Supplementary material for

“Cytotoxic and mutagenic properties of alkyl phosphotriester lesions in

*Escherichia coli* cells”

Jiabin Wu¹, Pengcheng Wang² and Yinsheng Wang¹,²,*

Environmental Toxicology Graduate Program and Department of Chemistry, University of

California Riverside, Riverside, CA 92521-0403

*To whom correspondence should be addressed: E-mail: Yinsheng.Wang@ucr.edu
Supplementary Materials and Methods:

**Mass Spectrometry (MS) and NMR**
Electrospray ionization-MS (ESI-MS) and tandem MS (MS/MS) experiments were carried out on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Acetonitrile and water (50:50, v/v) was used as the solvent for electrospray. The spray voltage was 3.0 kV, and the temperature of the ion transport tube was maintained at 275°C. High-resolution mass spectra (HRMS) were acquired on an Agilent 6510 Q-TOF LC/MS instrument (Agilent Technologies, Palo Alto, CA) equipped with an electrospray ionization (ESI) source.

$^1$H NMR spectra were recorded at 400 MHz on a Bruker Avance NEO 400 NMR spectrometer (Bruker Inc., Billerica, MA), and $^{31}$P NMR spectra were acquired at 80 MHz on a Varian Inova 300 NMR spectrometer (Varian Inc., Palo Alto, CA).

**Reaction yields, and NMR and mass spectrometric characterizations of the synthetic products:**

5'-O-(4,4'-dimethoxytrityl)-thymidine-3'-O-(ethyl-N,N'-diisopropyl)phosphoramidite: The compound was obtained as a colorless solid (yield 47%). $^1$H NMR (400 MHz, CDCl$_3$) δ 8.45 (s, 1H), 7.72-7.66 (d, $J$=0.92, 1H), 7.45-7.40 (m, 2H), 7.32-7.25 (m, 7H), 6.85 (m, 4H), 6.45 (m, 1H), 4.67 (m, 1H), 4.20 (m, 1H), 3.83 (s, 6H), 3.70-3.53 (m, 5H), 3.40-3.29 (m, 1H), 2.61-2.45 (m, 1H), 2.37-2.27 (m, 1H), 1.32-1.16 (m, 25H), 1.15-1.08 (m, 3H). $^{31}$P NMR (80 MHz, CDCl$_3$) δ = 147.9, 147.2. HRMS (ESI) calcd for C$_{39}$H$_{50}$N$_3$O$_8$P [M+H]$^+$ 720.3336, found 720.3349.

5'-O-(4,4'-dimethoxytrityl)-thymidine-3'-O-(n-propyl-N,N'-diisopropyl)phosphoramidite: The compound was obtained as a colorless solid (yield 40%). $^1$H NMR (400 MHz, CDCl$_3$) δ 9.22 (s, 1H), 7.63-7.53 (d, $J$=14.77 Hz, 1H), 7.39-7.35 (m, 2H), 7.27-7.10 (m, 7H), 6.40 (dd, $J$=7.45, 14.07 Hz, 1H), 4.69-4.55 (d, $J$=3.62 Hz, 1H), 4.25-4.0 (m, 1H), 3.74 (s, 6H), 3.54-3.44 (m, 5H), 2.58-2.4 (m, 1H), 2.28-2.19 (m, 1H), 1.68-1.52 (m, 2H), 1.43-1.10 (m, 25H), 0.9-0.75 (t, $J$=7.34 Hz, 3H). $^{31}$P NMR (80 MHz, CDCl$_3$) δ = 148.30, 147.69. HRMS (ESI) calcd for C$_{40}$H$_{52}$N$_3$O$_8$P [M+H]$^+$ 734.3492, found 734.3510.
Figure S1. The $^1$H NMR spectrum of dT-ethylphosphoramidite (400 MHz, CDCl$_3$, 25°C).

Figure S2. The $^1$H NMR spectrum of dT-$n$-propylphosphoramidite (400 MHz, CDCl$_3$, 25°C).
Figure S3. The $^{31}$P NMR spectrum of the phosphoramidite building block of dT-ethylphosphoramidite (80 MHz, CDCl$_3$, 25°C).

Figure S4. The $^{31}$P NMR spectrum of the phosphoramidite building block of dT-$n$-propylphosphoramidite (80 MHz, CDCl$_3$, 25°C).
Figure S5. (a) The HPLC trace for the purification of T(Me)T-dimer and $^{31}$P NMR spectrum of the $S_p$-T(Me)T-1 (b) and $R_p$-T(Me)T-2 (c) (80 MHz, MeOD, 25ºC).
**Figure S6.** The HPLC traces for the separation of $S_p$-T(Me)T and $R_p$-T(Me)T in the digestion mixtures of synthesized 12mer site-specific lesion-containing ODN (5'-ATGGCT(Me)TGCTAT-3'). ‘dC’, ‘dG’, ‘dT’, ‘dA’ represent 2'-deoxycytidine, 2'-deoxyguanosine, thymidine, and 2'-deoxyadenosine, respectively. (a) The HPLC trace for the separation of the synthesized T(Me)T dimer together with canonical 2'-deoxynucleosides. (b) The HPLC trace for separation of the $S_p$-T(Me)T standard. (c) The HPLC trace for the separation of the $R_p$-T(Me)T standard. (d) The HPLC trace for the separation of the digestion mixture of 5'-ATGGCT(Me$_{S_p}$)TGCTAT-3'. (e) The HPLC trace for the separation of the digestion mixture of 5'-ATGGCT(Me$_{R_p}$)TGCTAT-3'.
Figure S7. HPLC traces for the separations of the synthesized 12mer alkyl phosphotriester-bearing ODNs: (a) methyl; (b) ethyl; (c) n-propyl; (d) n-butyl.
Figure S8. ESI-MS & MS/MS characterizations of d(ATGGCT(X)TGCTAT), X = Methyl group: (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the [M-3H]^3- ion (m/z 1221.0).
Figure S9. ESI-MS & MS/MS characterizations of d(ATGGCT(X)TGCTAT), X = Ethyl group: (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the [M-3H]$^{3-}$ ion (m/z 1225.4)
Figure S10. ESI-MS & MS/MS characterizations of d(ATGGCT(X)TGCTAT), X = n-propyl group: (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the [M-3H]$^{3-}$ ion (m/z 1230.4)
Figure S11. ESI-MS & MS/MS characterizations of d(ATGGCT(X)TGCTAT), X = n-butyl group: (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the [M-3H]^3- ion (m/z 1235.2)
a. Construction of Lesion-containing genome:

5'-CAGGAAAGCTATGACCCATGATTCGATGAACTGCTGGAAACAGCAATGCT(X)TGCTATAGCTCAGATGATTCAGTGAGTGGAAGACATGGCTAATTCACTGCGTCGTTTTACAACGTCGTGACTGGGA-3'  
3'-CTTTTCGATGACTGAACTACTAGTCACCTACCTTTGACAAAGCACTGCTGGAAAGTTTACAACGTCGTGACTGGGA-3'  
3'-CGATATTAAGTGACCCGGCAGCAGACAAAA-5'

b. Construction of competitor genome:

5'-CAGGAAAGCTATGACCCATGATTCGATGAACTGCTGGAAACAGCAATGCT(X)TGCTATAGCTCAGATGATTCAGTGAGTGGAAGACATGGCTAATTCACTGCGTCGTTTTACAACGTCGTGACTGGGA-3'  
3'-CTTTTCGATGACTGAACTACTAGTCACCTACCTTTGACAAAGCACTGCTGGAAAGTTTACAACGTCGTGACTGGGA-3'  
3'-CGATATTAAGTGACCCGGCAGCAGACAAAA-5'

**Figure S12.** Schematic diagrams showing the construction of the lesion-containing (a) and competitor (b) genomes. Displayed are the partial sequence of the linearized M13 genome, the 22-mer lesion-containing insert or the 25-mer lesion-free insert (shown in red), and the two scaffolds employed for the ligation reactions. The lesion site is annotated with (X).
Figure S13. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of \( \delta_p \)- and \( \rho_p \)-methyl phosphotriester in wild-type (WT), SOS-WT, SOS-\( \Delta \)Pol II, IV, V and triple knockout (TKO) AB1157 E. coli cells. (a) Selective labelling of original lesion-containing strand and its complementary strand via sequential restriction digestion. ‘SAP’ and ‘PNK’ represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT, 10mer GC, and 10mer GT represent the \([5'-32P]\)-labeled standard ODNs \(5'-GGCMNGCTAT-3'\), with ‘MN’ being TT, GC and GT, respectively. (c) Gel images showing the 13-mer and 10-mer products released from the bottom-strand (opposite to the lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer A and 10mer G represent the \([5'\text{-}32P]\)-labeled standard ODNs \(5'-AATTATAGCA-3'\) and \(5'-AATTATAGCG-3'\), respectively.
Figure S14. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of $n$-propyl and $n$-butyl phosphotriester in wild-type AB1157 cells without SOS induction (WT), or wild-type AB1157 (SOS-WT) and isogenic cells that are deficient in Pol II, Pol IV, or PolV ($\Delta$Pol II-SOS, $\Delta$Pol IV-SOS, $\Delta$Pol V-SOS) and all three polymerases (TKO-SOS) with SOS induction. (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. ‘SAP’ and ‘PNK’ represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT represents the $[5'-^{32}P]$-labeled standard ODNs 5'-GGCTGCTAT-3'.

\[
\begin{align*}
\text{WT} & \\
\text{WT-SOS} & \\
\text{ΔPol II-SOS} & \\
\text{ΔPol IV-SOS} & \\
\text{ΔPol V-SOS} & \\
\text{TKO-SOS} & \\
\end{align*}
\]
Figure S15. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of alkylphosphotriester in triple knockout (TKO) AB1157 cells without SOS induction (WT). (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. ‘SAP’ and ‘PNK’ represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where where 10mer TT, 10mer GC, and 10mer GT represent the [5'-32P]-labeled standard ODNs 5'-GGCMNGCTAT-3', with ‘MN’ being TT, GC and GT, respectively.
Figure S16 Restriction enzymes digestion and LC-MS methods for identifying the mutation products of alkyl phosphotriesters in E. coli cells. X in 22-mer DNA strand designates the location of the alkyl phosphotriester. BbsI and MluCI digestion sites are indicated with broken arrows. Only partial sequences of the PCR products for the lesion-containing genome are shown, and the PCR products of the competitor genome are not shown.

For LC-MS analysis, the PCR products were digested with BbsI or MluCI prior to digestion with shrimp alkaline phosphatase, where the [32P]-labeling step involving the use of T4 polynucleotide kinase (T4 PNK) was omitted.
Figure S17. Higher-resolution “ultra-zoom scan” ESI-MS of the restriction fragments for the PCR products from the replication of $S_p$- and $R_p$-methyl phosphotriester-bearing single-stranded M13 genomes in SOS-induced wild-type AB1157 cells. Displayed are the [M – 3H]$^3^-$ ions for the lesion-containing top strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for $S_p$-methyl phosphotriester are shown in Figure S17.
Figure S18. LC-MS and MS/MS for the identification of restriction fragments of PCR products. MS/MS for the [M – 3H]⁻ ions of (a) 10 mer TT (non-mutagenic product), (b) 10 mer GT (TT → GT mutation), and 10 mer GC (TT → GC mutation).
Figure S19. Higher-resolution “zoom scan” ESI-MS of the restriction fragments for the PCR products from the replication of $S_p$- and $R_p$-methyl phosphotriester bearing single-stranded M13 genomes in SOS-induced wild-type AB1157 cells. Displayed are the [M – 3H]$^{3-}$ ions for the bottom strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for $S_p$-methyl phosphotriester are shown in Figures S19.
Figure S20. LC-MS and MS/MS for the identification of restriction fragments of PCR products. MS/MS for the $[M - 3H]^{-3}$ ions of (a) 10 mer A (non-mutagenic product and TT→GT mutation), (b) 10 mer G (TT→GC mutation),
Figure S21. Gel image showing the time-dependent CRAB assay of 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome. (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. ‘SAP’ and ‘PNK’ represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT represent the \([5'-32P]-labeled standard ODNs\ 5'-GGCMNGCTAT-3'\), with ‘MN’ being TT, respectively. (c) Bypass efficiencies of alkyl phosphotriester lesions in wild-type AB1157 Escherichia coli strains. The data represent the means and standard deviations of results from three independent replication experiments. **, \(0.001 < p < 0.01\). The \(p\) values were calculated by using an unpaired two-tailed t-test and referred to comparisons with the data obtained for control cells.