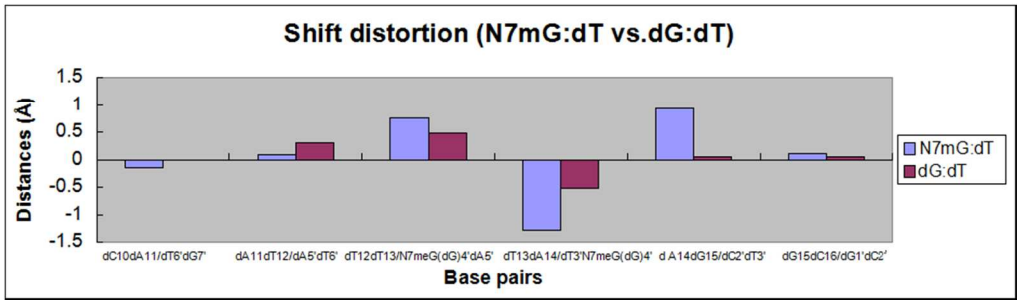


Figure S2.17. Slide distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

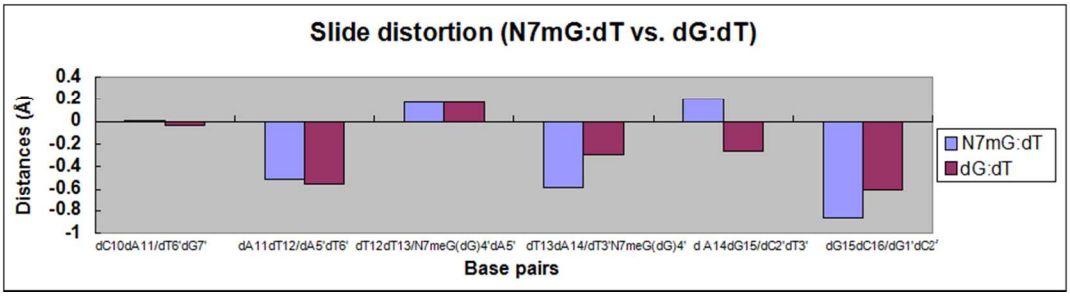


Figure S2.18. Tilt distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

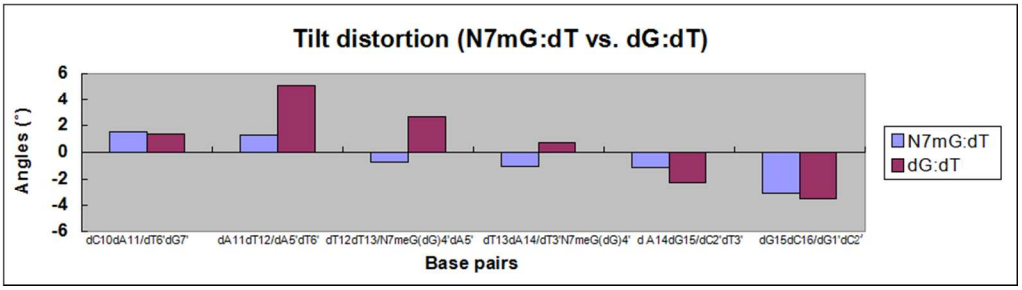
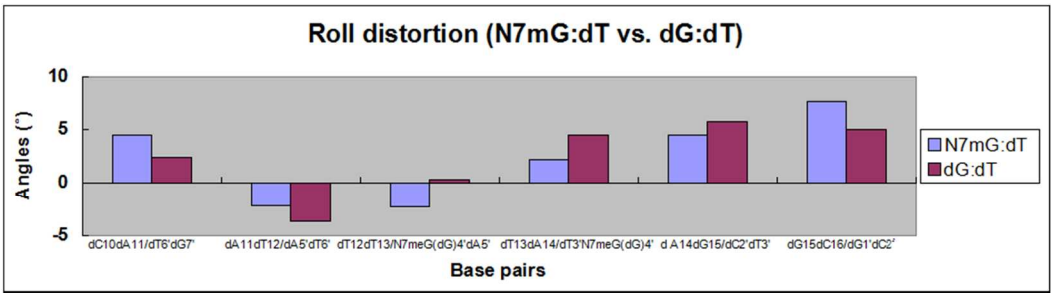


Figure S2.19. Roll distortion of base pairs in oligo containing N7mG:dT vs. dG:dT



3.4 Distortion parameters of N7meG:dG

Table S2.14. Local base pair distortion parameters of N7mG:dG vs. dG:dG

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (°)	Propeller (°)	Opening (°)
N7mG:dG	-1.94	-3.14	-0.58	14.48	15.17	77.37
dG:dG	-1.70	-3.31	-0.76	13.87	6.75	80.49

Table S2.15. Neighbouring base pair distortion parameters of N7mG:dG vs. dG:dG

Local base pair parameters	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (˚)	Propeller (˚)	Opening (˚)
dA5':T12 (N7mG:dG)	0.17	-0.04	-0.03	6.14	-16.57	-2.27
dT3':A14 (N7mG:dG)	0.03	-0.16	0.67	5.10	-9.39	1.13
dA5':T12 (dG:dG)	0.16	-0.07	0.18	4.58	-15.15	1.38
dT3':A14 (dG:dG)	0.04	-0.11	0.17	-1.02	-12.33	0.60

Figure S2.20. Propeller distortion of base pairs in oligo containing N7mG:dT vs. dG:dT

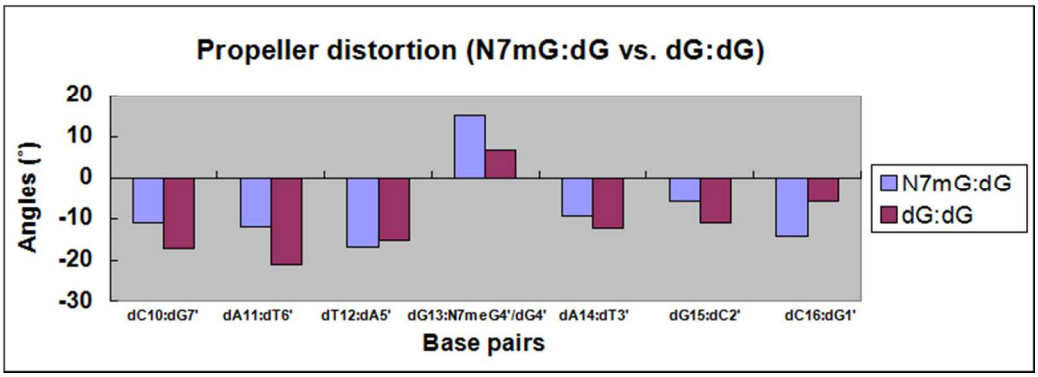


Figure S2.21. Stagger distortion of base pairs in oligo containing N7mG:dG vs. dG:dG

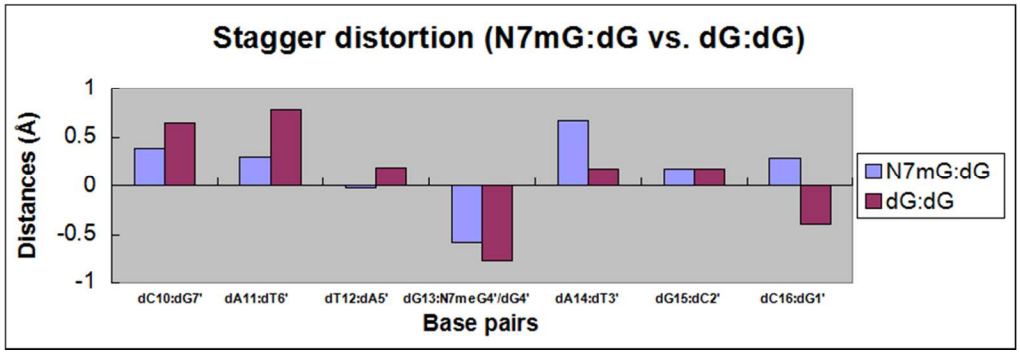


Figure S2.22. Buckle distortion of base pairs in oligo containing N7mG:dG vs. dG:dG

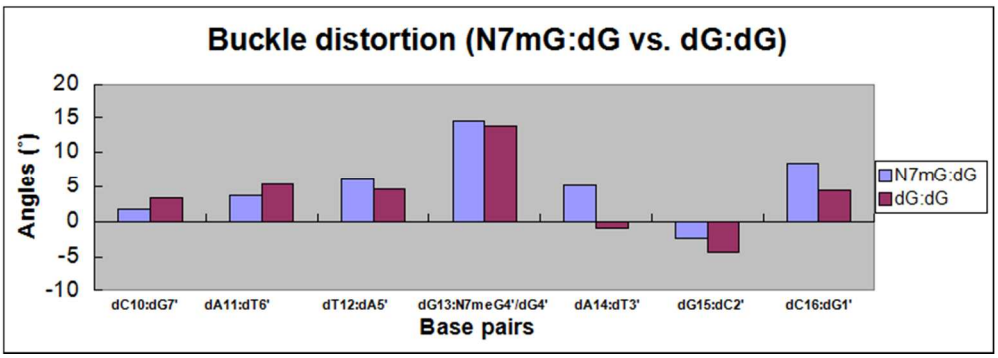


Figure S2.23. Opening distortion of base pairs in oligo containing N7mG:dG vs. dG:dG

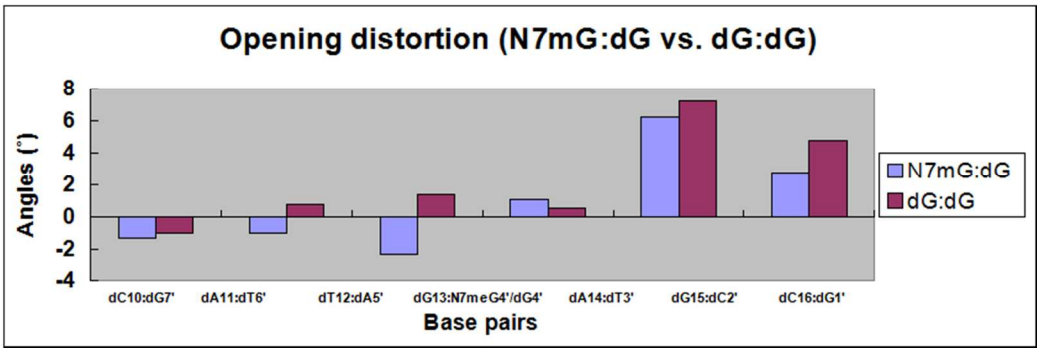
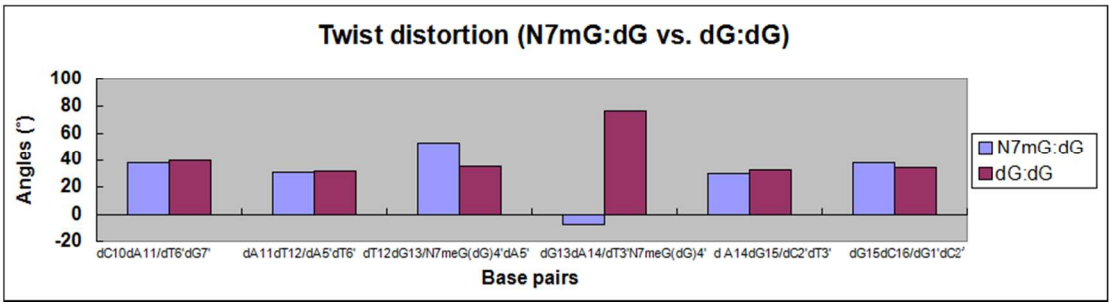


Table S2.16. Local base pair step distortion parameters of N7mG:dG vs. dG:dG

Local base pair step parameters	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (˚)	Roll (˚)	Twist (˚)
dT12dG13/N7mG4'dA5'	1.68	-3.62	1.88	-174.57	14.85	52.75
dG13dA14/dT3'N7mG4'	-3.47	-0.86	-1.40	129.08	-115.21	-7.21
dT12dG13/dG4'dA5'	1.16	-3.51	1.82	-175.10	19.19	35.52
dG13dA14/dT3'dG4'	-2.30	-1.99	-2.65	130.34	-117.82	76.82

Figure S2.24. Twist distortion of base pairs in oligo containing N7mG:dG vs. dG:dG



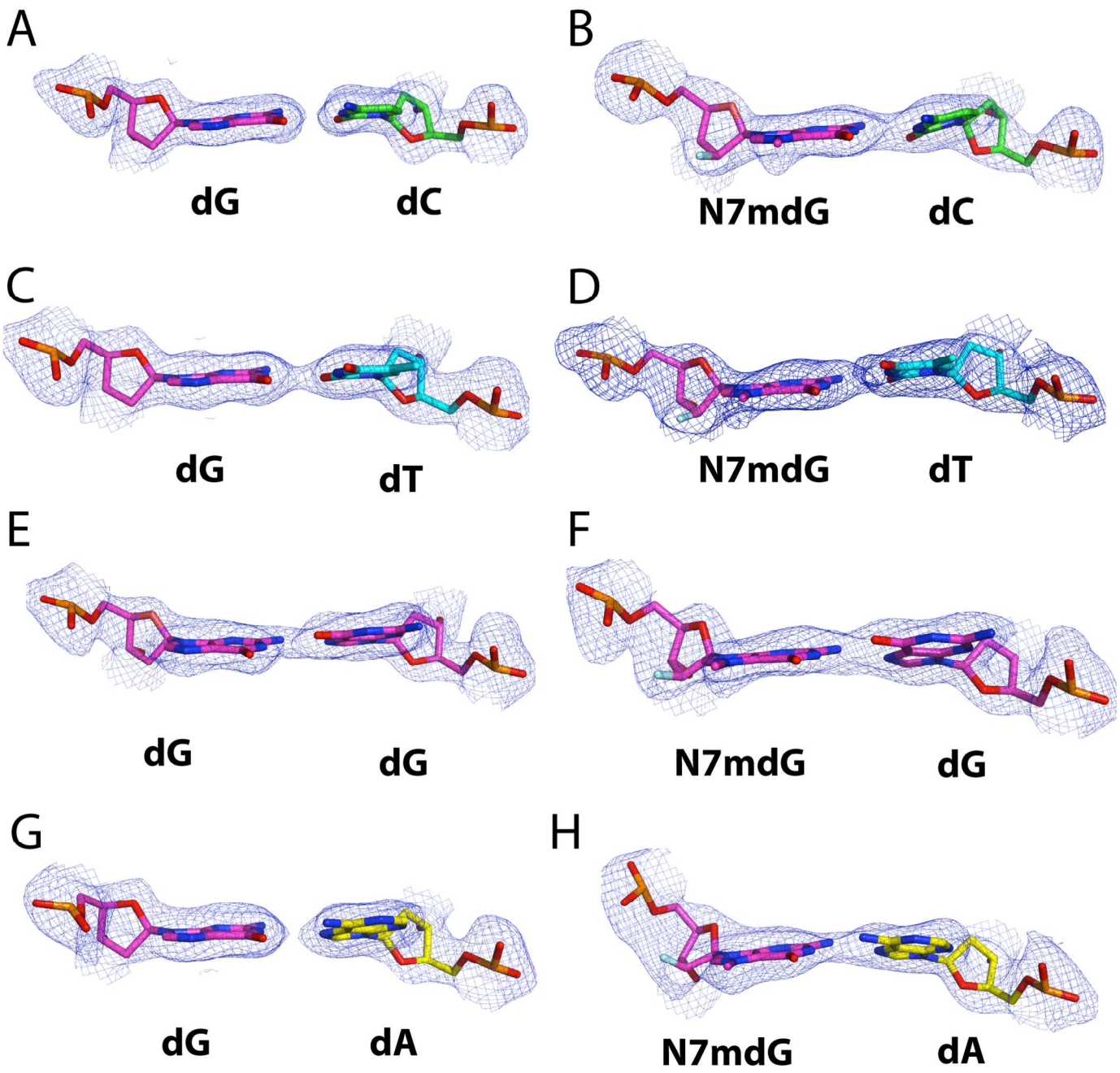


Figure S3. Side view of structures of dG:dN and N7mdG:dN base pairs determined by the polβ-HGC complex. Base pairing of (A) dG:dC, (B) N7mdG:dC, (C) dG:dT, (D) N7mdG:dT, (E) dG:dG, (F) N7mdG:dG, (G) dG:dA and (H) N7mdG:dA. $2F_o$-F_c electron density maps contoured at 1σ around the base pairs in polβ-HGC complexes.

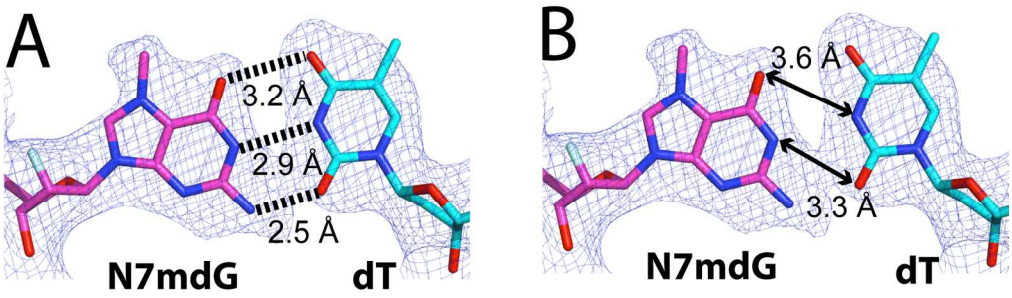


Figure S4. Distances between potential hydrogen bond donors and acceptors in the Watson-Crick-like N7mdG:dT base pair. (A) Interbase hydrogen bonds in the N7mdG:dT mismatch. (B) Distances between O6(N7mdG)-N3(dT) and N1(N7mdG)-O2(dT) are 3.6 Å and 3.3 Å, respectively, indicating that N7mdG's keto tautomer does not significantly contribute to the Watson-Crick-like N7mdG:dT conformation.

References.

1. Lee, S.; Bowman, B. R.; Ueno, Y.; Wang, S.; Verdine, G. L. *J. Am. Chem. Soc.* **2008**, *130*, 11570-11571
2. Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757–766
3. Koag, M.-C.; Lee, S. *J. Am. Chem. Soc.* **2014**, *136*, 5709–5721
4. Otwinowski, Z.; Minor, W. *Methods in Enzymology* 1997, 307–326, Elsevier [online] http://linkinghub.elsevier.com/retrieve/pii/S007668799776066X
5. Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132
6. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 1948– 1954
7. Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, W375–383
8. Xiang-Jun L.; Wilma K. O. *Nucleic Acids Res.* **2003**, *31*, 5108-5121.