SUPPORTING INFORMATION

The abundant DNA adduct N^7-methyl deoxyguanosine miscodes during replication by human DNA polymerase η

Olive J. Njuma, Yan Su, and F. Peter Guengerich

CONTENTS

FIGURE S11. **LC-MS sequence analysis of extension products opposite *N*^{7}-CH_3 2´-F dG with hpol η**.

FIGURE S12. **LC-MS sequence analysis of extension products opposite *N*^{7}-CH_3 2´-F dG with hpol η**.

TABLE S1. **Observed and theoretical CID ions of the 5´-pTCATGA-3´ extension product with hpol η**.

TABLE S2. **Observed and theoretical CID ions of the 5´-pTACTGA-3´ extension product with hpol η**.

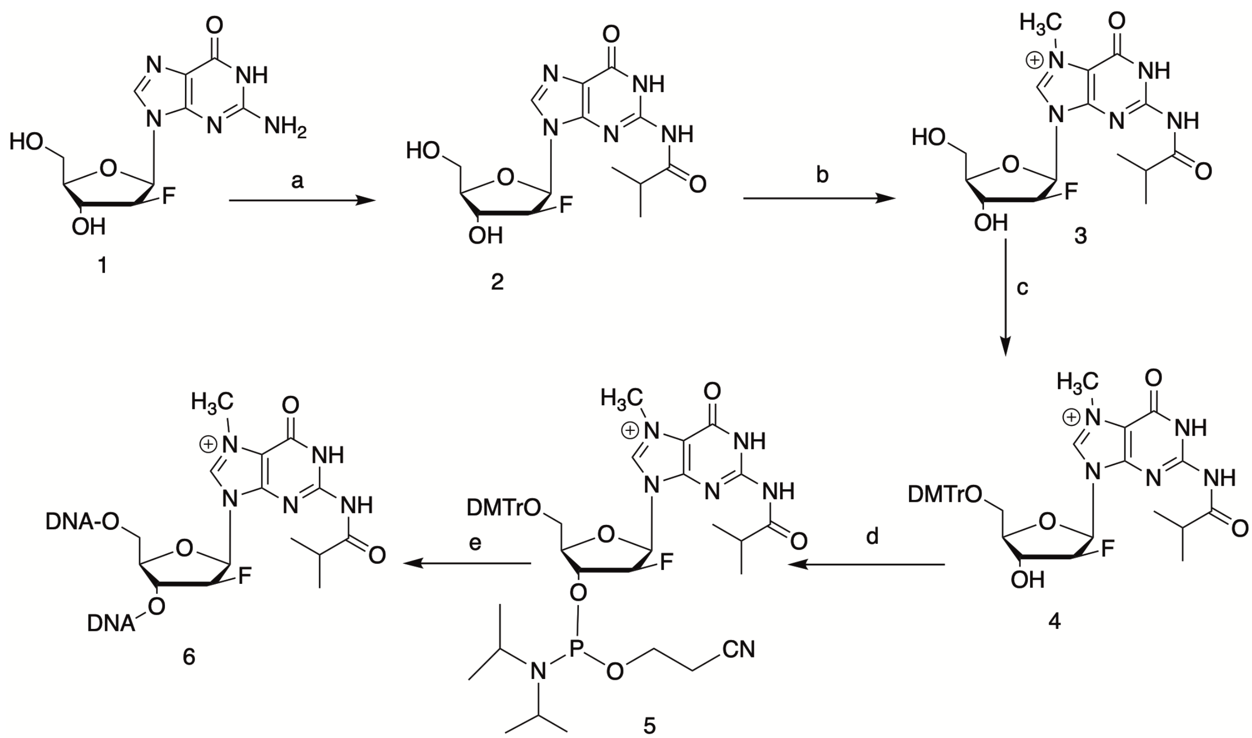
TABLE S3. **Observed and theoretical CID ions of the 5´-pTACT-3´ extension (stalled) product with hpol η**.

TABLE S4. **Observed and theoretical CID ions of the 5´-pTCATGAT-3´ extension product with hpol η**.

TABLE S5. **Observed and theoretical CID ions of the 5´-pTAGTCAT-3´ extension product with hpol η**.

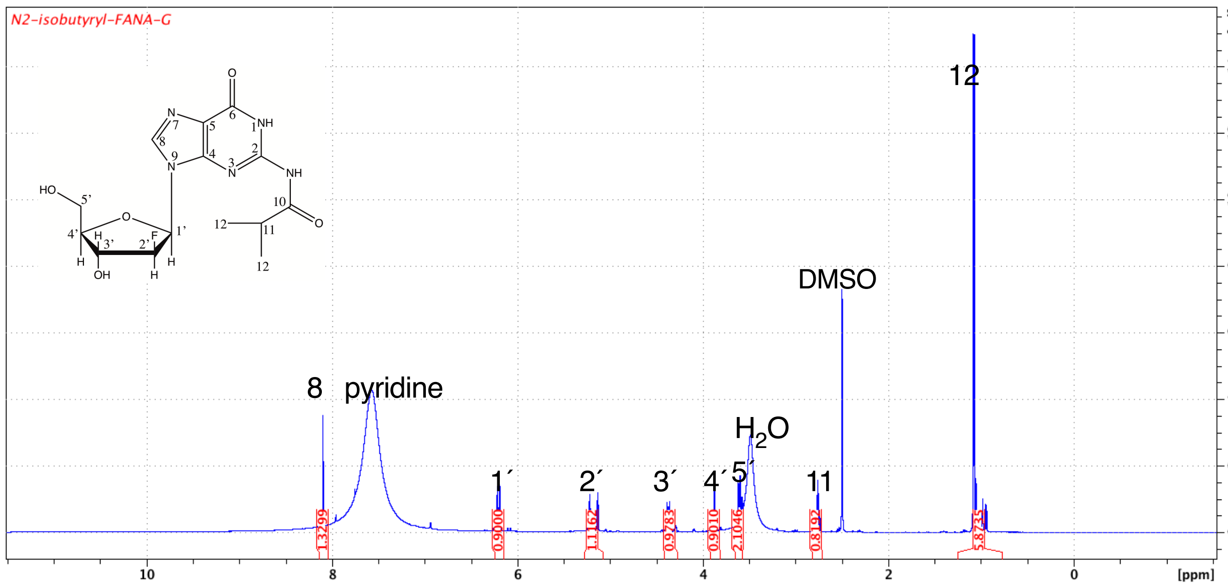
TABLE S6. **Observed and theoretical CID ions of the 5´-pTGATCAT-3´ extension product with hpol η**.

Scheme 1



SCHEME 1. **Synthesis of *N*[7]-CH_3 2´-F dG-containing oligonucleotide**. Reagents and conditions (all reactions were done at room temperature). (a) (i) chlorotrimethylsilane, pyridine, 2 h; (ii) isobutyryl chloride, 3 h, 80%; (b) CH_3I, *N*,*N*-dimethylformamide (DMF), 22 h, 80%; (c) 4,4´,4´´-trimethoxytrityl (DMT) chloride, *N*,*N*-diisopropylethylamine, pyridine, 2 h, 60 %; (d) 2-cyano-*N*,*N*,*N*,*N*-tetraisopropyl phosphane, tetrazole, CH_2Cl_2, 2 h, 68%; (e) solid phase DNA synthesis. (1, 2).

A



B

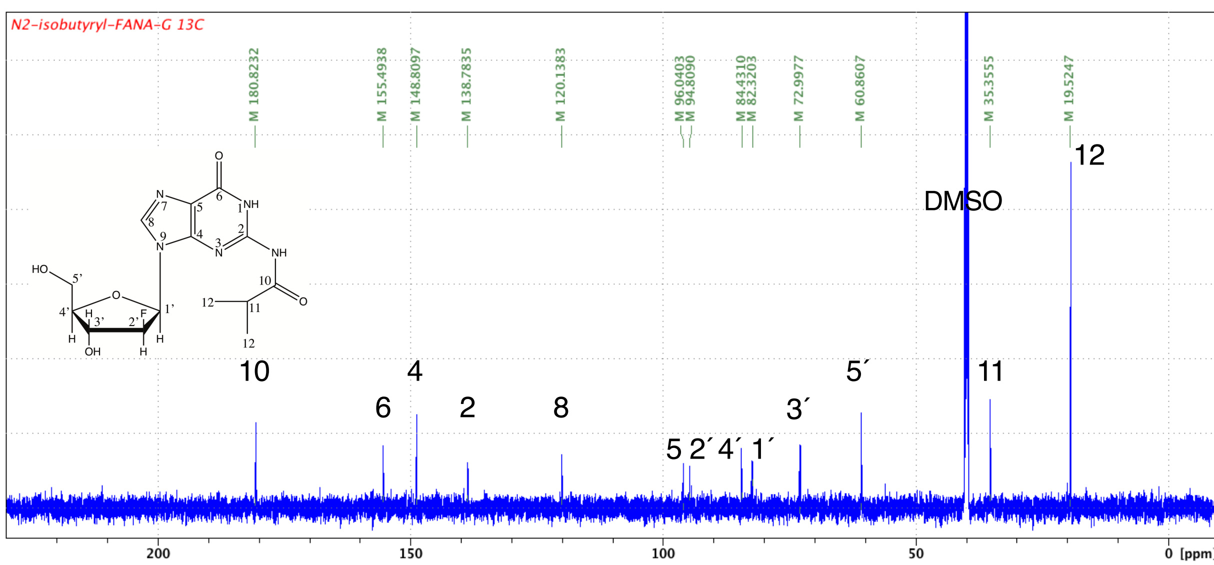


FIGURE S1. **NMR spectra for product of step 1 of synthesis of *N*^{7}-CH$_3$ 2´-F dG-containing oligonucleotide.** *A*, ^{1}H-spectrum; *B*, ^{13}C-spectrum.

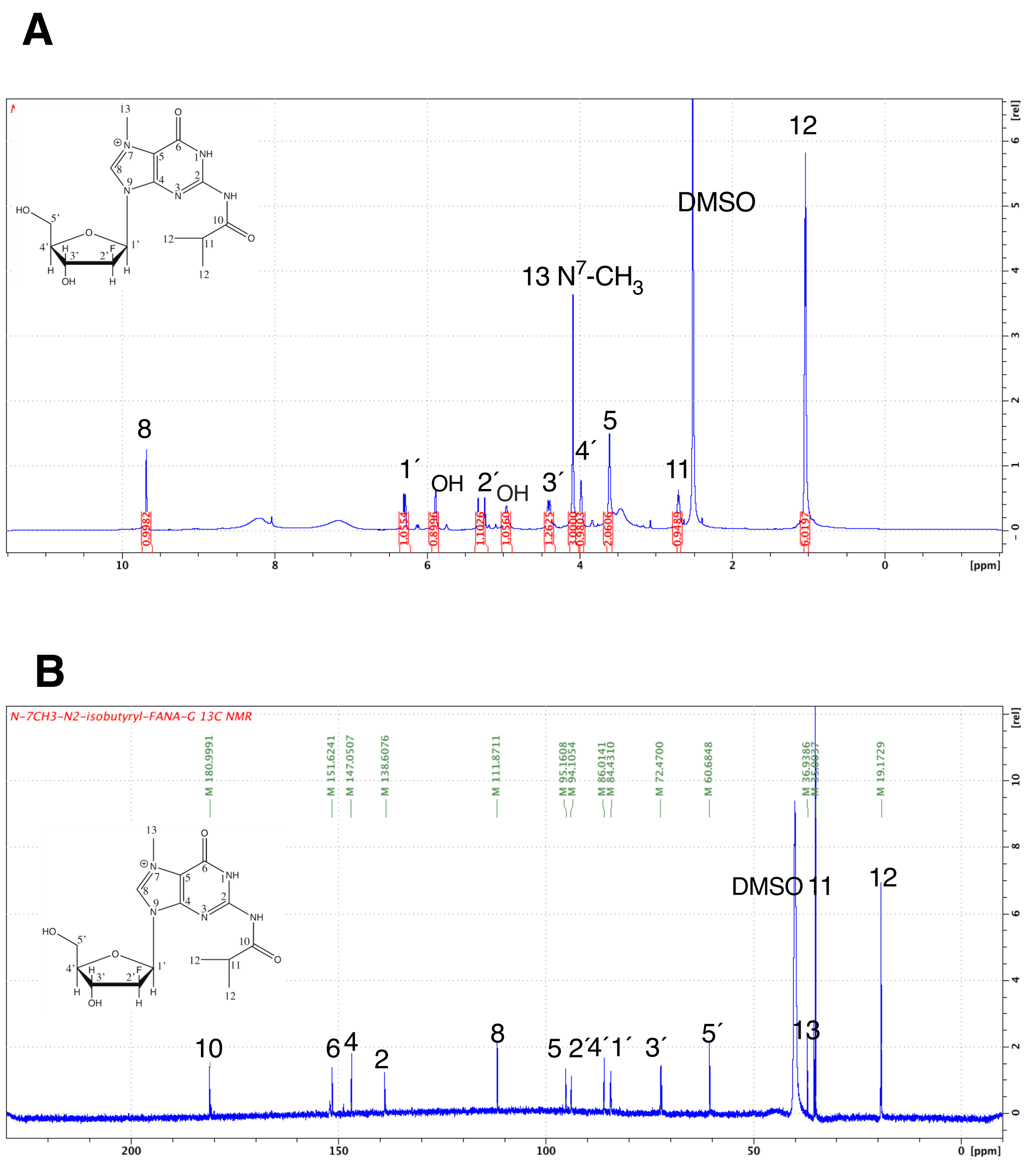


FIGURE S2. **NMR spectra for product of step 2 of synthesis of *N*^{7}-CH_3 2´-F dG-containing oligonucleotide.** *A*, ^{1}H-spectrum; *B*, ^{13}C-spectrum.

Fig. S3

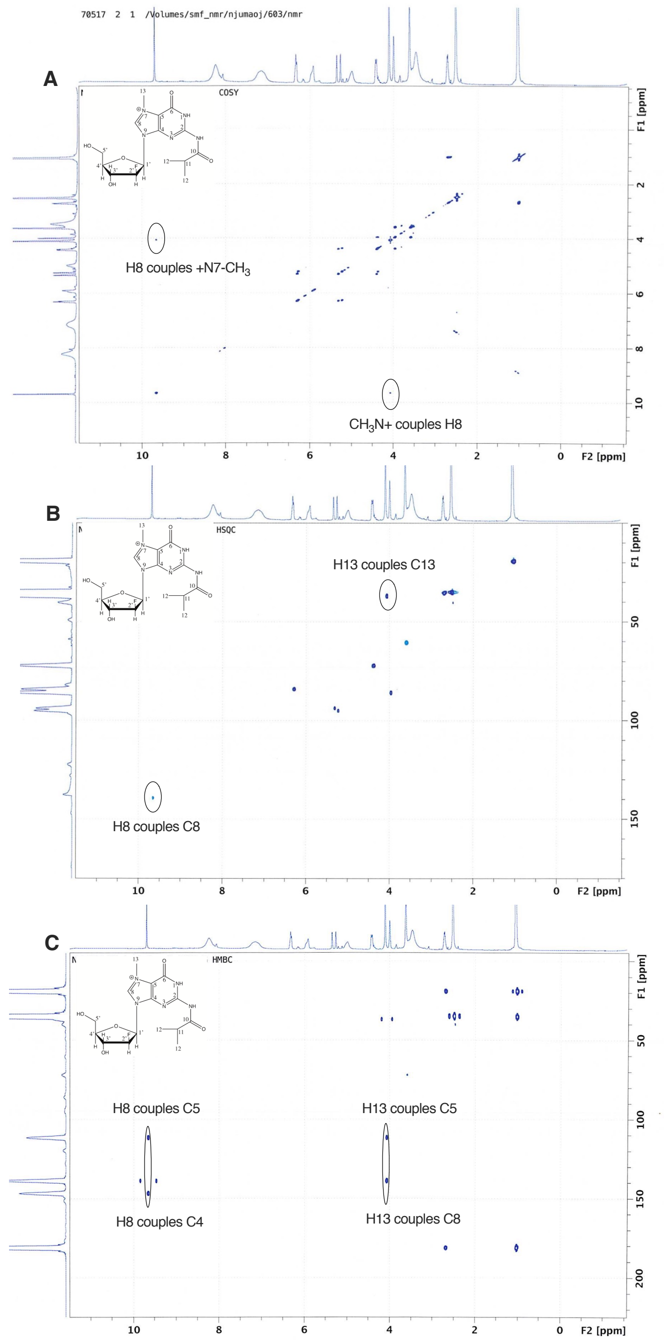


FIGURE S3. **COSY, HSQC, and HMBC NMR spectra for product of step 2 of synthesis of N^7-CH_3 2´-F dG-containing oligonucleotide**. *A*, COSY (homonuclear correlation spectroscopy); *B*, HSQC (heteronuclear single quantum coherence spectroscopy) ; *C*, HMBC (heteronuclear multiple bond correlation spectroscopy).

Fig. S4

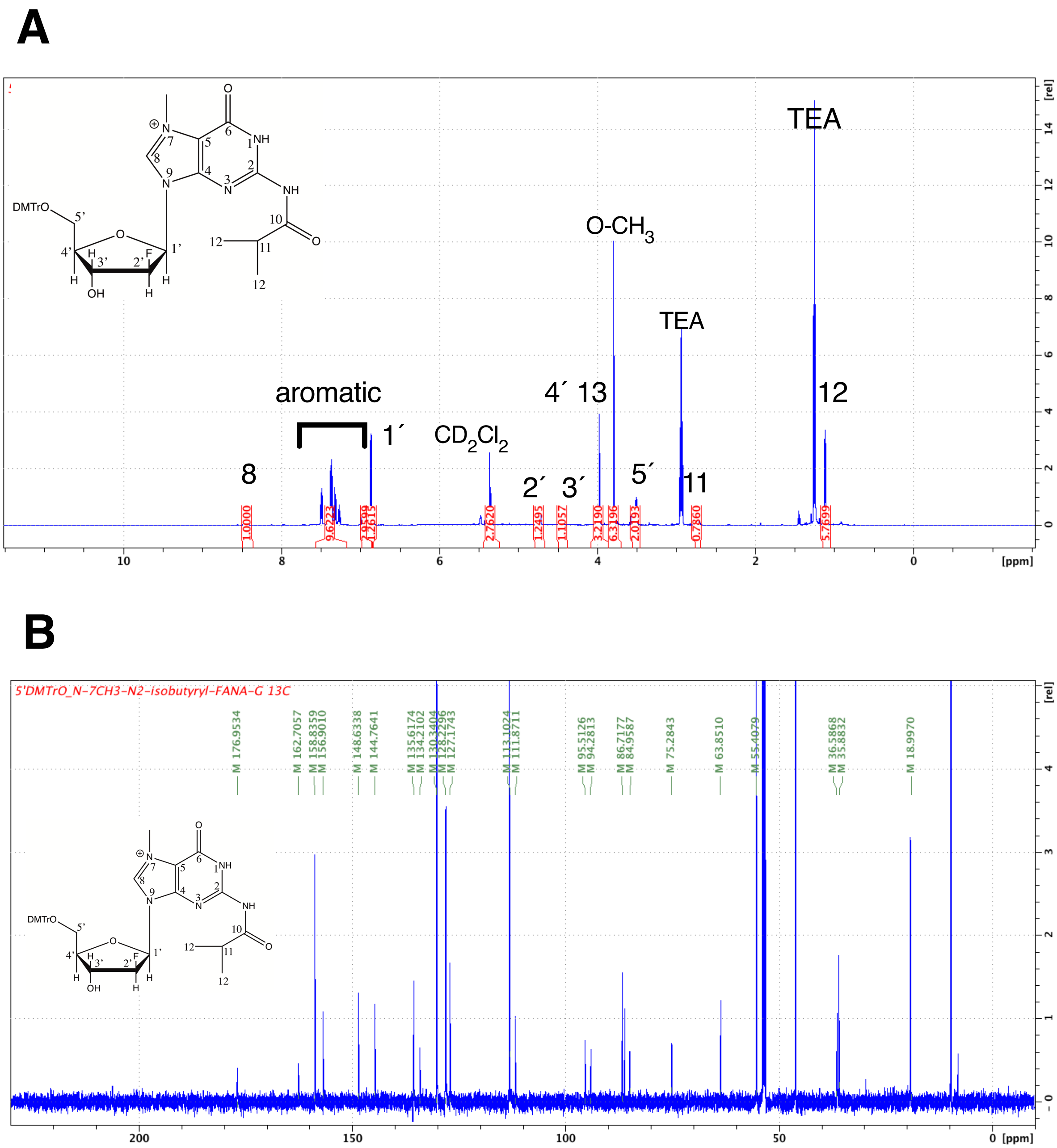
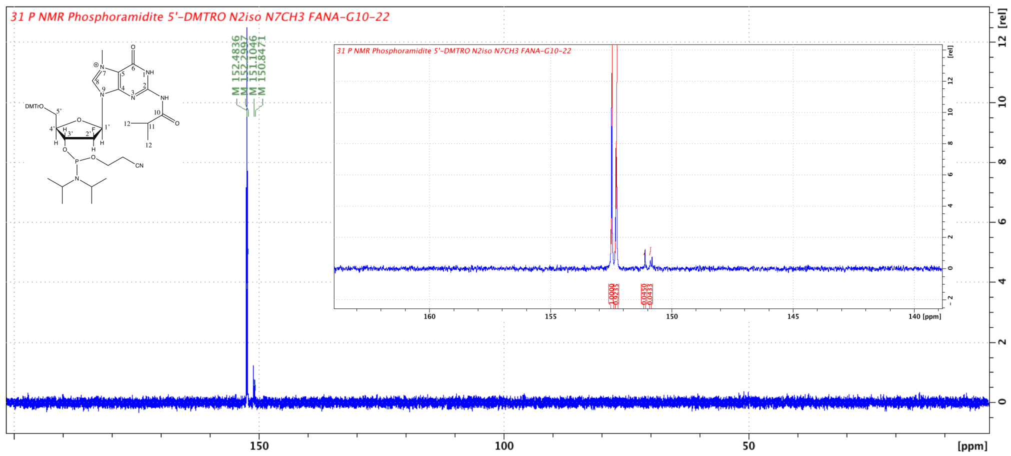


FIGURE S4. **NMR spectra for product of step 3 of synthesis of *N*^{7}-CH$_3$ 2´-F dG-containing oligonucleotide**. *A*, ^{1}H-spectrum; *B*, ^{13}C-spectrum.

Fig. S5

A



B

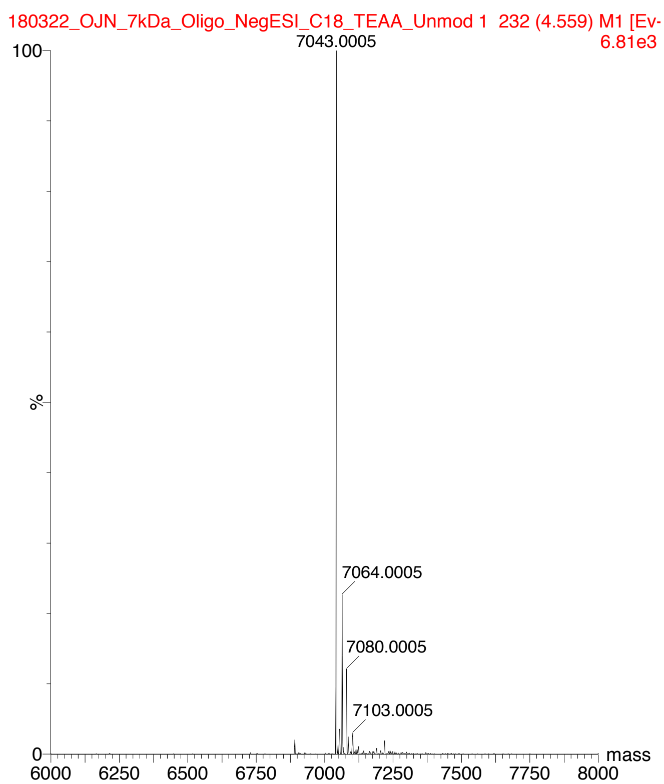


FIGURE S5: *A*, **^{31}P NMR spectrum for product of step 4 of synthesis of *N*^{7}-CH$_3$ 2´-F dG-containing oligonucleotide; *B*, LC-ESI-MS analysis of 2´-F dG-containing oligonucleotide.**

Fig. S6

A

B

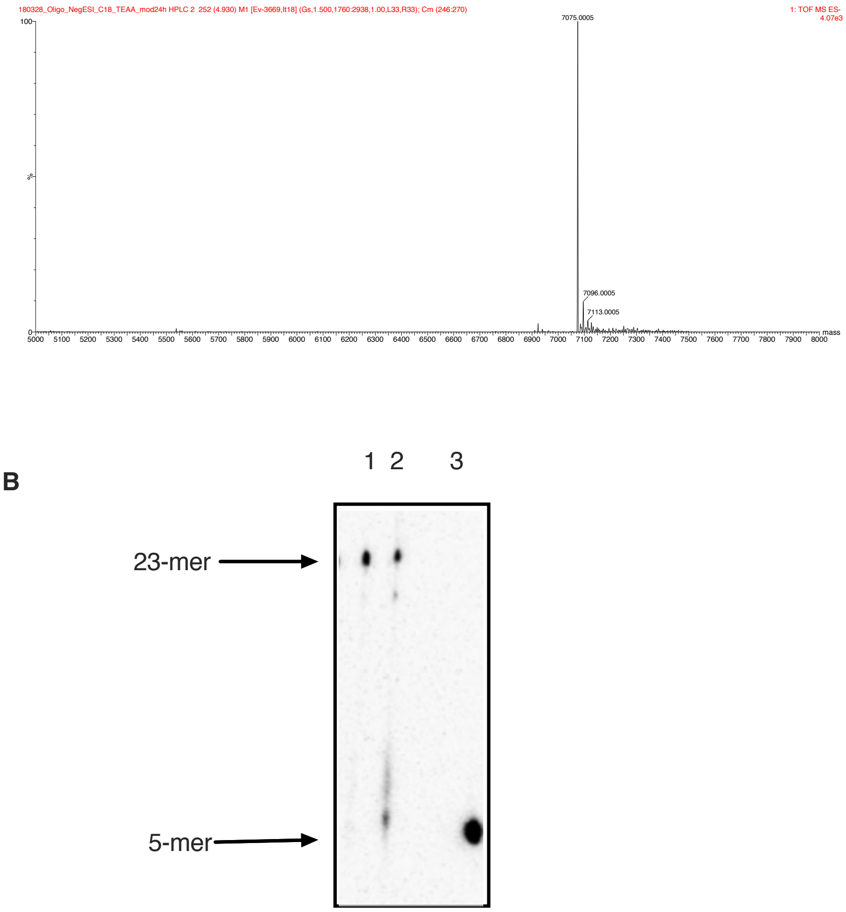


FIGURE S6: **Mass spectra and gel analysis of N^7-CH$_3$ 2´-FdG-containing oligonucleotide**. *A*, LC-ESI-MS analysis of N^7-CH$_3$ 2´-F dG-containing oligonucleotide; *B*, gel showing FPG glycosylase treatment of N^7-CH$_3$ 2´-F dG-containing oligonucleotide. Lane 1, N^7-CH$_3$ 2´-F dG-containing template, annealed, and treated with FPG glycosylase; lane 2, N^7-CH$_3$ 2´-F dG-containing template treated with NaOH, annealed, and treated with FPG glycosylase; lane 3, 5-mer standard (3´-GTACT-5´). The sequence used for the 23-mer was 3´-GCCCGAGCATTCGCAGTAXTACT-5´, where X is N^7-CH$_3$ 2´-F dG, which was annealed to its complementary sequence.

Fig. S7

5´-/FAM/CGGGCTCGTAAGCGTCAT
3´-GCCCGAGCATTCGCAGTAXTACT

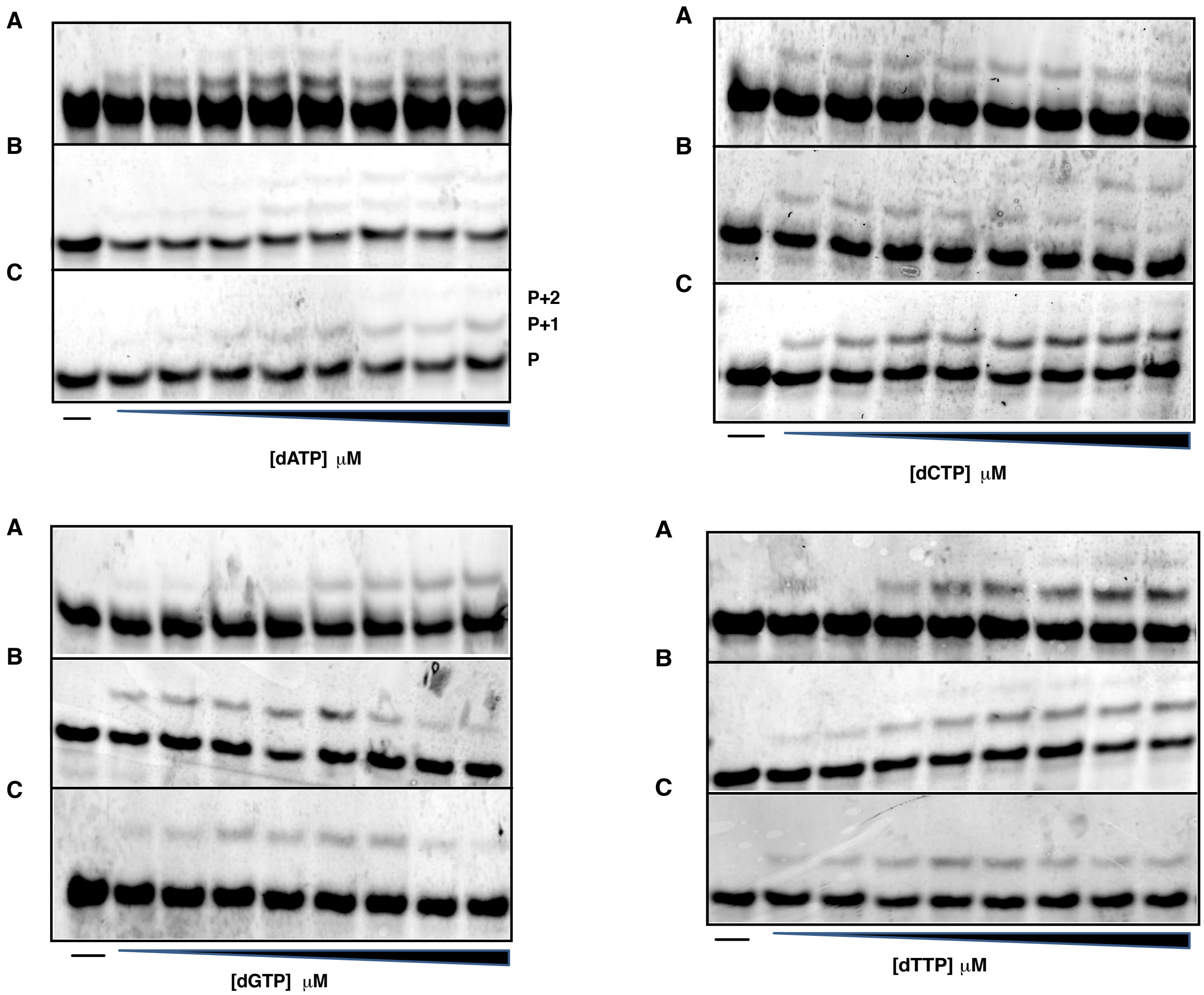


FIGURE S7. **Representative gels for steady-state kinetic analysis of single-nucleotide incorporation across dG, 2´-F dG, *N*⁷-CH₃ 2´-F dG by hpol η**. *A*, dG; *B*, 2´-F dG; and *C*, *N*⁷-CH₃ 2´-F dG. Reactions were done at 37 °C for 5-10 min by incubating 120 nM primer-template DNA complex, 2.5-10 nM hpol η, and varying concentrations of the indicated dNTP: ([dATP] 5-250 μM, [dCTP] 0.25-100 μM, [dGTP] 2.5-200 μM, and [dATP] 5-200 μM).

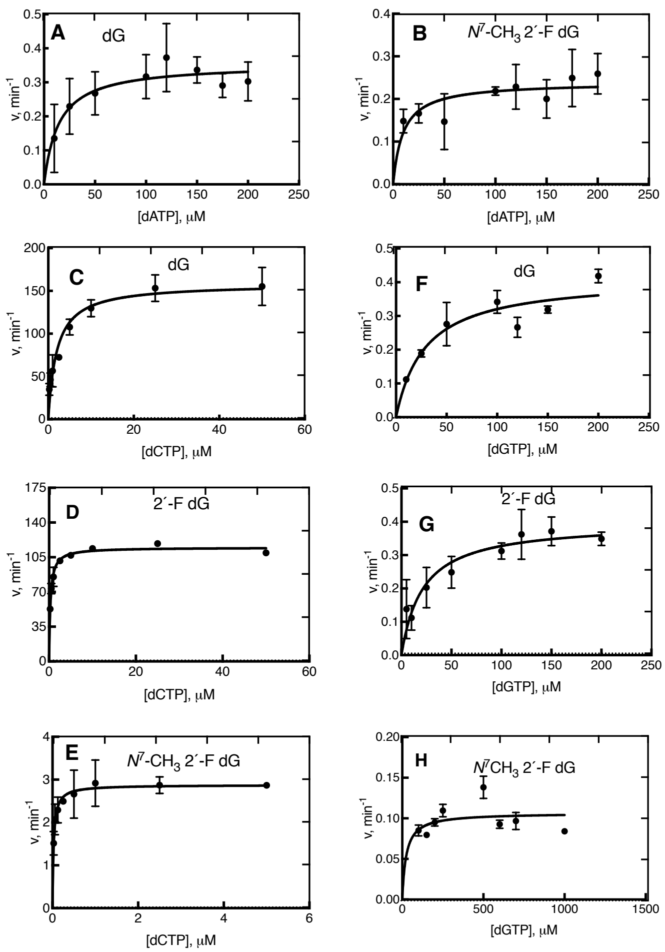


FIGURE S8. **Steady-state kinetic analysis of individual dATP, dCTP, and dGTP insertions by Dpo4**. Reactions contained templates dG (panels *A*, *C* and *F*), 2´-F dG (panels *D* and *G*), and N^7-CH_3 2´-F dG (panels *B*, *E*, and *H*) at position X in the sequences 5´-CGGGCTCGTAAGCGTCAT-3´ and 3´-GCCCGAGCATTCGCAGTAXTACT-5´. Reactions were done at 37 °C for 5-10 min by incubating 120 nM primer-template oligonucleotide complex, 0.15-10 nM Dpo4, and varying concentrations of dATP, dCTP and dGTP. Fitting was to a hyperbolic equation in GraphPad Prism v.8.0, CA, and k_{cat} and K_m values are presented in Table1.

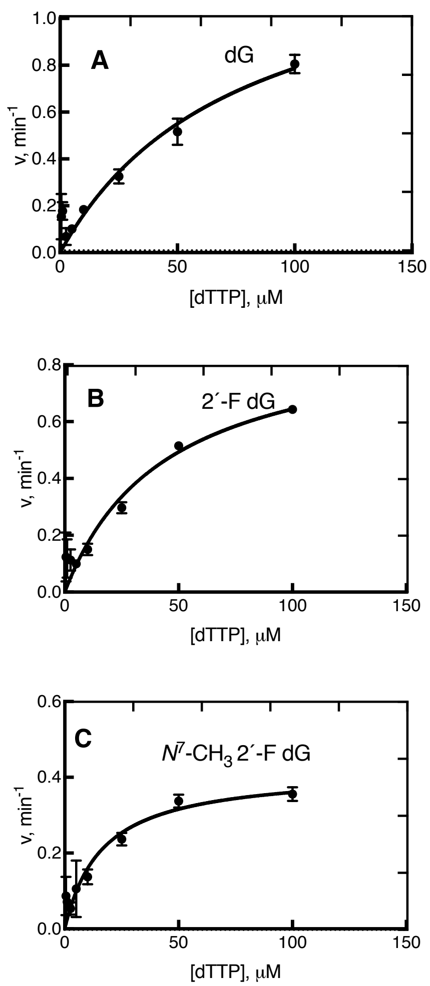


FIGURE S9. **Steady-state kinetic analysis of individual dTTP insertions by Dpo4**. Reactions contained templates dG (panel *A*), 2´-F dG (panel *B*), and N^7-CH_3 2´-F dG (panel *C*) at position X in the sequences 5´-CGGGCTCGTAAGCGTCAT-3´ and 3´-GCCCGAGCATTCGCAGTAXTACT-5´. Reactions were done at 37 °C for 3-10 min by incubating 120 nM primer-template oligonucleotide complex, 0.15-10 nM Dpo4, and varying concentrations of dATP, dCTP and dGTP. Fitting was to a hyperbolic equation in GraphPad Prism v.8.0, CA, and k_{cat} and K_m values are presented in Table 2.

Fig. S10

```
5´-/FAM/CGGGCTCGTAAGCGTCATC
      3´-GCCCGAGCATTCGCAGTAXTACT

5´-/FAM/CGGGCTCGTAAGCGTCATT
      3´-GCCCGAGCATTCGCAGTAXTACT
```

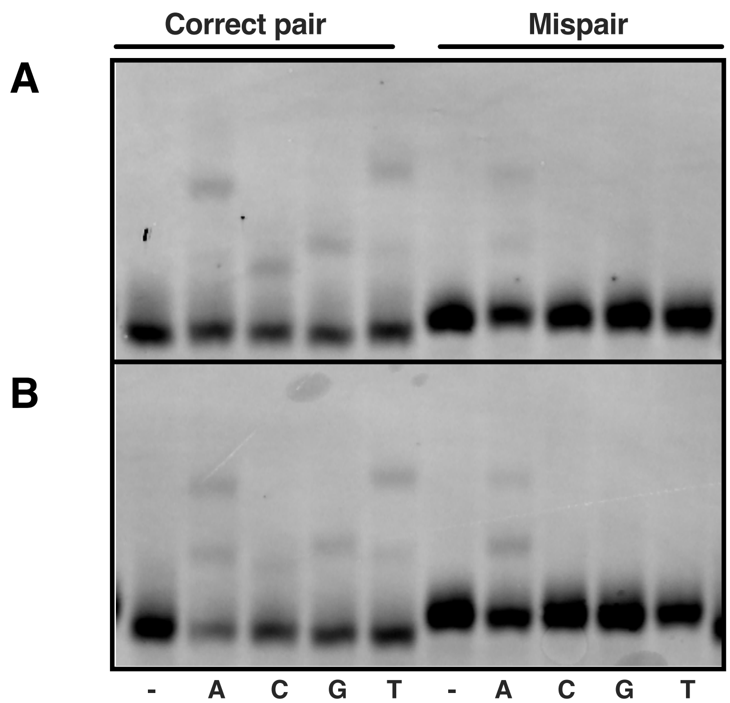


FIGURE S10. **Gel analysis for single nucleotide extension from 2´-F dG:dC and 2´-F dG:dT base-pairs and from *N*^{7}-CH$_3$ 2´-F dG:dC and *N*^{7}-CH$_3$ 2´-F dG:dT base-pairs**. *A*, 2´-F dG:dC and 2´-F dG:dT; *B*, *N*^{7}-CH$_3$ 2´-F dG:dC and *N*^{7}-CH$_3$ 2´-F dG:dT. Reactions were conducted with 120 nM FAM-labeled primer-template DNA complex, 250 μM dNTPs, and 5 nM hpol η for 5 minutes.

Fig. S11

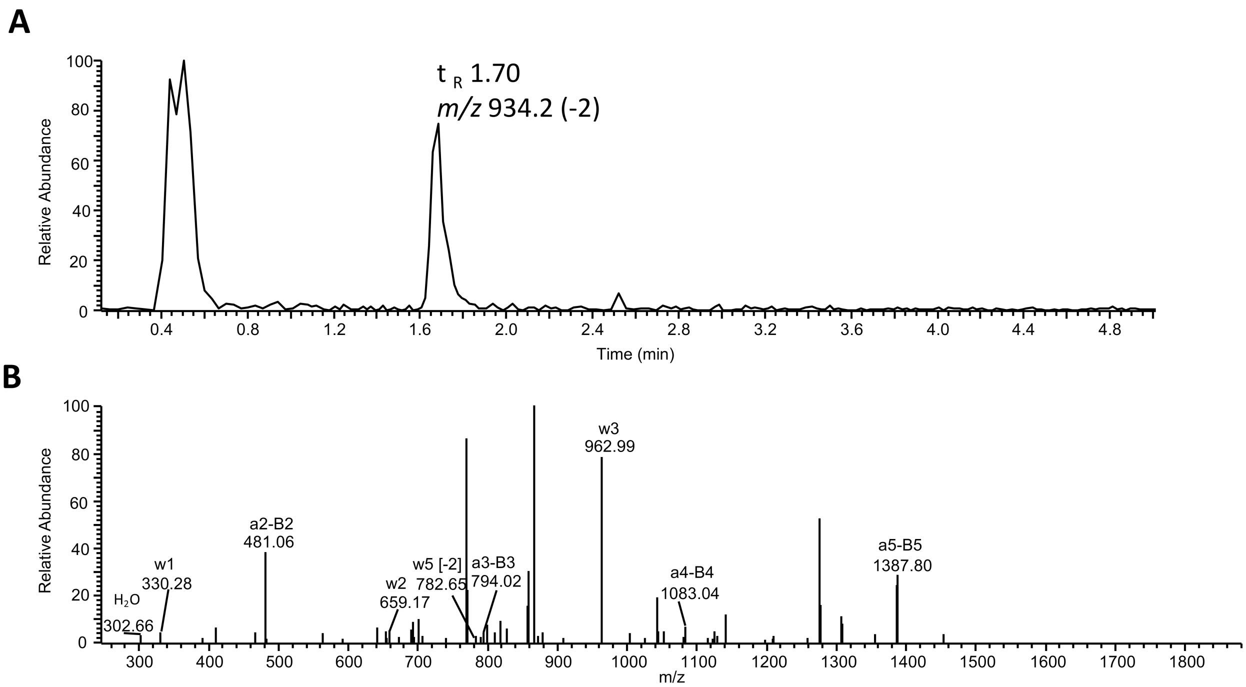
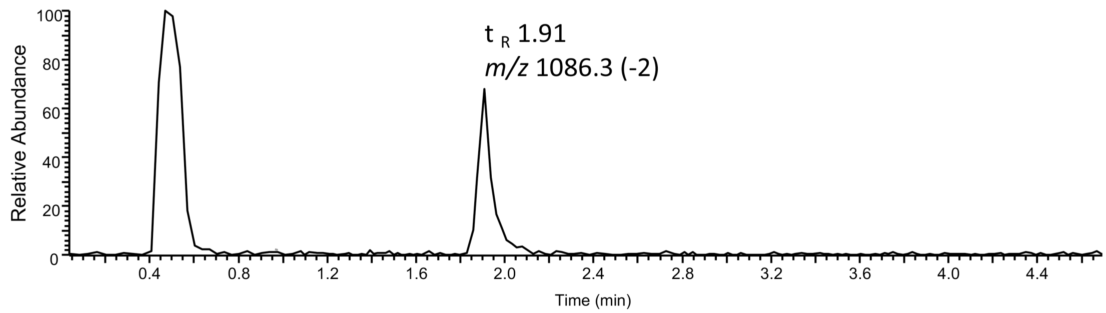


FIGURE S11. **LC-MS sequence analysis of extension products opposite *N*^{7}-CH_3 2´-F dG with hpol η**. *A*, extracted ion chromatograms of product ions at *m*/*z* 934.2 (-2) (misincorporation of A); *B*, CID spectrum of *m*/*z* 934.

Fig. S12

A



B

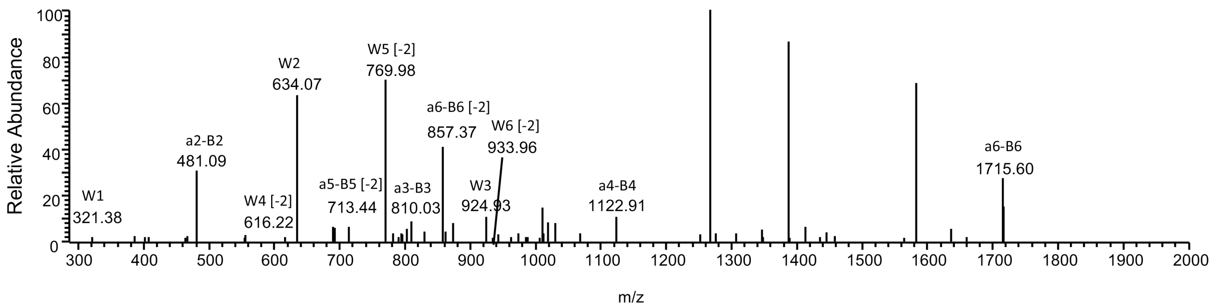


FIGURE S12. **LC-MS sequence analysis of extension products opposite *N*[7]-CH_3 2´-F dG with hpol η**. *A*, extracted ion chromatogram of product ions at *m/z* 1086.3 (-2) (misincorporation of G); *B*, CID spectrum of *m/z* 1086.3.

TABLE S1. **Observed and theoretical CID ions of the 5´-pTCATGA-3´ extension product with hpol η (*m/z* 934.3)**

Fragmentation assignment	*m/z*, observed (charge)	*m/z*, theoretical (charge)
5´-pT (a_2-B_2)	481.06	481.27
5´-pTC (a_3-B_3)	769.99	770.46
5´-pTCA (a_4-B_4)	1083.04	1083.67
5´-pTCAT (a_5-B_5)	1387.80	1387.87
pCATGA-3´ (w_5,-2)	782.65 (2)	782.51 (2)
pATGA-3´ (w_4)	1275.94	1276.83
pTGA-3´ (w_3)	962.99	963.62
pGA-3´ (w_2)	659.17	659.43
pA-3´ (w_1)	330.28	330.22

TABLE S2. **Observed and theoretical CID ions of the 5´-pTACTGA-3´ extension product with hpol η (*m*/*z* 934.3)**

Fragmentation assignment	*m*/*z*, observed	*m*/*z*, theoretical
5´-pT (a_2-B_2)	481.06	481.27
5´-pTA (a_3-B_3)	794.02	794.48
5´-pTAC (a_4-B_4)	1083.04	1083.67
5´-pTACT (a_5-B_5)	1387.80	1387.87
pACTGA-3´ (w_5,-2)	782.65 (2)	782.51 (2)
pTGA-3´ (w_3)	962.99	963.62
pGA-3´ (w_2)	659.17	659.43
pA-3´ (w_1)	330.28	330.22

TABLE S3. **Observed and theoretical CID ions of the 5´-pTACT-3´ extension stalled product with hpol η (*m*/*z* 613.2).**

Fragmentation assignment	*m*/*z*, observed	*m*/*z*, theoretical
5´-pT (a_2-B_2)	481.06	481.27
5´-pTC (a_3-B_3)	770.0	770.46
pCAT-3´ (w_3)	923.90	923.60
pAT-3´ (w_2)	634.23	634.41
pT-3´ (w_1)	321.07	321.20

TABLE S4. **Observed and theoretical CID ions of the 5´-pTCATGAT-3´ extension product with hpol η (*m/z* 1086.2).**

Fragmentation assignment	*m/z*, observed (charge)	*m/z*, theoretical (charge)
5´-pT (a_2-B_2)	481.20	481.27
5´-pTC (a_3-B_3)	770.13	770.45
5´-pTCAT (a_5-B_5)	1387.77	1387.87
5´-pTCATG (a_6-B_6)	1716.94	1717.08
pCATGAT-3´ (w_6,-2)	933.96 (-2)	934.6 (-2)
pATGAT-3´ (w_5,-2)	1579.94	1581.03
pTGAT-3´ (w_4)	1267.75	1267.82
pGAT-3´ (w_3)	962.91	963.62
pAT-3´ (w_2)	634.07	634.41
pT-3´ (w_1)	321.38	321.20

TABLE S5. Observed and theoretical CID ions of the 5´-pTAGTCAT-3´ extension product with hpol η (*m/z* 1086.2).

Fragmentation assignment	*m/z*, observed (charge)	*m/z*, theoretical (charge)
5´-pT (a_2-B_2)	481.20	481.27
5´-pTA (a_3-B_3)	794.22	794.48
5´-pTAG (a_4-B_4)	1122.91	1123.69
5´-pTAGT (a_5-B_5)	713.44 (-2)	713.44 (-2)
5´-pTAGTC (a_6-B_6)	1716.94, 858.18 (-2)	1717.08, 858.03 (-2)
pTAGTA-3´ (w_6,-2)	933.96 (-2)	934.6 (-2)
pTCAT-3´ (w_4, -2)	616.22 (-2)	613.39 (-2)
pCAT-3´ (w_3)	924.93	923.60
pAT-3´ (w_2)	634.07	634.41
pT-3´ (w_1)	321.38	321.20

TABLE S6. Observed and theoretical CID ions of the 5´-pTGATCAT-3´ extension product by hpol η (*m/z* 1086.2).

Fragmentation assignment	*m/z*, observed (charge)	*m/z*, theoretical (charge)
5´-pT (a_2-B_2)	481.20	481.27
5´-pTG (a_3-B_3)	810.03	810.48
5´-pTGA (a_4-B_4)	1122.91	1123.69
5´-pTGAT (a_5-B_5)	713.44 (-2)	713.44 (-2)
5´-pTGATC (a_6-B_6)	1716.94, 858.18 (-2)	1717.08, 858.03 (-2)
pGATCAT-3´ (w_6, -2)	933.96 (-2)	934.6 (-2)
pATCAT-3´ (w_5, -2)	769.88 (-2)	769.99 (-2)
pTCAT-3´ (w_4,-2)	616.22 (-2)	613.39 (-2)
pCAT-3´ (w_3)	924.93	923.60
pAT-3´ (w_2)	634.07	634.41
pT-3´ (w_1)	321.38	321.20

References

1. Wilson, T. J., Li, N. S., Lu, J., Frederiksen, J. K., Piccirilli, J. A., and Lilley, D. M. (2010) Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11751-11756

2. Lee, S., Bowman, B. R., Ueno, Y., Wang, S., and Verdine, G. L. (2008) Synthesis and structure of duplex DNA containing the genotoxic nucleobase lesion N^7-methylguanine. *J. Am Chem. Soc.* **130**, 11570-11571