Supporting Information

Backbone Flexibility Influences Nucleotide Incorporation by Human Translesion DNA Polymerase *η* Opposite Intrastrand Cross-linked DNA

Derek K. O'Flaherty[§], F. Peter Guengerich[‡], Martin Egli[‡]*, Christopher J. Wilds[§]*

[§] Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke St. West, Montréal, Québec (Canada) H4B 1R6

[‡] Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee (USA) 37232-0146

Contents **Page**

Supporting Methods

Supporting Tables and Figures

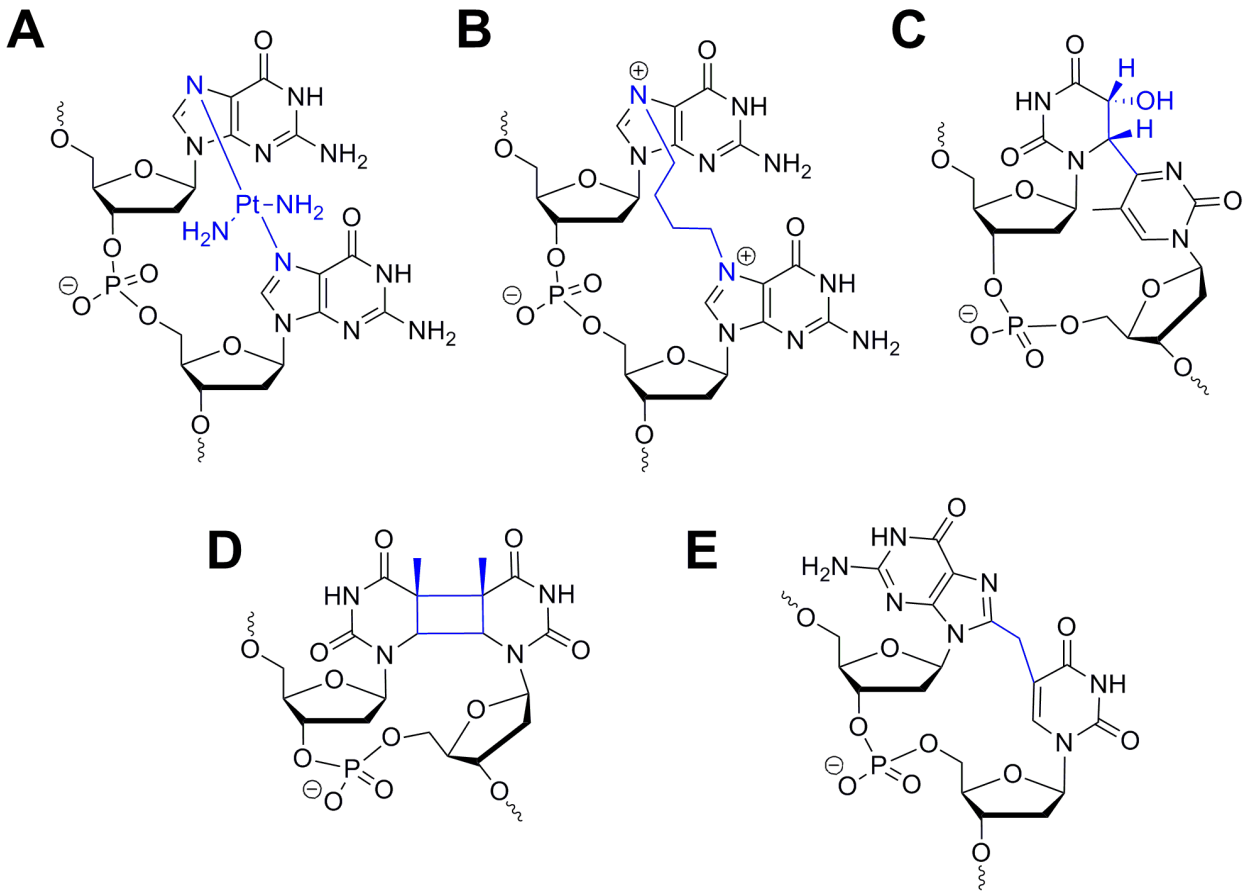
Supporting Methods

Chemical Synthesis and Characterization of Modified DNA Oligonucleotides

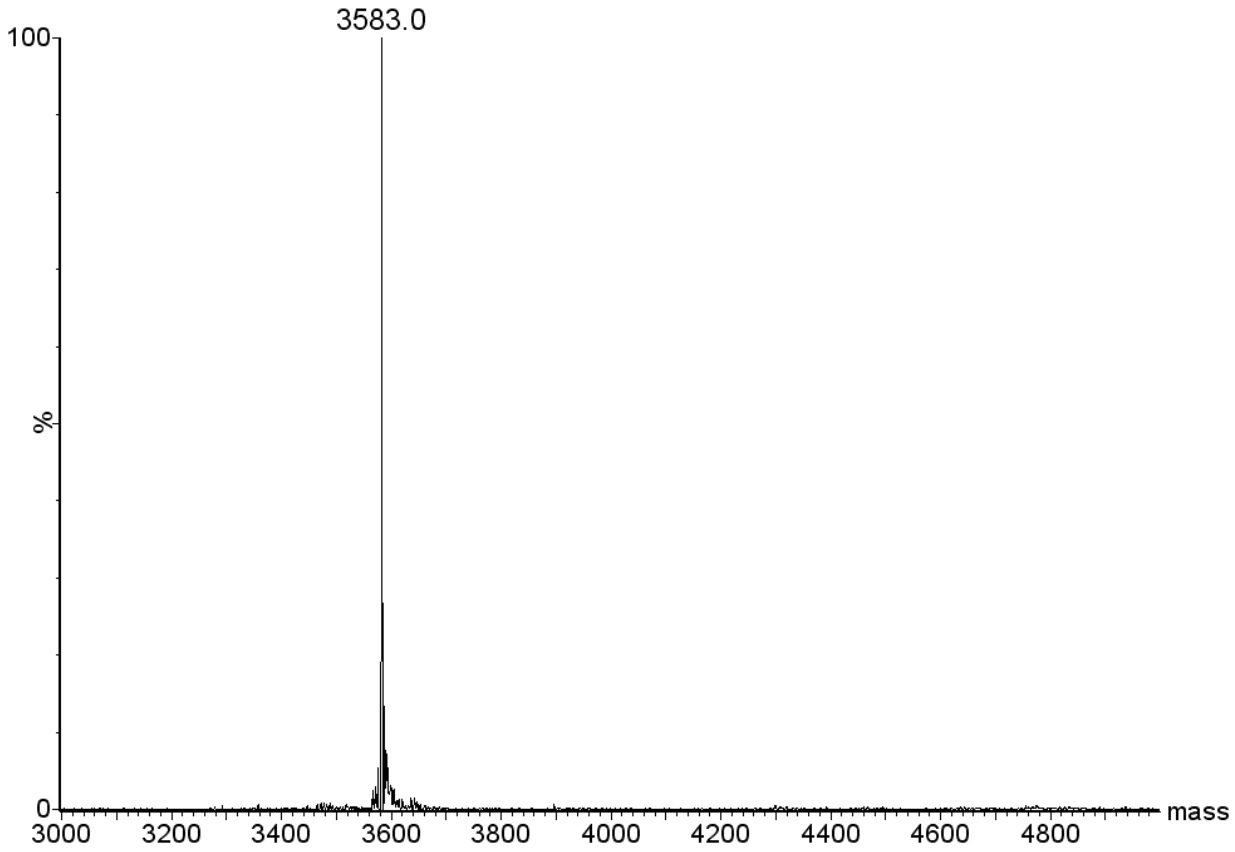
The chemical synthesis of the dimer phosphoramidites required to synthesize **GG4** and **GG7** have been described in previously published work.[1] Those required for **GpG4**, and **GpG7** are described in a manuscript submitted for review with synthesizes adapted from previously described procedures.[2] The cross-linked duplexes, whose sequences are 5´- AC **XX** CT CAC ACT (where **XX** denotes the cross linked dGG residues or dGG for the unmodified control), were assembled on an Applied Biosystems Model 3400 synthesizer on a 1.5 μmol scale using standard β-cyanoethylphosphoramidite chemistry supplied by the manufacturer with slight modifications to coupling times. The nucleoside phosphoramidites containing fast deprotecting groups were prepared in anhydrous CH_3CN at a concentration of 0.1 M for the 3´-*O*-deoxyphosphoramidites, and 0.15 M for the cross-linked 3´-*O*-deoxyphosphoramidite. Oligomer sequence assembly was carried out according to published procedures by our group.[1]

Protecting groups and cleavage from the solid support was carried out by treatment with aqueous (28% v/v) NH_4OH in C_2H_5OH (1 mL of a 3 : 1 v/v solution) for 4h at 55 ^{o}C in 2 mL screw cap microfuge tubes fitted with Teflon lined caps for **GG4** and **GG7**. Modified oligomers **GpG4** and **GpG7** were first deprotected for 16 h in aqueous (28% v/v) ammonium hydroxide at room temperature with gentle rocking, followed by an additional 4 h at 55 ^{o}C. Crude oligomers were transferred and the solvent removed using a Savant SC110A SpeedVac Concentrator (Thermo) followed by purification by strong-anion exchange HPLC with a Dionex DNAPAC PA-100 column (0.4 cm × 25 cm) purchased from Dionex, Sunnyvale, CA using a linear gradient of 0-52% buffer B (v/v) over 24 min (buffer A: 100 mM Tris HCl, pH 7.5, 10% MeCN and buffer B: 100mMTris HCl, pH 7.5, 10% CH_3CN, 1 M NaCl) at 55 ^{o}C. The columns were monitored at 260 nm for analytical runs or 280 nm for preparative runs. The purified oligomers were desalted using C-18 SEP PAK cartridges (Waters) as previously described.[3] ESI mass spectra for oligonucleotides were obtained at the Concordia University Centre for Biological Applications of Mass Spectrometry using a Micromass Qtof2 mass spectrometer (Waters) equipped with a nanospray ion source. The mass spectrometer was operated in full scan, negative ion detection mode. The molecular mass of the modified oligomers were identifed by ESI-MS and the measured values were in agreement with the expected masses (see Supporting Figures S2-S5 for MS spectra).

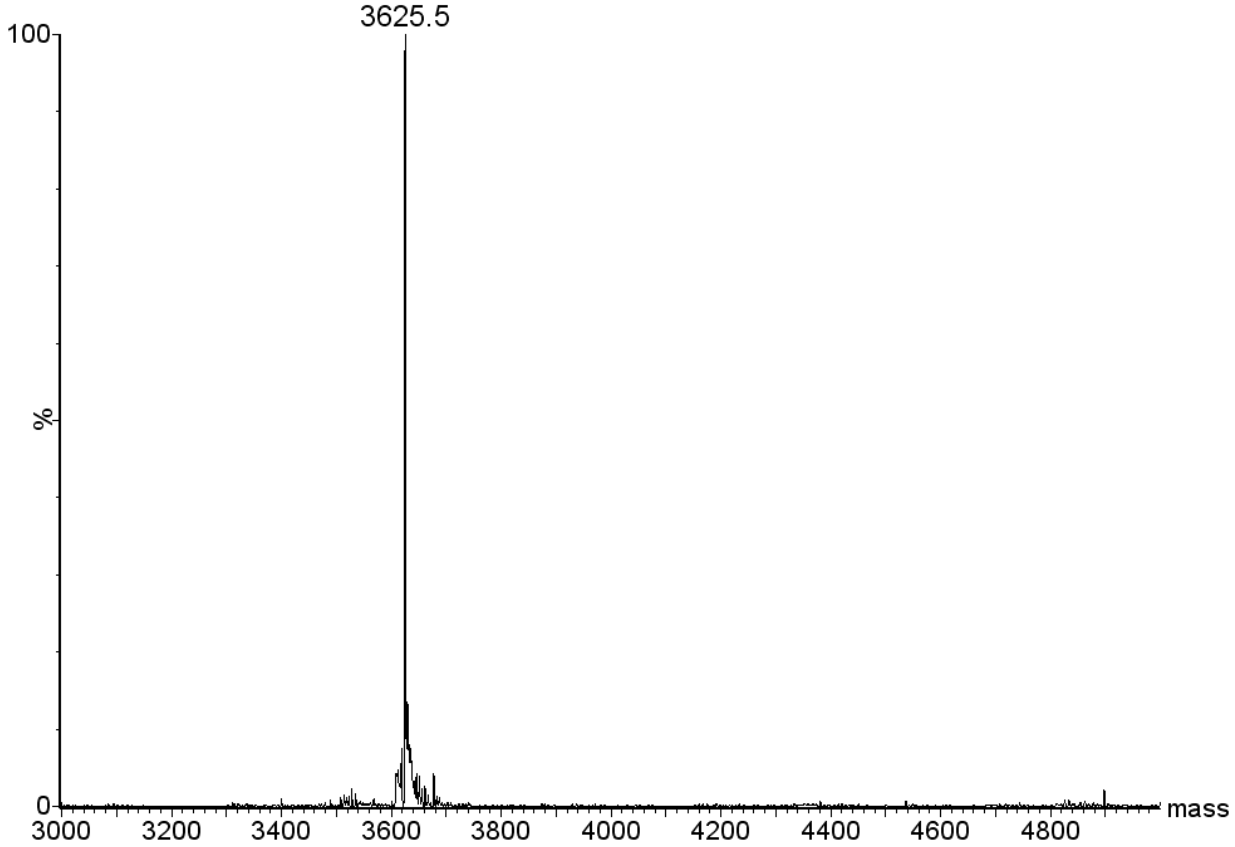
Supporting Figure S1: Examples of different IaCL DNA adducts mentioned in the main text. **A** An example of an IaCL formed from exposure of DNA to the chemotherapeutic drug cisplatin.[4,5] **B** An example of an IaCL formed from exposure of DNA to the chemotherapeutic drug busulfan.[6] **C** An example of a UV-induced (6-4) pyrimidone photoproduct lesion.[7] **D** An example of a UV-induced cyclobutane pyrimidine dimer. [8] **E** The structure of dG[8,5-Me]dT IaCL.[9] Modifications are highlighted in blue.



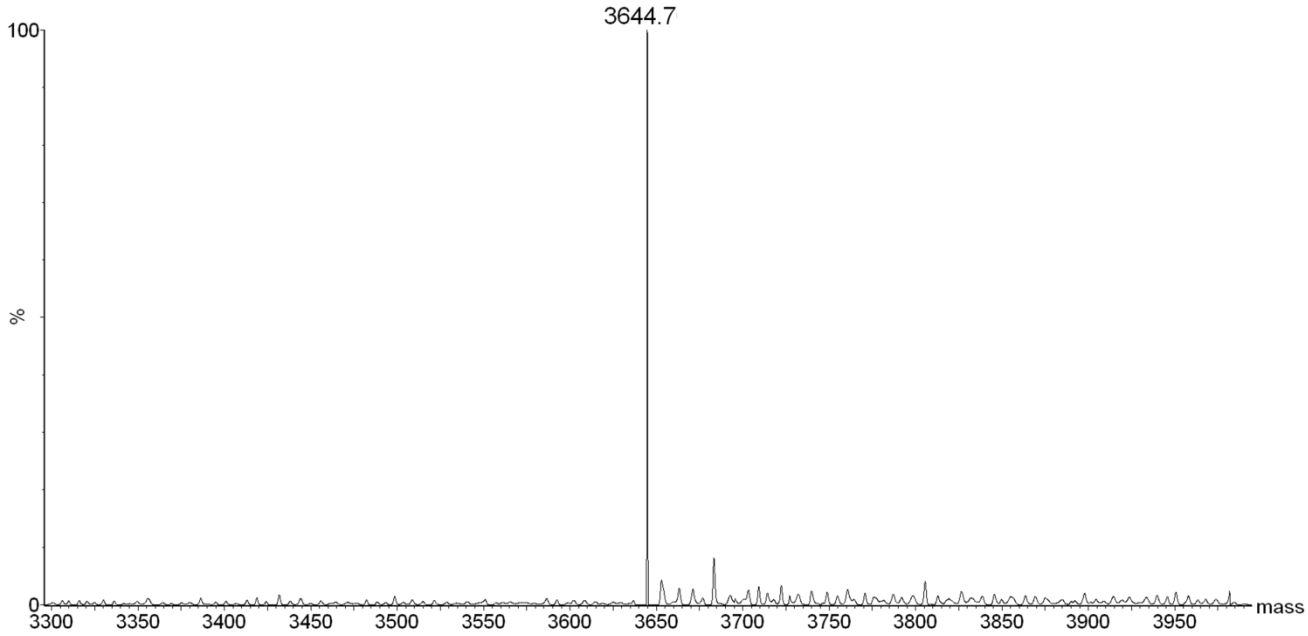
Supporting Figure S2 - ESI MS spectrum of oligonucleotide **GG4** (expected mass of 3582.5).



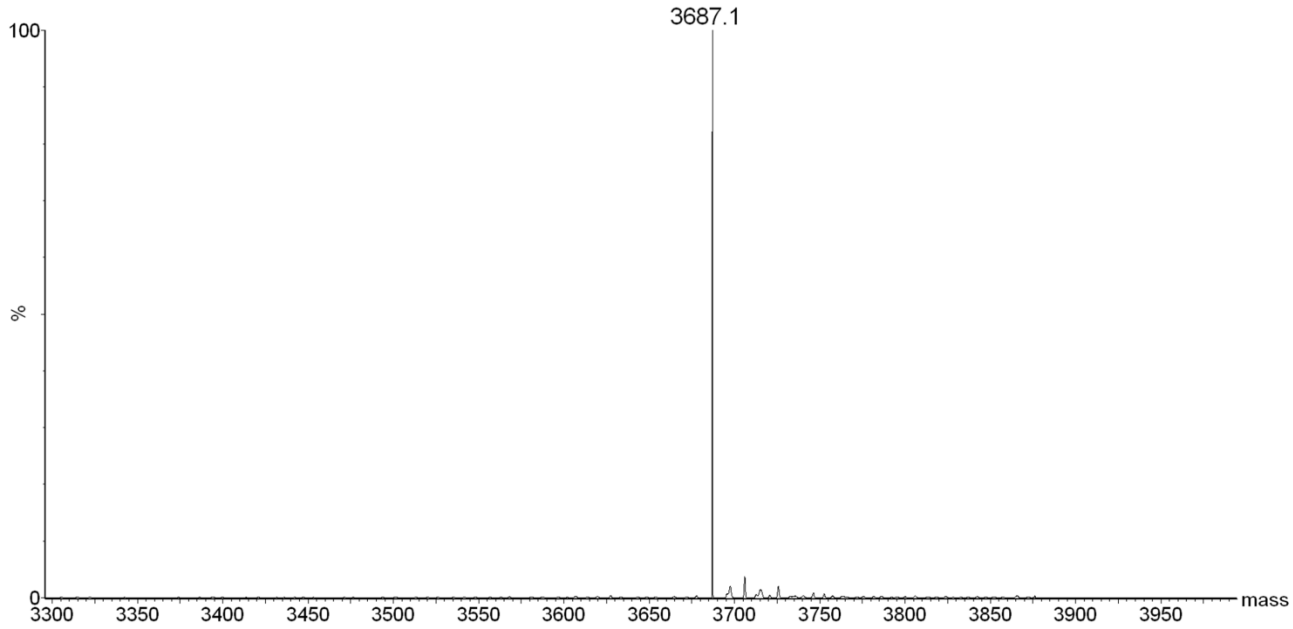
Supporting Figure S3 - ESI MS spectrum of oligonucleotide **GG7** (expected mass of 3624.6).



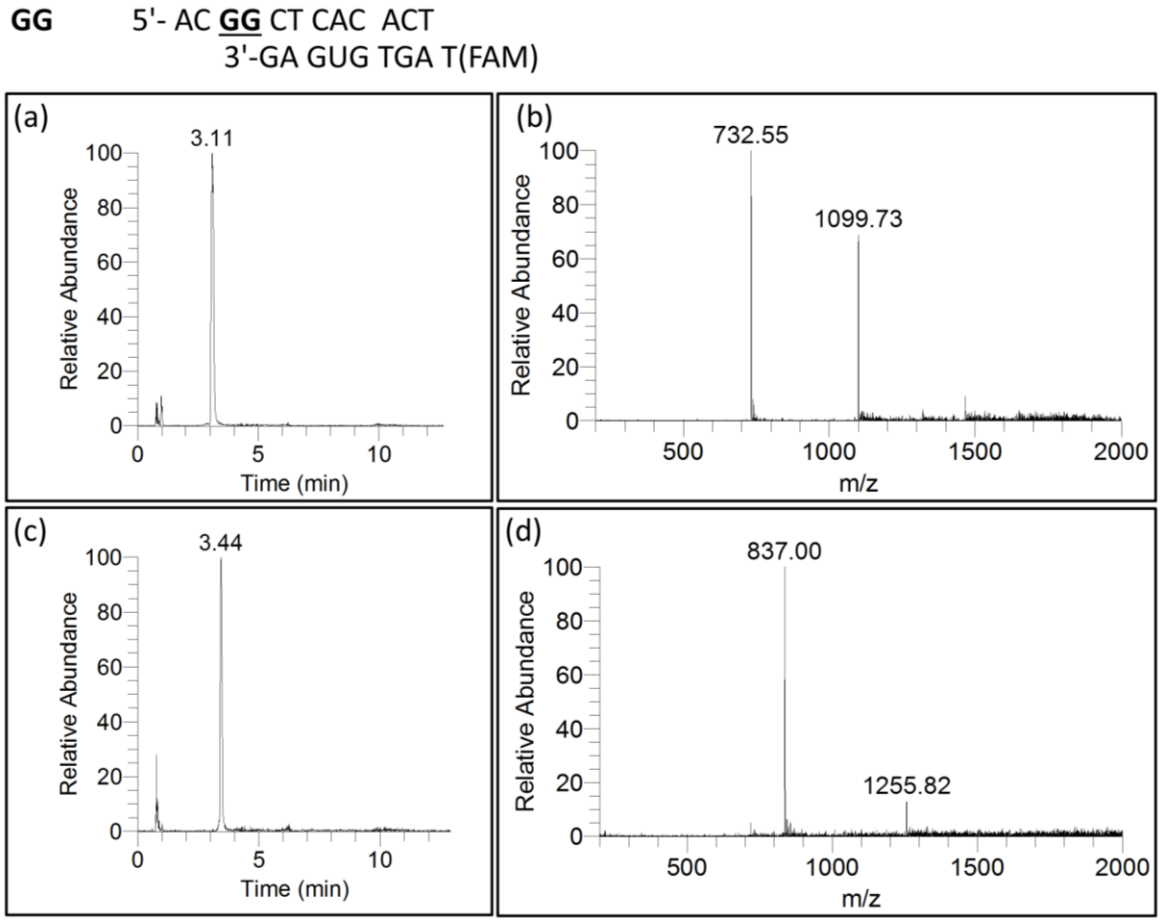
Supporting Figure S4 - ESI MS spectrum of oligonucleotide **GpG4** (expected mass of 3644.5).



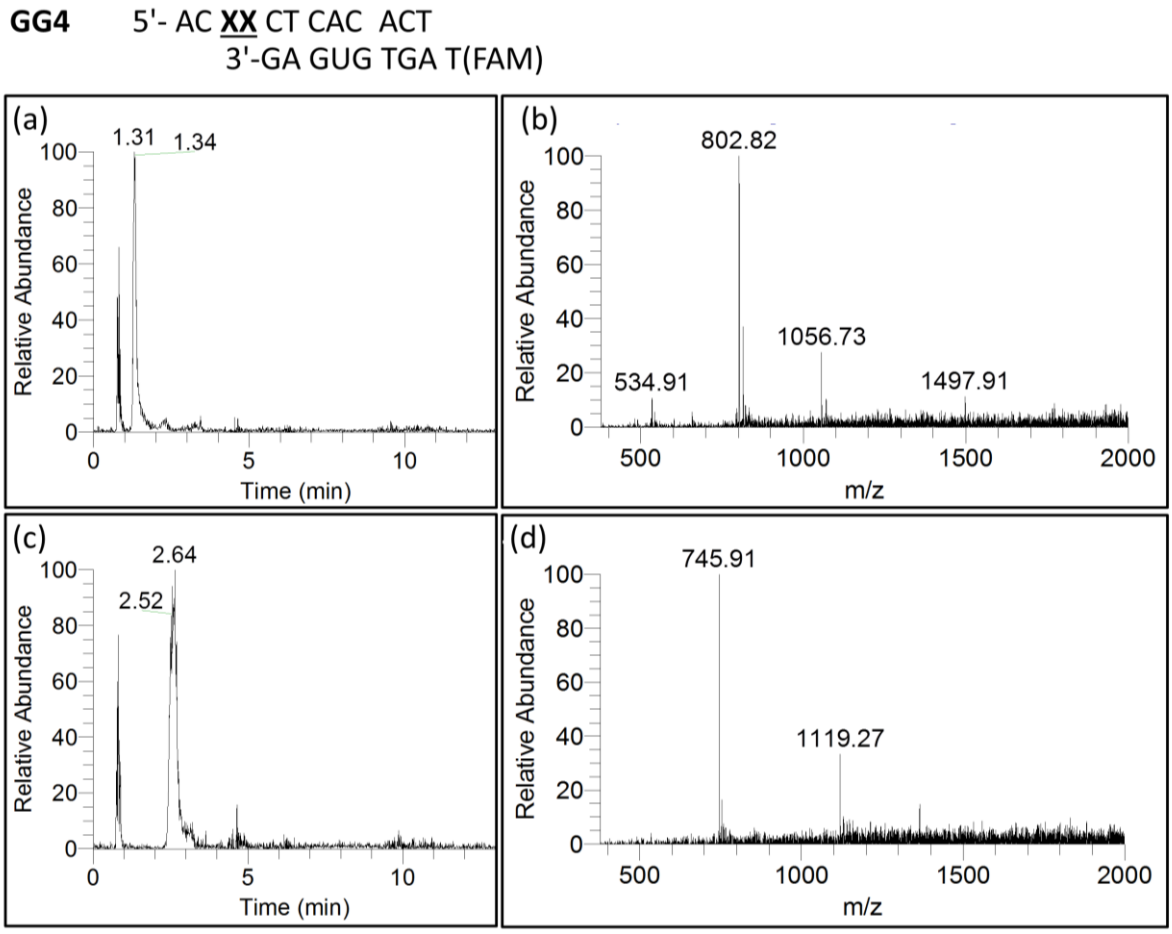
Supporting Figure S5 - ESI MS spectrum of oligonucleotide **GpG7** (expected mass of 3686.6).



Supporting Figure S6 - LC-MS analysis of most abundant full-length extension products opposite **GG** in DNA template by hPol *η* in the presence of all four dNTPs. (a) Sample reconstructed extracted ion chromatogram for *m/z* 732.80 for product with sequence 5´-pGAGCCGT and (b) mass spectrum of peak at retention time 3.11 min. (c) Sample reconstructed extracted ion chromatogram for *m/z* 837.20 for product with sequence 5´-pGAGCCGTA and (d) mass spectrum of peak at retention time 3.44 min. See Supporting Table S4 for full list of products and respective *m/z* assignments. See above for methodology.

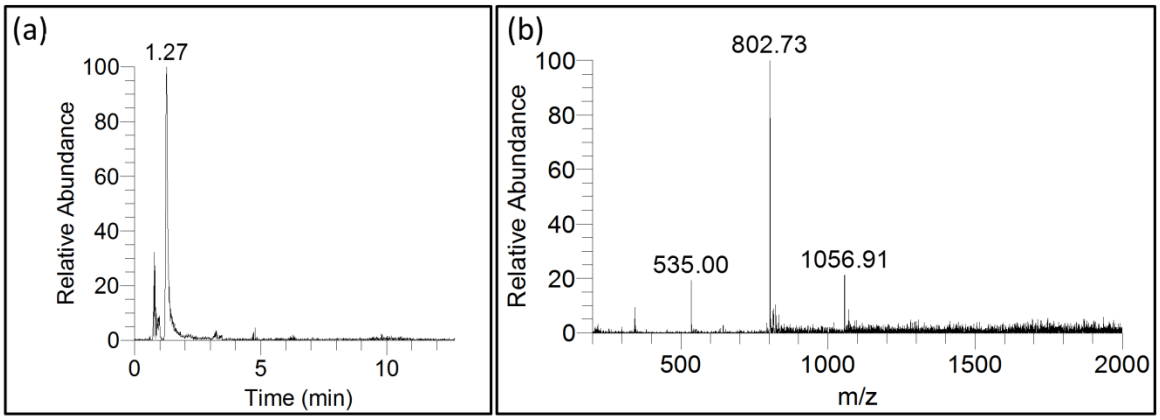


Supporting Figure S7 - LC-MS analysis of most abundant full-length extension products opposite **GG4** in DNA template by hPol *η* in the presence of all four dNTPs. (a) Sample reconstructed extracted ion chromatogram for *m/z* 803.0 for product with sequence 5´-pGAGCG and (b) mass spectrum of peak at retention time 1.31 min. (c) Sample reconstructed extracted ion chromatogram for *m/z* 746.14 for product with sequence 5´-pGAGC_GTG and (d) mass spectrum of peak at retention time 2.64 min. See Supporting Table S4 for full list of products and respective *m/z* assignments. See above for methodology.

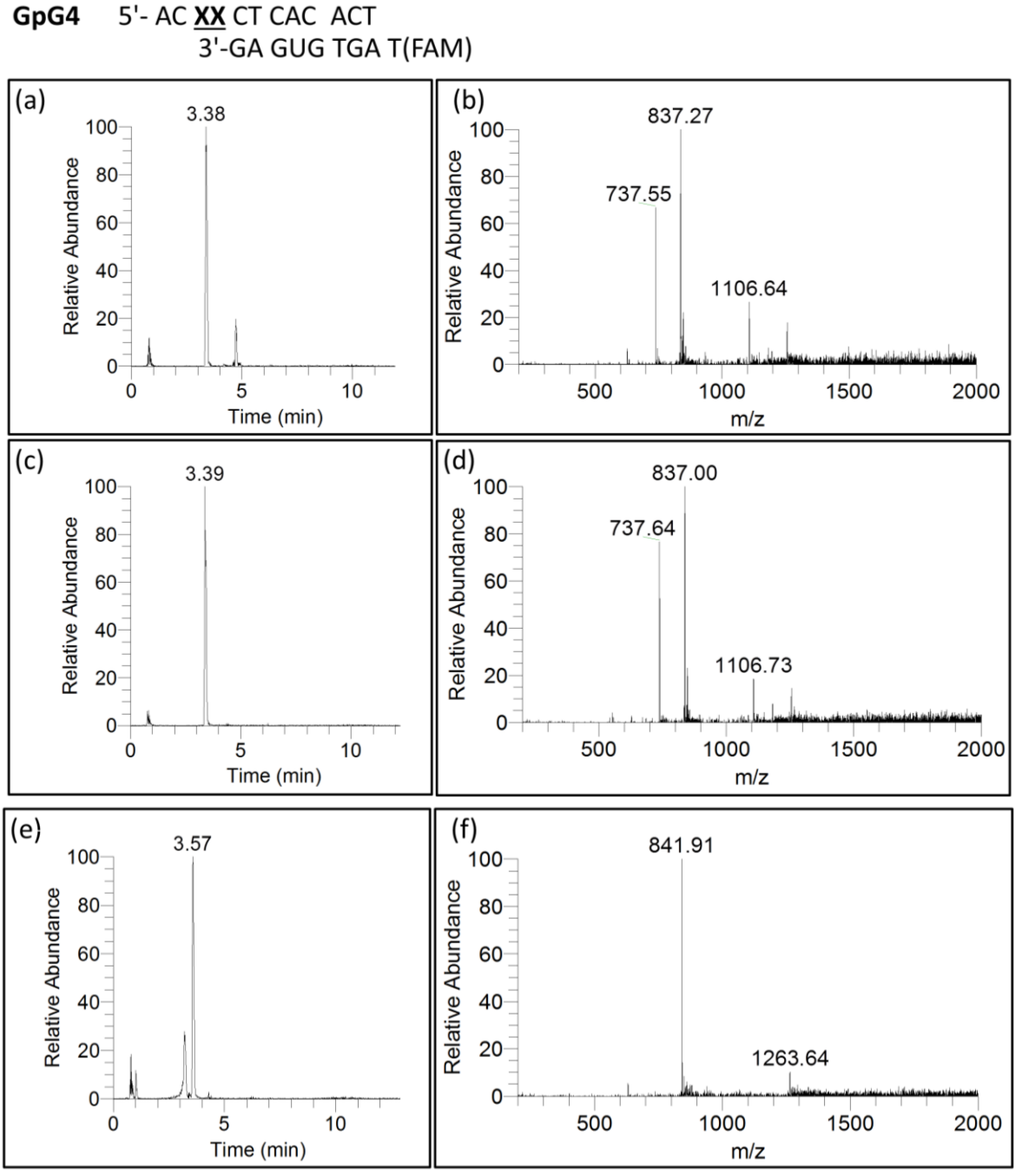


Supporting Figure S8 - LC-MS analysis of most abundant full-length extension products opposite **GG7** in DNA template by hPol *η* in the presence of all four dNTPs. (a) Sample reconstructed extracted ion chromatogram for *m/z* 803.0 for product with sequence 5´-pGAGCG and (b) mass spectrum of peak at retention time 1.27 min. See Supporting Table S4 for full list of products and respective *m/z* assignments. See above for methodology.

GG7 5'- AC **XX** CT CAC ACT
3'-GA GUG TGA T(FAM)

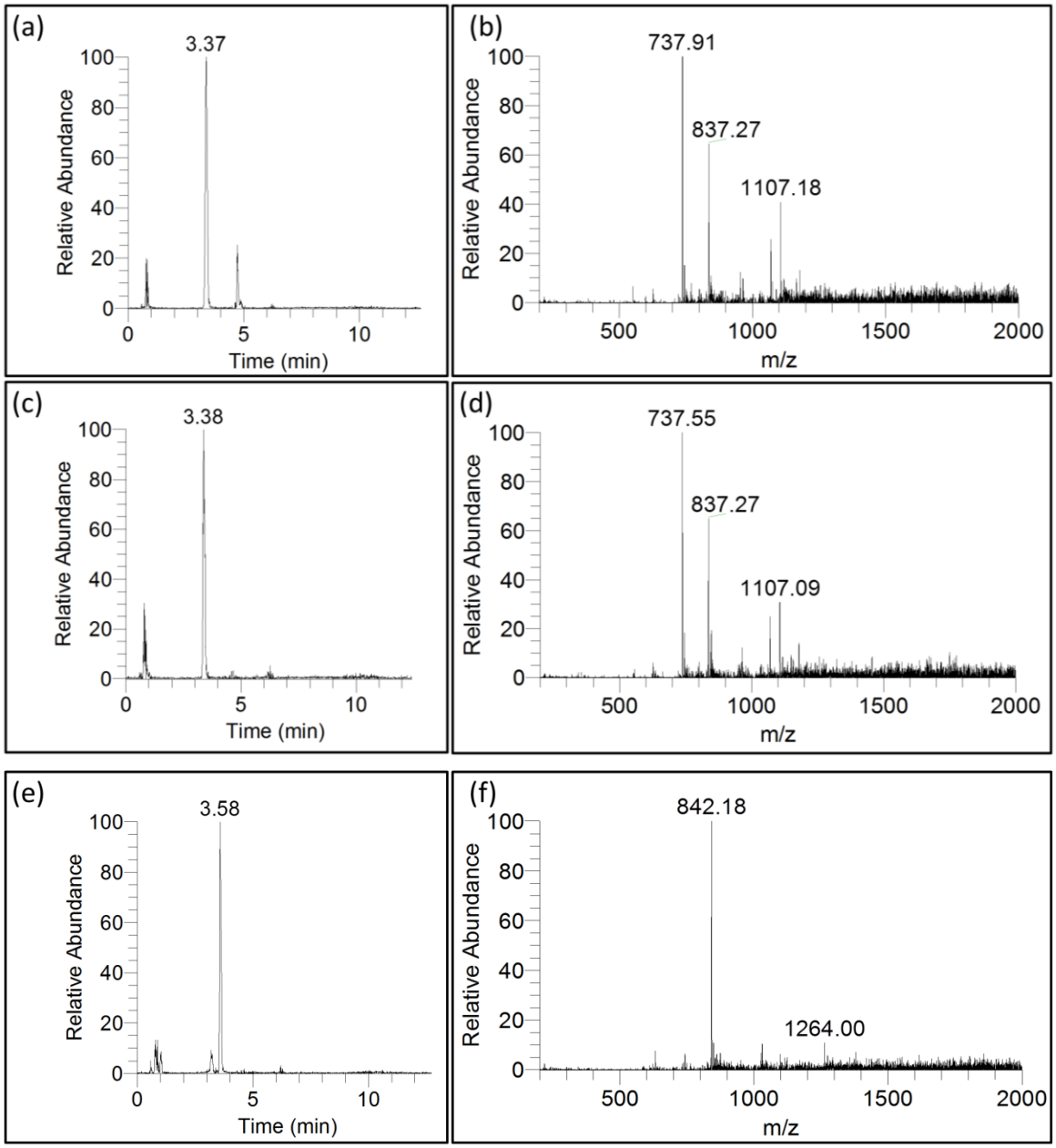


Supporting Figure S9 - LC-MS analysis of most abundant full-length extension products opposite **GG** in DNA template by hPol *η* in the presence of all four dNTPs. (a) Sample reconstructed extracted ion chromatogram for *m/z* 737.80 for product with sequence 5´-pGAGCTGT and (b) mass spectrum of peak at retention time 3.38 min. (c) Sample reconstructed extracted ion chromatogram for *m/z* 837.20 for product with sequence 5´-pGAGCCGTA and (d) mass spectrum of peak at retention time 3.39 min. (e) Sample reconstructed extracted ion chromatogram for *m/z* 842.21 for product with sequence 5´-pGAGCTGTA and (f) mass spectrum of peak at retention time 3.57 min. See Supporting Table S4 for full list of products and respective *m/z* assignments. See above for methodology.



Supporting Figure S10 - LC-MS analysis of most abundant full-length extension products opposite **GG** in DNA template by hPol *η* in the presence of all four dNTPs. (a) Sample reconstructed extracted ion chromatogram for *m/z* 737.80 for product with sequence 5´-pGAGCTGT and (b) mass spectrum of peak at retention time 3.37 min. (c) Sample reconstructed extracted ion chromatogram for *m/z* 837.20 for product with sequence 5´-pGAGCCGTA and (d) mass spectrum of peak at retention time 3.38 min. (e) Sample reconstructed extracted ion chromatogram for *m/z* 842.21 for product with sequence 5´-pGAGCTGTA and (f) mass spectrum of peak at retention time 3.58 min. See Supporting Table S4 for full list of products and respective *m/z* assignments. See above for methodology.

GpG7 5'- AC **XX** CT CAC ACT
3'-GA GUG TGA T(FAM)



Supporting Table S1 - Steady-state kinetics of incorporation of dNTP opposite 3´-end O^6-alkylated-dG of IaCL-containing template **GG4**, **GG7**, **GpG4**, **GpG7** (5´- AC **XX** CT CAC ACT) and unmodified template (**GG**) by hPol *η*. DNA primer sequence identity was 3´- GA GUG TGA T(FAM)-5´.

Template base	dNTP	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M x 10^{-3} ($\mu M^{-1}s^{-1}$)	f^1
G	dATP	0.045 ± 0.002	62 ± 9	0.7 ± 0.1	0.001
G	dGTP	0.052 ± 0.003	74 ± 13	0.7 ± 0.1	0.001
G	dCTP	2.32 ± 0.09	4.8 ± 0.5	480 ± 60	
G	dTTP	0.18 ± 0.01	111 ± 22	1.6 ± 0.3	0.003
GG4	dATP	0.037 ± 0.001	66 ± 5	0.56 ± 0.05	0.37
GG4	dGTP	0.029 ± 0.001	56 ± 4	0.52 ± 0.04	0.34
GG4	dCTP	0.12 ± 0.01	76 ± 17	1.5 ± 0.4	
GG4	dTTP	0.040 ± 0.003	111 ± 22	0.36 ± 0.08	0.24
GG7	dATP	0.031 ± 0.001	45 ± 5	0.70 ± 0.08	0.40
GG7	dGTP	0.025 ± 0.001	63 ± 7	0.40 ± 0.05	0.23
GG7	dCTP	0.12 ± 0.01	69 ± 16	1.7 ± 0.4	
GG7	dTTP	0.07 ± 0.01	92 ± 21	0.8 ± 0.2	0.47
GpG4	dATP	0.025 ± 0.001	80 ± 11	0.31 ± 0.04	0.20
GpG4	dGTP	0.0074 ± 0.0002	69 ± 5	0.107 ± 0.009	0.07
GpG4	dCTP	0.078 ± 0.003	51 ± 5	1.5 ± 0.2	
GpG4	dTTP	0.046 ± 0.006	150 ± 36	0.30 ± 0.08	0.20
GpG7	dATP	0.021 ± 0.001	83 ± 8	0.25 ± 0.02	0.16
GpG7	dGTP	0.0103 ± 0.0002	83 ± 5	0.13 ± 0.01	0.08
GpG7	dCTP	0.066 ± 0.003	43 ± 6	1.5 ± 0.2	
GpG7	dTTP	0.040 ± 0.002	72 ± 9	0.55 ± 0.08	0.36

[1] $(k_{cat}/K_M)_i / (k_{cat}/K_M)_{max}$

Supporting Table S2 - Steady-state kinetics of incorporation of dNTP opposite 5´-end O^6-alkylated-dG of IaCL-containing template **GG4**, **GG7**, **GpG4**, **GpG7** (5´- AC **XX** CT CAC ACT) and unmodified template (**GG**) by hPol *η*. DNA primer sequence identity was 3´- **C**GA GUG TGA T(FAM)-5´.

Template base	dNTP	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M x 10^{-3} ($\mu M^{-1}s^{-1}$)	f^1
G	dATP	0.028 ± 0.001	76 ± 5	0.36 ± 0.03	0.0009
G	dGTP	0.033 ± 0.001	38 ± 3	0.87 ± 0.07	0.002
G	dCTP	2.19 ± 0.11	5.7 ± 0.8	390 ± 60	
G	dTTP	0.09 ± 0.01	260 ± 34	0.35 ± 0.05	0.001
GG4	dATP	0.0021 ± 0.0001	93 ± 9	0.022 ± 0.002	0.01
GG4	dGTP	0.059 ± 0.002	32 ± 4	1.9 ± 0.2	
GG4	dCTP	0.0014 ± 0.0001	236 ± 26	0.0059 ± 0.0007	0.003
GG4	dTTP	0.009 ± 0.001	274 ± 42	0.033 ± 0.006	0.02
GG7	dATP	0.00176 ± 0.00004	76 ± 6	0.023 ± 0.002	0.02
GG7	dGTP	0.029 ± 0.001	27 ± 3	1.1 ± 0.1	
GG7	dCTP	0.00075 ± 0.00005	195 ± 35	0.0038 ± 0.0007	0.003
GG7	dTTP	0.0057 ± 0.0003	177 ± 24	0.032 ± 0.005	0.03
GpG4	dATP	0.0139 ± 0.0004	81 ± 8	0.17 ± 0.02	0.20
GpG4	dGTP	0.0057 ± 0.0001	26 ± 2	0.22 ± 0.02	0.25
GpG4	dCTP	0.060 ± 0.006	103 ± 20	0.6 ± 0.1	0.68
GpG4	dTTP	0.047 ± 0.003	55 ± 8	0.9 ± 0.1	
GpG7	dATP	0.0070 ± 0.0002	54 ± 5	0.13 ± 0.01	0.75
GpG7	dGTP	0.0055 ± 0.0001	32 ± 3	0.17 ± 0.02	
GpG7	dCTP	0.025 ± 0.002	375 ± 57	0.07 ± 0.01	0.38
GpG7	dTTP	0.049 ± 0.006	307 ± 60	0.16 ± 0.04	0.94

[1] $(k_{cat}/K_M)_i$ / $(k_{cat}/K_M)_{max}$

Supporting Table S3 - Steady-state kinetics of incorporation of dNTP opposite 5´-end O^6-alkylated-dG of IaCL-containing template **GG4**, **GG7**, **GpG4**, **GpG7** (5´- AC **XX** CT CAC ACT) and unmodified template (**GG**) by hPol *η*. DNA primer sequence identity was 3´- **T**GA GUG TGA T(FAM)-5´.

Template base	dNTP	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M x 10^{-3} ($\mu M^{-1}s^{-1}$)	f^1
G	dATP	0.00151 ± 0.00003	43 ± 3	0.035 ± 0.003	0.002
G	dGTP	0.00247 ± 0.00005	49 ± 4	0.051 ± 0.004	0.003
G	dCTP	0.168 ± 0.004	11 ± 1	15 ± 2	
G	dTTP	0.0034 ± 0.0003	91 ± 19	0.037 ± 0.008	0.002
GG4	dATP	0.00218 ± 0.00003	37 ± 2	0.059 ± 0.004	0.25
GG4	dGTP	0.0168 ± 0.0003	72 ± 5	0.23 ± 0.02	
GG4	dCTP	0.00134 ± 0.00005	196 ± 18	0.0068 ± 0.0007	0.03
GG4	dTTP	0.0016 ± 0.0001	144 ± 19	0.011 ± 0.001	0.05
GG7	dATP	0.00155 ± 0.00002	80 ± 4	0.019 ± 0.001	0.12
GG7	dGTP	0.0092 ± 0.0001	56 ± 2	0.162 ± 0.007	
GG7	dCTP	0.00103 ± 0.00004	198 ± 23	0.0052 ± 0.0006	0.032
GG7	dTTP	0.0016 ± 0.0001	189 ± 21	0.008 ± 0.001	0.05
GpG4	dATP	0.0048 ± 0.0002	65 ± 7	0.07 ± 0.01	0.04
GpG4	dGTP	0.0036 ± 0.0001	34 ± 2	0.104 ± 0.007	0.06
GpG4	dCTP	0.14 ± 0.01	78 ± 13	1.8 ± 0.3	
GpG4	dTTP	0.066 ± 0.006	98 ± 18	0.7 ± 0.1	0.38
GpG7	dATP	0.00206 ± 0.00004	57 ± 4	0.036 ± 0.003	0.64
GpG7	dGTP	0.00259 ± 0.00003	46 ± 2	0.056 ± 0.003	
GpG7	dCTP	0.0053 ± 0.0002	140 ± 13	0.038 ± 0.004	0.68
GpG7	dTTP	0.0057 ± 0.0003	117 ± 16	0.049 ± 0.007	0.87

[1] $(k_{cat}/K_M)_i$ / $(k_{cat}/K_M)_{max}$

Supporting Table S4 - LC-MS analysis of products of hpol *η* replicating template **GG4**, **GG7**, **GpG4**, **GpG7** (5´- AC **XX** CT CAC ACT) and unmodified **GG**. DNA primer sequence identity was 3´- GA GUG TGA T(FAM)-5´.

Name	**Fragment**	***m/z* calculated**	**GG**	**GG4**	**GG7**	**GpG4**	**GpG7**
4A	GAGA	(-2) 650.42		(-2) 650.18	(-2) 650.27		
4G	GAGG	(-2) 658.42		(-2) 658.27			
4T	GAGT	(-2) 645.92			(-2) 645.91		
5AA	GAGAA	(-2) 807.02 (-3) 537.68		(-2) 806.82 (-3) 537.73	(-2) 807.00		
5A_G	GAGAG	(-2) 815.02 (-3) 543.01		(-2) 814.73 (-3) 543.09	(-2) 814.82		(-2) 815.36 (-3) 543.18
5GA	GAGGA						
5CA	GAGCA	(-2) 795.01 (-3) 529.67	(-2) 794.73 (-3) 529.36	(-2) 794.91 (-3) 529.45	(-2) 794.73		
5AC	GAGAC						
5C_G	GAGCG	(-2) 803.01 (-3) 535.01	(-2) 802.73	(-2) 802.82 (-3) 534.64	(-2) 802.82 (-3) 534.68		
5TA	GAGTA	(-2) 802.52 (-3) 534.68			(-2) 802.82 (-3) 534.68		
5T_G	GAGTG	(-2) 810.52 (-3) 540.01		(-2) 810.36 (-3) 540.09	(-2) 810.36		
5__	GAGGT						
5CT	GAGCT	(-2) 790.51			(-2) 790.27		
6AGG	GAGAGG	(-2) 979.63		(-2) 979.45			
6TAG	GAGTAG	(-2) 967.12		(-2) 967.55			
6ATT	GAGATT	(-2) 954.62					(-2) 954.36
6TAT	GAGTAT						
6C_	GAGCGT	(-2) 955.11		(-2) 954.82	(-2) 954.91		
6C_G	GAGCGG	(-2) 967.62		(-2) 968.00			
6CTA	GAGCTA	(-2) 947.11 (-3) 631.07					(-2) 946.82 (-3) 631.09
6T_	GAGTGT	(-2) 962.62		(-2) 962.45			
7CC	GAGCCGT	(-2) 1099.7	(-2) 1099.73	(-2) 1099.70		(-2) 1099.36	(-3) 733.09

		(-3) 732.80	(-3) 732.55	(-3) 732.64		(-3) 733.00	
7CT	GAG CTG T	(-2) 1107.21 (-3) 737.80				(-2) 1106.73 (-3) 737.55	(-2) 1107.18 (-3) 737.91
7TC	GAG TCG T						
7TT	GAGTTGT	(-2) 1114.72 (-3) 742.80					(-2) 1114.73 (-3) 742.82
7C_A	GAGCGTA	(-2) 1111.72 (-3) 740.81	(-2) 1111.91 (-3) 740.36	(-2) 1111.91 (-3) 740.73			
7C_G	GAGC_GTG	(-2) 1119.72 (-3) 746.14		(-2) 1119.82 (-3) 745.91			
7AT	GAGATGT	(-2) 1119.22 (-3) 745.81				(-2) 1118.73 (-3) 745.36	
8CCA	GAG CCG TA	(-2) 1256.31 (-3) 837.20	(-2) 1255.82 (-3) 837.00	(-3) 836.91		(-2) 1256.36 (-3) 837.00	(-2) 1256.36 (-3) 837.27
8CCG	GAG CCG TG	(-2) 1264.31 (-3) 842.54	(-2) 1264.45 (-3) 842.36	(-3) 842.45		(-2) 1264.31 (-3) 842.27	(-2) 1264.00 (-3) 842.18
8CCC	GAG CCCG T	(-3) 829.19	(-3) 828.82				
8CCT	GAG CCG CT	(-2) 1244.30 (-3) 829.19				(-2) 1244.45 (-3) 829.36	
8CTA	GAG CTG TA	(-2) 1263.81 (-3) 842.21	(-2) 1264.27 (-3) 842.36			1263.64 841.91	(-2) 1264.00 (-3) 841.82
8TCA	GAG TCG TA						
8CTG	GAG CTG TG	(-2) 1271.81 (-3) 847.54				(-2) 1272.27 (-3) 847.54	(-2) 1272.18 (-3) 847.55
8TCG	GAG TCG TG						
8CTC	GAG CTG TC	(-2) 1251.80 (-3) 834.20				(-2) 1251.91 (-3) 834.18	(-2) 1251.55 (-3) 834.09
8CCT	GAG CCG TT						
8CTT	GAG CTG TT	(-2) 1259.30 (-3) 839.20				(-3) 839.09	(-3) 839.27
8TCT	GAG TCG TT						
8TTC	GAG TTG TC						
8TTA	GAG TTG TA	(-2) 1271.32 (-3) 847.21				(-2) 1271.18 (-3) 847.00	(-2) 1271.18 (-3) 846.91
8TTT	GAG TTG TT	(-2) 1266.81 (-3) 844.21				(-2) 1266.64 (-3) 844.64	

Supporting Table S5 - Percentages of full-length extension products determined by LC-MS/MS analysis.

Name	Fragment	GG	GG4	GG7	GpG4	GpG7
4A	GAGA		3	4		
4G	GAGG		1			
4T	GAGT			6		
5AA	GAGAA		7	3		
5A_G	GAGAG		12	6		2
5GA	GAGGA		< 1	< 1		
5CA	GAGCA	1	1	< 1		
5AC	GAGAC		1	< 1		
5CT	GAGCT			2		
5C_G	GAGCG	< 1	21	64		
5TA	GAGTA			4		
5T_G	GAGTG			3		
5__	GAGGT		3	3		
6AGG	GAGAGG		< 1			
6TAG	GAGTAG		1			
6ATT	GAGATT					< 1
6TAT	GAGTAT					< 1
6C_G	GAGCGG		5			
6CTA	GAGCTA					1
6C_	GAGCGT		2	5		
6T_	GAGTGT		2			
7CC	GAGCCGT	79	8		2	2
7CT	GAG CTG T				21	36
7TC	GAG TCG T				< 1	
7TT	GAGTTGT					2
7C_A	GAGC_GTA	1	< 1			
7C_G	GAGC_GTG		28			
7AT	GAGATGT				1	
8CCA	GAG CCG TA	11	2		31	16
8CCG	GAG CCG TG	6	1		9	3
8CCC	GAG CCCG TC	< 1				
8CCT	GAG CCG CT				1	
8CTA	GAG CTG TA	< 1			18	24
8TCA	GAG TCG TA				< 1	< 1
8CTG	GAG CTG TG				8	6
8TCG	GAG TCG TG				< 1	< 1
8CTC	GAG CTG TC				1	1
8CCT	GAG CCG TT				1	< 1
8CTT	GAG CTG TT				< 1	1

8TCT	GAG TCG TT					1
8TTC	GAG TTG TC				< 1	
8TTA	GAG TTG TA				2	2
8TTT	GAG TTG TT				1	

References

(1) O'Flaherty, D. K., and Wilds, C. J. (2015). *Euro. Chem. J. 21*, 10522–10529.

(2) Murata, S., Mizumura, Y., Hino, K., Ueno, Y., Ichikawa, S., and Matsuda, A. (2007). *J. Am. Chem. Soc. 129*, 10300–10301.

(3) Noll, D. M., Noronha, a M., and Miller, P. S. (2001). *J. Am. Chem. Soc. 123*, 3405–3411.

(4) Vaisman, A., and Chaney, S. G. (2000). *J Biol. Chem. 275*, 13017–13025.

(5) Zhao, Y., Biertumpfel, C., Gregory, M. T., Hua, Y.-J., Hanaoka, F., and Yang, W. (2012). *Proc. Natl. Acad. Sci. USA 109*, 7269–7274.

(6) Iwamoto, T., Hiraku, Y., Olkawa, S., Mizutani, H., Kojima, M., and Kawanishi, S. (2004). *Cancer Sci. 95*, 454–458.

(7) Yao, J., Dixon, K., and Carty, M. P. (2001). *Environ. Mol. Mutagen. 38*, 19–29.

(8) Masutani, C., Kusumoto, R., and Yamada, A. (1999). *Nature 399*, 700–704.

(9) Raychaudhury, P., and Basu, A. K. (2010). *J. Nucleic Acids 2010*, doi:10.4061/2010/101495.