SUPPORTING INFORMATION

Mechanism of Error-Free DNA Replication Past Lucidin-Derived DNA Damage by Human DNA Polymerase κ

Oliver P. Yockey[†,#], Vikash Jha[¶,#], Pratibha P. Ghodke[§,#], Tianzuo Xu[†], Wenyan Xu[†], Hong Ling[¶,*], P. I. Pradeepkumar[§,*] and Linlin Zhao[†,‡,*]

[†]*Department of Chemistry and Biochemistry and* [‡]*Science of Advanced Materials Program, Central Michigan University, Mount Pleasant, Michigan 48859, USA*

[¶] *Department of Biochemistry, Schulich School of Medicine & Dentistry, University of Western Ontario, London, Ontario, N6A 5C1, Canada*

[§] *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

TABLE OF CONTENTS

METHODS

Solid phase synthesis of LdG modified oligodeoxynucleotides. All the LdG modified DNA sequences were synthesized on 1 μmol scale using CPG solid support in a MerMade4 DNA synthesizer.[1] The concentration used for the LdG modified phosphoramidite solution was 0.1 M and the coupling time used was 6 min. The activator 5-(ethylthio)-*1H*-tetrazole (ETT) was used for the coupling step. Deprotection of the LdG modified oligodeoxynucleotides was carried out in four different steps. Initially, the deprotection of cyanoethyl groups was carried out using 10% diethylamine in ACN (800 μL, v/v) for 5 min. Supernatant layer was collected and solid support was washed with ACN (2 × 400 μL). The CPG beads were air dried and treated with 1 M DBU in ACN (1 mL, v/v) for 1 h at room temperature to remove *p*-nitrophenylethyl group.[2] The supernatant was removed and the CPG beads were washed with dry MeOH (2 × 1 mL) and dry ACN (3 × 1 mL). Further, the CPG was treated with 30% aq. NH_3 (1 mL, v/v) for 3 h for the cleavage of the solid support. Supernatant layer was collected separately and the CPG beads were washed with water (2 × 100 μL) and the combined aqueous layer was evaporated on a speedvac apparatus. To remove the base protecting groups, crude DNA was resuspended in 30% aq. NH_3 (1 mL, v/v) and heated at 55 °C for 16 h. Further, the NH_3 layer was evaporated on speedvac to obtain the DNA pellet and finally, the crude oligodeoxynucleotides were dissolved in water (100 μL). Crude DNAs were purified by using 20% (7 M urea) denaturing PAGE (30 W, 3 h) with 1 X TBE running buffer (89 mM each Tris and boric acid and 2 mM EDTA, pH ~ 8.3). The gel thickness was 1 mm, and gel dimension was 20 × 30 cm. The desired oligomer bands were visualized and marked under the UV lamp (260 nm). Further, the gel bands were cut and crushed into fine particles. The crushed gels were placed in a falcon tube and the oligomers were then extracted from the gel three times with 15 mL of TEN buffer (10 mM Tris, 1 mM EDTA, 300 mM NaCl, pH ~ 8.0) at 37 °C. Finally desalting of oligomers was carried out using C18 Sep-Pak column. Sep-Pak column was washed with 50% ACN in water (v/v), 5 mL of water and 5 mL of TEN buffer. After this, DNA samples were loaded onto the column followed by washing with water (5 mL) and TEN buffer (5 mL). Finally, DNA was eluted with 50% ACN in water (4 mL, v/v) and evaporated on a speedvac apparatus. The DNA pellet was dissolved in water and the concentration was measured at 260 nm in UV-Vis spectrophotometer using appropriate molar extinction coefficients (ε). The molar extinction coefficients were determined

employing the oligo analyzer from IDT (http://biophysics.idtdna.com/UVSpectrum.html). The integrity of all the modified DNAs was confirmed by using MALDI-TOF spectrometry (Table S1).

TABLES

Table S1. DNA oligodeoxynucleotides used in this study.

Experiment	Sequences (5'-3')	MW (calc.)	MW (found)
Synthetic LdG-containing oligodeoxynucleotides			
primer-extension assays (**D1**)	TCCTACCGTGCCTACCTGAACAGCTGGTCA CACTG*ATGCCTACGAGTACG	15547	15547
LC-MS (**D2**)	AGCTGGTCACACTG*AAGCGTTAGCCATTA	9183	9183
crystallization; insertion (**D3**)	ATGG*CTGATCCGCGCGGATCAG	7058	7057
crystallization; extension (**D4**)	CTATG*TCGATCCGCGGATCGAC	6993	6994
DNA oligodeoxynucleotides obtained commercially			
primer-extension assays	CGTACTCGTAGGCAT (15-mer oligomer)		
	CGTACTCGTAGGCATC (16-mer oligomer)		
LC-MS	TAATGGCTAACGCUT		

LdG-modified DNAs were characterized using MALDI-TOF positive reflectron or linear mode (calculated and found). G* denotes LdG.

Table S2. Steady-state kinetic parameters for single-nucleotide incorporations opposite the LdG lesion or unmodified dG catalyzed by pol κ, pol η, pol ι, or Rev1. Assays were performed at 37 °C with 80-100 nM 15/50-mer primer-template DNA (unmodified or adducted), 0.5-20 nM DNA polymerase, varying concentrations of dNTP, 4 % (v/v) glycerol, 5 mM DTT, 50 mM NaCl, 5 mM $MgCl_2$, and 100 μg mL^{-1} bovine serum albumin (BSA) in 50 mM Tris-HCl pH 7.4. Correct base pairs are in italic. Results are shown as the mean ± standard error of data from four experiments.

Polymerase	Template:dNTP pair	k_{cat} (min^{-1})	$K_{m,dNTP}$ (μM)	$k_{cat}/K_{m,dNTP}$ (min^{-1} μM^{-1})	f [b]
pol κ	*dG:C*	26 ± 1	2.0 ± 0.1	13 ± 0.8	
	dG:T	1.9 ± 0.1	470 ± 20	$(3.6 \pm 0.2) \times 10^{-3}$	2.7×10^{-4}
	dG:A	0.26 ± 0.03	32 ± 13	$(8.1 \pm 3.4) \times 10^{-3}$	6.2×10^{-4}
	dG:G	0.48 ± 0.13	1700 ± 600	$(2.8 \pm 1.2) \times 10^{-4}$	2.1×10^{-5}
	LdG:C	7.5 ± 2.1	5.6 ± 1.2	1.3 ± 0.5	
	LdG:T	0.16 ± 0.03	180 ± 10	$(8.9 \pm 1.8) \times 10^{-4}$	7.0×10^{-4}
	LdG:A	0.025 ± 0.004	55 ± 14	$(4.5 \pm 1.4) \times 10^{-4}$	3.5×10^{-4}
	LdG:G	0.022 ± 0.007	170 ± 40	$(1.3 \pm 0.5) \times 10^{-4}$	1.0×10^{-4}
pol η	*dG:C*	17 ± 3	0.72 ± 0.07	24 ± 4.8	
	dG:T	6.9 ± 1.8	78 ± 18	$(8.8 \pm 3.1) \times 10^{-2}$	3.7×10^{-3}
	dG:A	2.7 ± 0.1	38 ± 16	$(7.1 \pm 3.0) \times 10^{-2}$	3.0×10^{-3}
	dG:G	6.2 ± 1.0	49 ± 7	$(1.3 \pm 0.3) \times 10^{-1}$	5.3×10^{-3}
	LdG:C	1.2 ± 0.1	10 ± 7	$(1.2 \pm 0.9) \times 10^{-1}$	
	LdG:T	2.3 ± 0.3	100 ± 8	$(2.3 \pm 0.4) \times 10^{-2}$	1.9×10^{-1}
	LdG:A	2.5 ± 0.7	290 ± 30	$(8.6 \pm 2.6) \times 10^{-3}$	7.2×10^{-2}
	LdG:G	0.95 ± 0.07	53 ± 17	$(1.8 \pm 0.3) \times 10^{-2}$	1.5×10^{-1}
pol ι	*dG:C*	0.59 ± 0.16	3.6 ± 0.5	$(1.6 \pm 0.5) \times 10^{-1}$	
	dG:T	1.8 ± 0.3	260 ± 30	$(6.9 \pm 1.4) \times 10^{-3}$	4.3×10^{-2}
	dG:A	0.021 ± 0.001	97 ± 27	$(2.2 \pm 0.6) \times 10^{-4}$	1.4×10^{-3}
	dG:G	0.022 ± 0.001	120 ± 20	$(1.8 \pm 0.3) \times 10^{-4}$	1.1×10^{-3}
	LdG:C	0.045 ± 0.002	20 ± 2	$(2.3 \pm 0.3) \times 10^{-3}$	
	LdG:T	0.20 ± 0.01	45 ± 4	$(4.4 \pm 0.5) \times 10^{-3}$	1.9
	LdG:A	0.0030 ± 0.001	33 ± 4	$(9.1 \pm 3.2) \times 10^{-5}$	3.9×10^{-2}
	LdG:G	0.010 ± 0.003	110 ± 40	$(9.1 \pm 4.3) \times 10^{-5}$	3.9×10^{-2}
Rev1	*dG:C*	0.064 ± 0.001	0.065 ± 0.002	$(9.8 \pm 0.3) \times 10^{-1}$	
	dG:T	0.028 ± 0.002	8.0 ± 1.0	$(3.5 \pm 0.5) \times 10^{-3}$	3.6×10^{-3}
	dG:A	0.0072 ± 0.001	48 ± 29	$(1.5 \pm 0.9) \times 10^{-4}$	1.5×10^{-4}
	dG:G	0.11 ± 0.01	100 ± 6	$(1.1 \pm 0.1) \times 10^{-3}$	1.1×10^{-3}
	LdG:C	0.0078 ± 0.0001	0.033 ± 0.004	$(2.4 \pm 0.3) \times 10^{-1}$	
	LdG:T	0.024 ± 0.002	3.7 ± 0.7	$(6.5 \pm 1.3) \times 10^{-3}$	2.7×10^{-2}
	LdG:A	0.017 ± 0.001	19 ± 2	$(8.9 \pm 1.1) \times 10^{-4}$	3.7×10^{-3}
	LdG:G	0.065 ± 0.001	21 ± 4	$(3.1 \pm 0.6) \times 10^{-3}$	1.3×10^{-2}

[a] Misinsertion frequency $f = (k_{cat}/K_{m,dNTP})_{incorrect}/(k_{cat}/K_{m,dNTP})_{correct}$.

Table S3. Steady-state kinetic parameters for single-nucleotide incorporations opposite the 5′-neighboring dT of LdG (or unmodified dG) in a 16/50-mer DNA. Assay conditions were similar to those described in **Supporting Information Table S2**. Correct base pairs are in italic. Results are shown as the mean ± standard error of data from four experiments.

Polymerase	Template:dNTP pair	k_{cat} (min^{-1})	$K_{m,dNTP}$ (μM)	$k_{cat}/K_{m,dNTP}$ (min^{-1} μM^{-1})	f [b]
pol κ	3′-G—*dT:A*	72 ± 20	3.6 ± 2.0	20 ± 12	
	dT:C	1.5 ± 0.1	129 ± 30	$(1.2 \pm 0.3) \times 10^{-2}$	6.0×10^{-4}
	dT:T	0.65 ± 0.10	280 ± 50	$(2.3 \pm 0.5) \times 10^{-3}$	1.2×10^{-4}
	dT:G	0.51 ± 0.13	85 ± 7	$(6.0 \pm 1.6) \times 10^{-3}$	3.0×10^{-4}
	3′-LdG—*dT:dA*	16 ± 9	4.6 ± 1.2	3.5 ± 2.2	
	dT:C	0.34 ± 0.02	150 ± 12	$(2.3 \pm 0.2) \times 10^{-3}$	6.5×10^{-4}
	dT:T	0.27 ± 0.03	620 ± 130	$(4.4 \pm 1.0) \times 10^{-4}$	1.3×10^{-4}
	dT:G	2.7 ± 0.8	160 ± 40	$(1.7 \pm 0.7) \times 10^{-2}$	4.9×10^{-3}
pol η	3′-G—*dT:A*	7.4 ± 1	3.2 ± 0.9	2.3 ± 0.7	
	dT:C	0.38 ± 0.09	160 ± 20	$(2.4 \pm 0.6) \times 10^{-3}$	1.0×10^{-3}
	dT:T	2.1 ± 0.2	500 ± 110	$(4.2 \pm 1.0) \times 10^{-3}$	1.8×10^{-3}
	dT:G	0.91 ± 0.11	28 ± 5	$(3.2 \pm 0.7) \times 10^{-2}$	1.4×10^{-2}
	3′-LdG—*dT:dA*	0.071 ± 0.023	0.18 ± 0.09	$(3.6 \pm 2.1) \times 10^{-1}$	
	dT:C	0.025 ± 0.004	210 ± 80	$(1.2 \pm 0.5) \times 10^{-4}$	3.3×10^{-3}
	dT:T	0.18 ± 0.01	193 ± 38	$(9.3 \pm 1.9) \times 10^{-4}$	2.5×10^{-3}
	dT:G	0.017 ± 0.007	9.1 ± 5.0	$(1.9 \pm 1.3) \times 10^{-3}$	5.2×10^{-3}
pol ι	3′-G—*dT:A*	0.23 ± 0.01	4.9 ± 0.9	$(4.7 \pm 0.9) \times 10^{-2}$	
	dT:C	5.7 ± 1.7	2900 ± 900	$(2.0 \pm 0.9) \times 10^{-3}$	4.2×10^{-2}
	dT:T	0.76 ± 0.07	180 ± 40	$(4.2 \pm 1.0) \times 10^{-3}$	8.9×10^{-2}
	dT:G	0.49 ± 0.11	22 ± 7	$(2.2 \pm 0.9) \times 10^{-2}$	4.7×10^{-1}
	3′-LdG—*dT:dA*	0.081 ± 0.001	4.7 ± 0.4	$(1.7 \pm 0.2) \times 10^{-2}$	
	dT:C	0.0023 ± 0.0002	160 ± 50	$(1.4 \pm 0.5) \times 10^{-5}$	8.2×10^{-4}
	dT:T	0.13 ± 0.01	270 ± 10	$(4.8 \pm 0.4) \times 10^{-4}$	2.8×10^{-2}
	dT:G	0.038 ± 0.013	7.8 ± 3.3	$(4.9 \pm 0.3) \times 10^{-3}$	2.9×10^{-1}

[a] Misinsersion frequency $f = (k_{cat}/K_{m,dNTP})_{incorrect}/(k_{cat}/K_{m,dNTP})_{correct}$.

Table S4. Root-mean-square deviations (RMSDs) of ternary pol ϰ complexes.

Structure (space group)	LDG-I ($P2_12_12_1$)	LDG-E ($P2_12_12_1$)	5T14 ($P2_12_12_1$)	4U7C[3] ($P2_12_12_1$)	2OH2[4] ($C222_1$)	2W7P[5] ($C222_1$)	3IN5[6] ($C222_1$)
LDG-I	0.040[a]	0.29/0.34[b] 0.33/0.29	0.32/0.31 0.32/0.31	0.31/0.34 0.31/0.33	1.16/0.97 1.16/0.97	0.85/0.82 0.85/0.82	1.06/0.85 1.06/0.85
LDG-E		0.268	0.26/0.31 0.40/0.28	0.36/0.41 0.38/0.36	1.12/0.95 1.10/0.97	0.89/0.85 0.88/0.81	1.03/0.84 0.97/0.83
5T14			0.31	0.42/0.42 0.41/0.46	1.16/1.15 0.94/0.97	0.90/0.90 0.87/0.83	1.01/1.05 0.84/0.85
4U7C				0.18	1.18/1.17 0.98/1.00	0.85/0.85 0.80/0.79	1.15/1.10 0.85/0.86
2OH2					0.72	1.01/1.30 0.89/1.14	0.48/0.7 0.81/0.48
2W7P						0.85	0.97/0.7 1.16/1.00
3IN5							0.67

[a] RMSD values along the diagonal line are the rmsds between the two complexes (A and B) in the same structure. Each crystal structure has two complexes (A and B) in an asymmetric unit.
[b] RMSD values off the diagonal line are for the superposition of proteins over 423 Cα atoms in pair-wise comparisons among two (A and B) independent complexes (AB vs A'B') in different ternary complex structures. The order of the rmsd values: AA'/AB'/BA'/BB'; A and B are from one structure, and A' and B' are from another.

FIGURES

Figure S1. MALDI spectrum of **D1**, 5′-TCCTACCGTGCCTACCTGAACAGCTGGTCACACTG*ATG CCTACGAGTACG-3′ Calculated mass, $[M+H]^+$ 15547; Observed mass, $[M+H]^+$ 15547

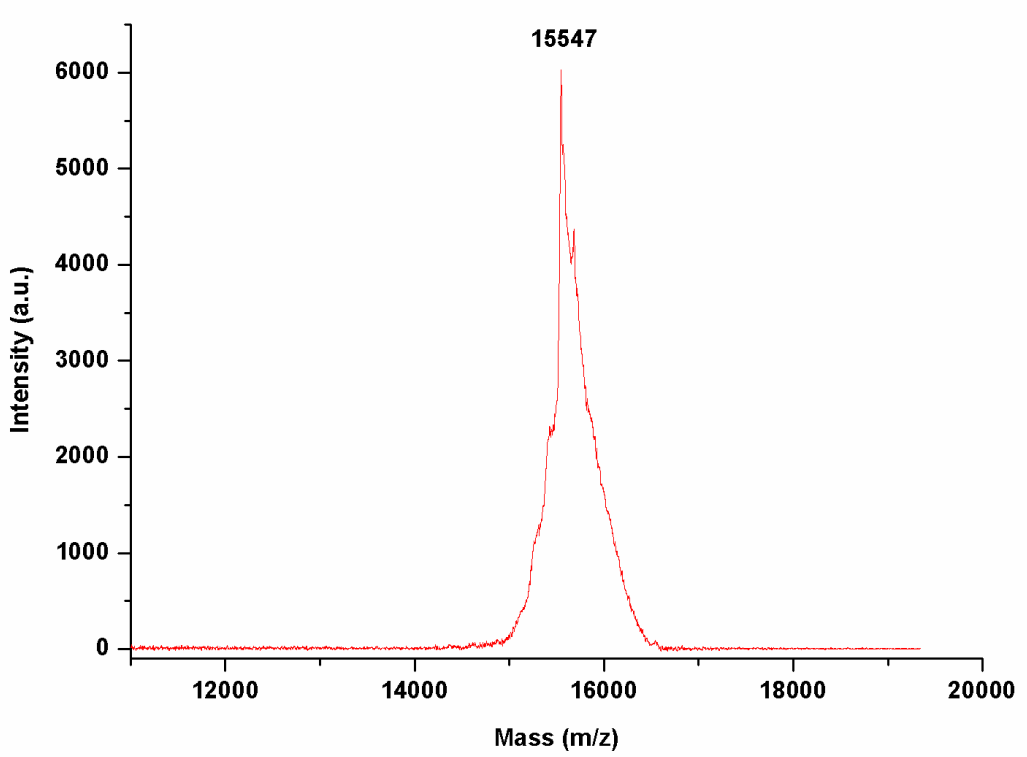


Figure S2. MALDI Spectrum of **D2**, 5′-AGCTGGTCACACTG*AAGCGTTAGCCATTA-3′
Calc. mass, $[M+H]^+$ 9183; Obs. mass, $[M+H]^+$ 9183

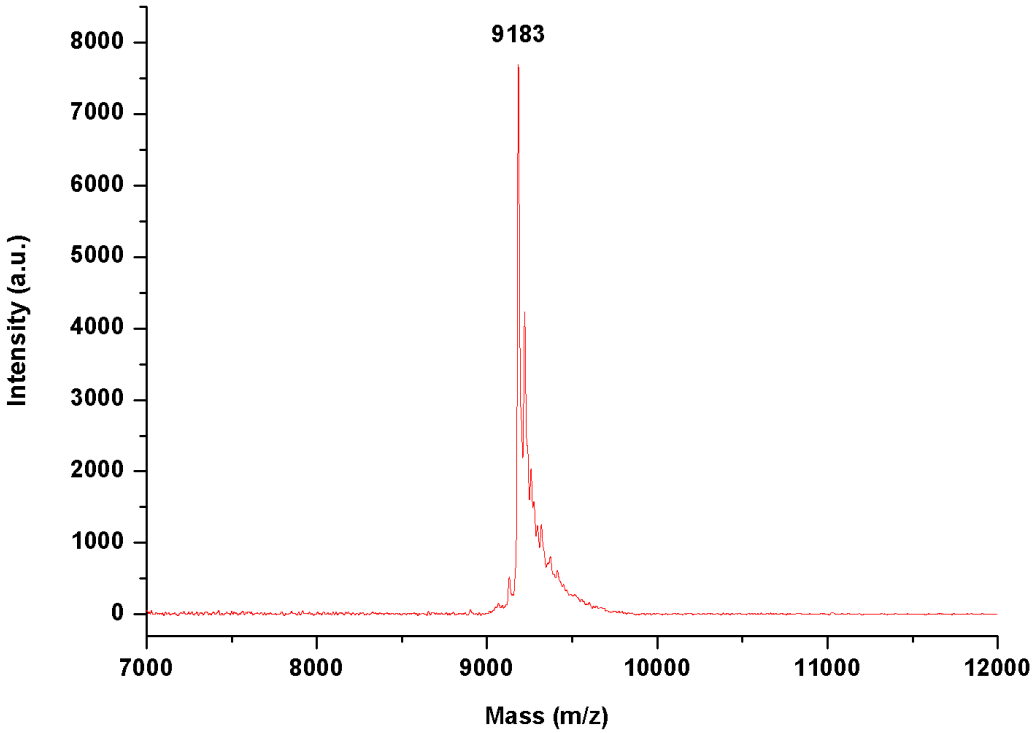


Figure S3. MALDI spectrum of **D3**, 5′-ATGG*CTGATCCGCGCGGATCAG-3′ Calc. mass, $[M+H]^+$ 7058; Obs. mass, $[M+H]^+$ 7057

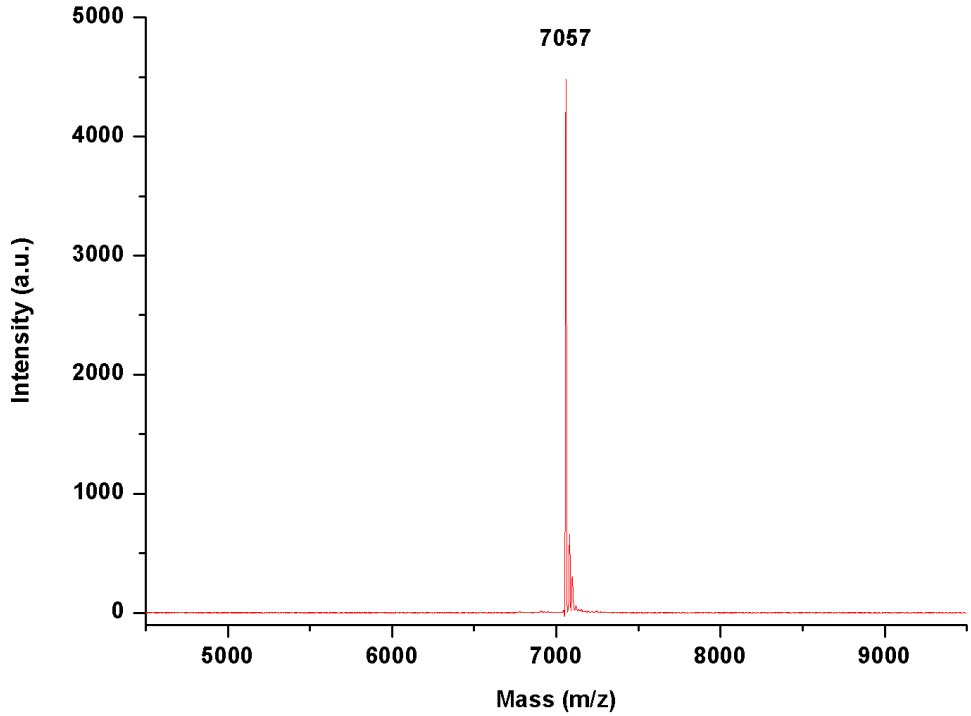


Figure S4. MALDI Spectrum of **D4**, 5′-CTATG*TCGATCCGCGGATCGAC-3′ Calc. mass, $[M+H]^+$ 6993; Obs. mass, $[M+H]^+$ 6994

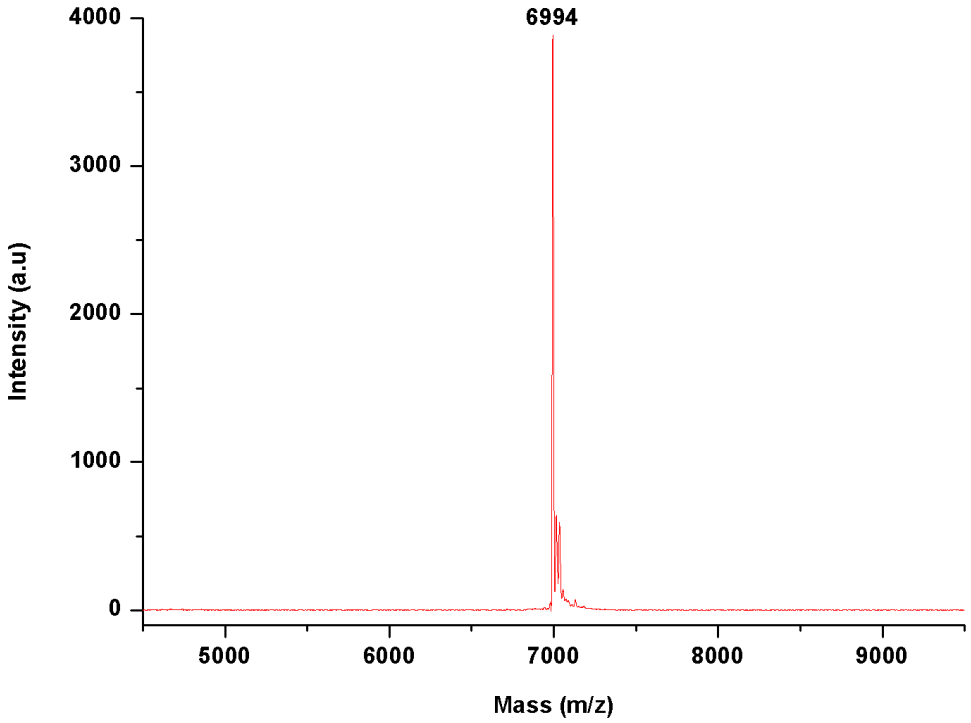


Figure S5. LC-MS analysis of pol κ-catalyzed primer-extension reaction with a dG-containing DNA substrate. (A) Extracted ion chromatograms of the cleaved primer (*m/z* 672.1, $[M\text{-}6H^+]^{6-}$), a 13-mer oligomer product (*m/z* 674, $[M\text{-}6H^+]^{6-}$), and a 12-mer oligomer primer-extension product (*m/z* 620, $[M\text{-}6H^+]^{6-}$). (B), (C), and (D) are LC-MS/MS collision-induced dissociation mass spectra for precursor ions 672.1, 674.7, and 620.0, respectively. The observed product ions that match the theoretical fragmented ions are labeled. Reaction conditions are described in the Experimental section.

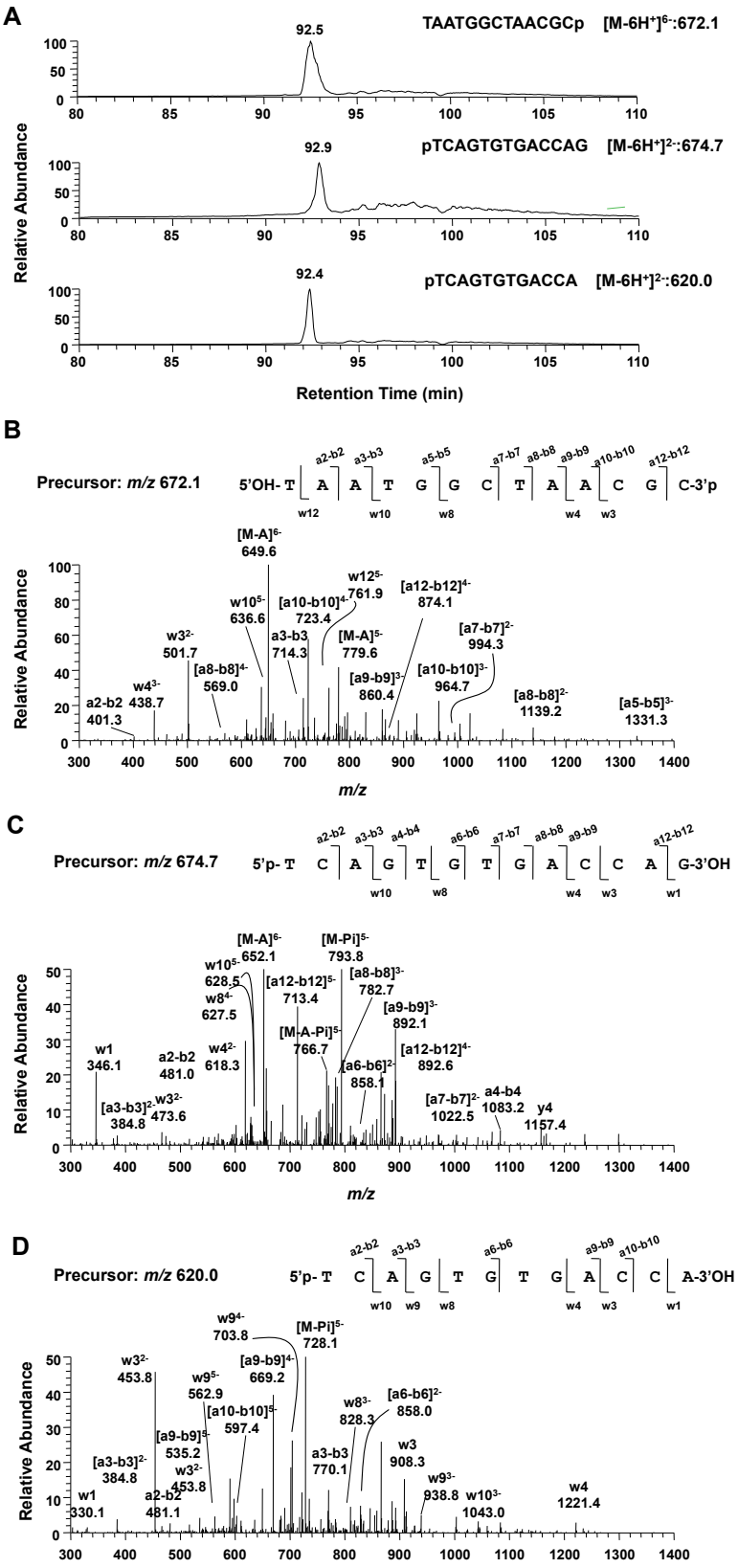
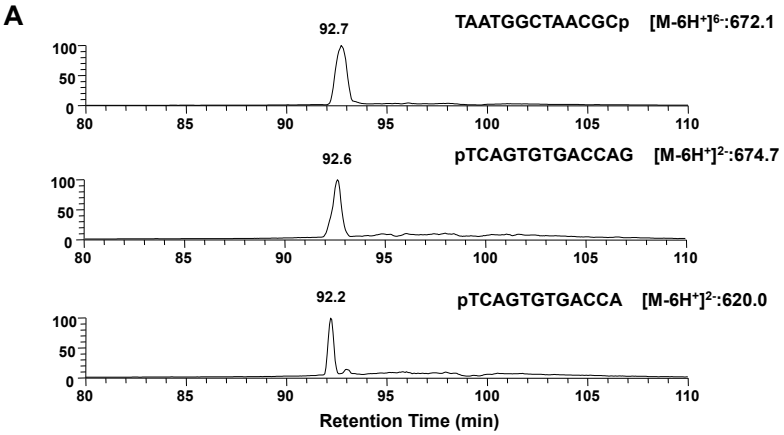
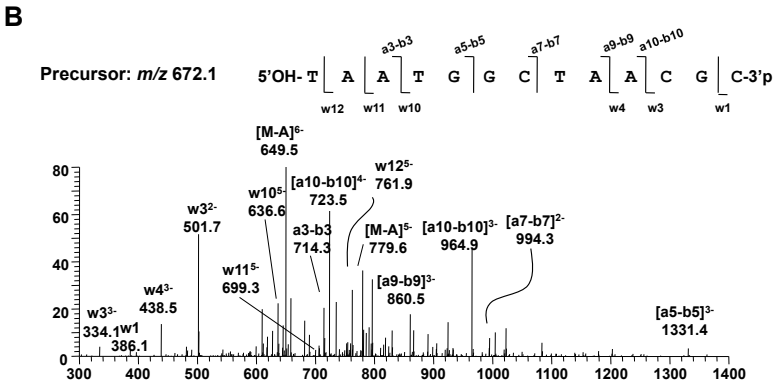
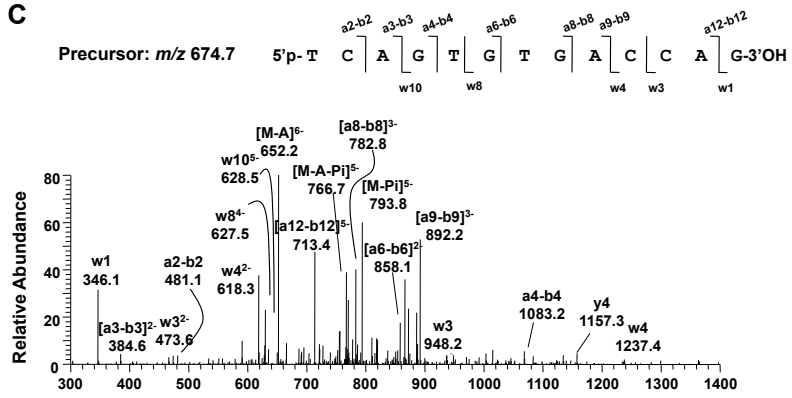
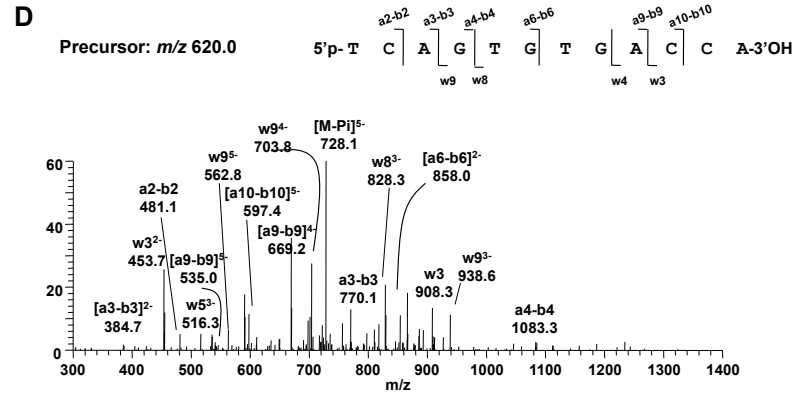

A
TAATGGCTAACGCp [M-6H+]6-:672.1
92.5
pTCAGTGTGACCAG [M-6H+]2-:674.7
92.9
pTCAGTGTGACCA [M-6H+]2-:620.0
92.4
Relative Abundance
Retention Time (min)
B
Precursor: m/z 672.1
5'OH- T A A T G G C T A A C G C-3'p
[M-A]6- 649.6
w125- 761.9
[a12-b12]4- 874.1
w105- 636.6
[a10-b10]4- 723.4
w32- 501.7
a3-b3 714.3
[M-A]5- 779.6
[a7-b7]2- 994.3
[a8-b8]4- 569.0
[a9-b9]3- 860.4
[a10-b10]3- 964.7
w43- 438.7
[a8-b8]2- 1139.2
a2-b2 401.3
[a5-b5]3- 1331.3
m/z
C
Precursor: m/z 674.7
5'p- T C A G T G T G A C C A G-3'OH
[M-A]6- 652.1
[M-Pi]5- 793.8
w105- 628.5
[a12-b12]5- 713.4
[a8-b8]3- 782.7
w84- 627.5
[a9-b9]3- 892.1
[M-A-Pi]5- 766.7
w1 346.1
a2-b2 481.0
w42- 618.3
[a12-b12]4- 892.6
[a6-b6]2- 858.1
[a3-b3]2- 384.8
w32- 473.6
[a7-b7]2- 1022.5
a4-b4 1083.2
y4 1157.4
m/z
D
Precursor: m/z 620.0
5'p- T C A G T G T G A C C A-3'OH
w94- 703.8
[M-Pi]5- 728.1
w32- 453.8
[a9-b9]4- 669.2
w95- 562.9
[a6-b6]2- 858.0
w83- 828.3
[a10-b10]5- 597.4
[a9-b9]5- 535.2
a3-b3 770.1
w3 908.3
[a3-b3]2- 384.8
w32- 453.8
w93- 938.8
w103- 1043.0
w4 1221.4
w1 330.1
a2-b2 481.1
Relative Abundance

Figure S6. LC-MS analysis of pol κ-catalyzed primer-extension reaction with an LdG-containing DNA substrate. (A) Extracted ion chromatograms of the cleaved primer (*m/z* 672.1, $[M\text{-}6H^+]^{6-}$), a 13-mer oligomer product (*m/z* 674, $[M\text{-}6H^+]^{6-}$), and a 12-mer oligomer primer-extension product (*m/z* 620, $[M\text{-}6H^+]^{6-}$). (B), (C), and (D) are LC-MS/MS collision-induced dissociation mass spectra for precursor ions 672.1, 674.7, and 620.0, respectively. The observed product ions that match the theoretical fragmented ions are labeled. Reaction conditions are described in the Experimental section.


A
92.7
TAATGGCTAACGCp [M-6H+]6-:672.1
92.6
pTCAGTGTGACCAG [M-6H+]2-:674.7
92.2
pTCAGTGTGACCA [M-6H+]2-:620.0
Retention Time (min)


B
Precursor: m/z 672.1
5'OH- T A A T G G C T A A C G C-3'p
[M-A]6- 649.5
w12 5- 761.9
[a10-b10]4- 723.5
w10 5- 636.6
w3 2- 501.7
a3-b3 714.3
[M-A]5- 779.6
[a10-b10]3- 964.9
[a7-b7]2- 994.3
w11 5- 699.3
[a9-b9]3- 860.5
w4 3- 438.5
w3 3- 334.1
w1 386.1
[a5-b5]3- 1331.4


C
Precursor: m/z 674.7
5'p- T C A G T G T G A C C A G-3'OH
[M-A]6- 652.2
[a8-b8]3- 782.8
w10 5- 628.5
[M-A-Pi]5- 766.7
[M-Pi]5- 793.8
w8 4- 627.5
[a12-b12]5- 713.4
[a9-b9]3- 892.2
[a6-b6]2 858.1
w1 346.1
a2-b2 481.1
w4 2- 618.3
a4-b4 1083.2
y4 1157.3
[a3-b3]2- 384.6
w3 2- 473.6
w3 948.2
w4 1237.4
Relative Abundance


D
Precursor: m/z 620.0
5'p- T C A G T G T G A C C A-3'OH
w9 4- 703.8
[M-Pi]5- 728.1
w9 5- 562.8
w8 3- 828.3
[a6-b6]2- 858.0
a2-b2 481.1
[a10-b10]5- 597.4
[a9-b9]4- 669.2
w3 2- 453.7
[a9-b9]5- 535.0
w9 3- 938.6
a3-b3 770.1
w3 908.3
[a3-b3]2- 384.7
w5 3- 516.3
a4-b4 1083.3
m/z

REFERENCES

1. Ghodke, P. P., Harikrishna, S., and Pradeepkumar, P. I. (2015) Synthesis and Polymerase-Mediated Bypass Studies of the N 2-Deoxyguanosine DNA Damage Caused by a Lucidin Analogue, *J. Org. Chem. 80*, 2128-2138.
2. Choi, J.-Y., and Guengerich, F. P. (2004) Analysis of the effect of bulk at N2-alkylguanine DNA adducts on catalytic efficiency and fidelity of the processive DNA polymerases bacteriophage T7 exonuclease-and HIV-1 reverse transcriptase, *J. Biol. Chem. 279*, 19217-19229.
3. Jha, V., Bian, C., Xing, G., and Ling, H. (2016) Structure and mechanism of error-free replication past the major benzo[*a*]pyrene adduct by human DNA polymerase ϰ, *Nucleic Acids Res. 44*, 4957-4967.
4. Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A., Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. (2007) Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass, *Mol Cell 25*, 601-614.
5. Irimia, A., Eoff, R. L., Guengerich, F. P., and Egli, M. (2009) Structural and functional elucidation of the mechanism promoting error-prone synthesis by human DNA polymerase kappa opposite the 7,8-dihydro-8-oxo-2'-deoxyguanosine adduct, *J Biol Chem 284*, 22467-22480.
6. Vasquez-Del Carpio, R., Silverstein, T. D., Lone, S., Swan, M. K., Choudhury, J. R., Johnson, R. E., Prakash, S., Prakash, L., and Aggarwal, A. K. (2009) Structure of human DNA polymerase kappa inserting dATP opposite an 8-OxoG DNA lesion, *PLoS One 4*, e5766.