



OPEN

Replicative bypass studies of L-deoxyribonucleosides in Vitro and in *E. coli* cell

Yuhe Kan^{1,4,5,6}, Zhaoyang Jin^{1,6}, Yongqi Ke¹, Dao Lin¹, Liang Yan¹, Li Wu^{1,2}✉ & Yujian He^{1,2,3}✉

L-nucleosides were the most important antiviral lead compounds because they can inhibit viral DNA polymerase and DNA synthesis of many viruses, whereas they may lead to mutations in DNA replication and cause genomic instability. In this study, we reported the replicative bypass of L-deoxynucleosides in recombinant DNA by restriction enzyme-mediated assays to examine their impact on DNA replication in vitro and in *E. coli* cells. The results showed that a template L-dC inhibited Taq DNA polymerase reaction, whereas it can be bypassed by Vent (exo⁻) DNA polymerase as well as in cell replication, inserting correct nucleotides opposite L-dC. L-dG can be bypassed by Taq DNA polymerase and in *E. coli* cells, maintaining insertion of correct incoming nucleotides, and L-dG induced mutagenic replication by Vent (exo⁻) DNA polymerase. In contrast, L-dA can induced mutagenic replication in vitro and in *E. coli* cells. MD simulations were performed to investigate how DNA polymerase affected replicative bypass and mutations when D-nucleosides replaced with L-nucleosides. This study will provide a basis for the ability to assess the cytotoxic and mutagenic properties of the L-nucleoside drugs.

Commonly, D-nucleotides were the basic building blocks of natural nucleic acids in all known forms of life. To our knowledge, genetic polymers (DNA/RNA) with the same chirality were functional as templates in replication, transcription and translation. DNA polymerase composed of L-amino acids has strict chiral selectivity for D-deoxynucleotides. However, studies have reported that the template-directed polymerization of DNA can be catalysed by a chemically synthesized D-amino acid polymerase on an L-DNA template¹. In addition, amount studies indicated modified or chiral nucleotides can be incorporated into DNA catalysed by various polymerases²⁻⁷. For example, L-nucleosides can be functioned as specific substrates for herpes simplex 1 (HSV1) thymidine kinase (TK), but not human TK⁸. Next, phosphorylated nucleoside was recognized by viral polymerases, which may provide clear clues for designing safe and effective antiviral drugs. L-nucleosides, such as β -L-thymidine (L-T, telbivudine), β -L-2'-deoxycytidine (L-dC) and β -L-2'-deoxyadenosine (L-dA) had most potent, selective and specific antiviral activity against HBV during nucleic acid synthesis^{9,10}. Although L-nucleosides can be recognized by viral enzymes rather than host (bacterial or human) enzymes, once L-nucleosides were recognized by host translesion synthetase (i.e. DNA polymerase IV and V for bacterial or DNA polymerase ζ , κ , η , and ι for human) and incorporated into DNA, the tautomeric equilibrium was re-established, which might bring about cytotoxic and mutagenic effects. Wherein, L-nucleosides could be regarded as chemical mutagen leading to DNA damage.

Although various DNA repair systems (such as BER, NER, TLS, MMR, SSBR, HR, NHEJ and DNA interstrand crosslink repair pathway) were to independently or synergistically repair damaged DNA¹¹, erroneous DNA repair resulted in mutations or chromosomal aberrations¹². Additionally, when the DNA strand encounters damage during replication and the replication steps, the translesion synthesis (TLS) polymerases in cells can bypass DNA lesions to continue DNA replication. However, it might led to the mutation¹³. Fatal gene mutations caused by DNA damage have attracted more and more attention. For example, 5-Guanidino-4-nitroimidazole (NI) induced mainly G \rightarrow T, G \rightarrow A and some G \rightarrow C mutation and strongly blocked replication in *E. coli* cells¹⁴. Two strong mutagenic agents, stereoisomers of spiroiminodihydantoin (Sp1 and Sp2), caused mainly G \rightarrow C (72%

¹School of Chemical Sciences, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China. ²State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, People's Republic of China. ³School of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China. ⁴School of Life Sciences, Inner Mongolia University, Hohhot 010021, Inner Mongolia, People's Republic of China. ⁵Qilu Pharmaceutical (Inner Mongolia) CO., LTD., Hohhot 010080, Inner Mongolia, People's Republic of China. ⁶These authors contributed equally: Yuhe Kan and Zhaoyang Jin. ✉email: wuli@ucas.ac.cn; heyujian@ucas.ac.cn

for Sp1 and 57% for Sp2) and G → T (27% for Sp1 and 41% for Sp2) mutations with strong replication inhibition (91%)¹⁵. 3-methyldeoxycytidine (m3C) and 3-ethyldeoxycytidine (e3C) were 30% mutagenic in SOS⁻ cells, but rising to 70% in SOS⁺ cells¹⁶. Aflatoxin B1-N7-dG adduct (AFB1-N7-dG) caused mutation frequency of 4%, which dominantly give rise to G → T mutation¹⁷. O⁴-alkyl dT can induce substantial frequencies T → C transition mutation and moderately inhibited DNA replication with the bypass efficiencies ranging from 20 to 33% in HEK293T cells¹⁸. Previously, we reported that L-thymidine induced self-repair by B family polymerases¹⁹ whereas L-T → dA (13%), T (22%), dG (19%) and dC (46%) mutations in *E. coli* cells^{20,21}. The results inspire us to explore the fidelity and cytotoxicity for L-dC, L-dG and L-dA during replication, facilitating understanding of the chiral transfer in the process of nucleic acid evolution.

In the current study, the recombinant pUC19 vector containing L-dA, L-dG and L-dC were constructed and subsequently transfected into *E. coli* cells. The replicative bypass of L-dC, L-dG and L-dA in recombinant DNA in vitro and in *E. coli* cells was to evaluate the mutagenic effect of L-deoxynucleotides during DNA replication using restriction enzyme-mediated assays^{20,21}. The result showed that the bypass efficiency of L-nucleosides in *E. coli* cells was relatively high compared to Taq DNA polymerases or Vent (exo⁻) DNA polymerases, and even the replicative bypass of L-dC was completely inhibited by Taq DNA polymerase. In addition, L-dA induced more mutation types in cells, including purine → purine conversion and purine → pyrimidine transversion. This was similar to our previously reported results for L-T in *E. coli* cells replication. However, L-dG did not cause a mismatch in *E. coli* cells but did mutate with Vent (exo⁻) DNA polymerases. Effect of L-nucleosides on replicative bypass and mutations with DNA polymerase were further investigated by using MD simulations. Our study will provide a basis for the ability to assess the cytotoxic and mutagenic properties of the L-nucleoside drugs.

Results

Construction of L-deoxynucleoside-bearing plasmids for DNA replication. The aim of this study was to elucidate comprehensively the impact of L-nucleosides on the bypass and fidelity of DNA replication in vitro and in *E. coli* cells. To this end, we first synthesized a series of site-specific L-nucleoside-carrying ODNs for each type of L-nucleosides and a complement strand COM20 (Table 1). For example, nucleoside T, dG, dC, dA on *Pst*I recognition sequence within STD strand (5'-AATTCTCCTGCAGGCTAGCACTGA-3') was substituted by L-dG, respectively, of which "lesion-containing strands" were named G13 (5'-AATTCTCCG^LGCAGGCTAGCACTGA-3'), G14 (5'-AATTCTCCTG^LCAGGCTAGCACTGA-3'), G15 (5'-AATTCTCCTGG^LAGGCTAGCACTGA-3') and G16 (5'-AATTCTCCTGCG^LGGCTAGCACTGA-3'). Then, the complement strand (COM20) was annealing with lesion-containing strands to give original lesion-containing double strand, and a C:C mismatch was inserted around the recognition site of *Pst*I of the double strand, which was to distinguish the replication products of lesion-containing strands from the natural strands as noted previously²¹. To assess the bypass efficiencies and mutation frequencies of the L-nucleosides, we ligated the aforementioned inserts containing L-nucleosides into pUC19 genome and performed a restriction enzyme-mediated assays to examine how these lesions inhibit DNA replication and induce mutations in vitro and in *E. coli* cells (Fig. 1). Prior to assessing the cytotoxicity and fidelity induced by L-nucleosides, the digestion products corresponding to 24-mer ODNs cleaved by *Eco*RI and *Hind*III restriction endonuclease were detected, indicating successful construction of L-nucleosides-containing genomes during ligation process. Then, we employed two restriction enzymes, i.e. *Pst*I and *Hind*III, to digest the PCR products of the progeny genome, affording 20mer ODN fragments from the lesion-containing or control genome (STD) (Fig. 1). The released ODNs were subjected to denature PAGE analyses to identify the replication products, as described elsewhere²¹. As illustrated in Fig. 1, in this vein, if the correct nucleotide opposite L-nucleosides was incorporated during DNA replication, which restore *Pst*I recognition sequence. *Pst*I restriction enzyme was allowed for the selective digestion of the progeny genomes emanating from the replication of the strand that initially contained the L-nucleosides due to C:C mismatch²¹. The complementary opposite strand (COM20, 5'-CTGCAC-3') in a recombinant vector produced a strand of 3'-CACGTC-5'/5'-CTGCAC-3' in replication, which is different from restriction digestion fragments of 3'-GACGTC-5'/5'-CTGCAG-3' formed from lesion-containing strand if a corrected nucleotide was incorporated opposite the lesion site. The amplified strand from COM20 could not be cleaved by *Pst*I and thus was not detected by analyzing 20mer bands in the gel. In this respect, sequential digestion of L-nucleoside-containing strands with *Hind*III and *Pst*I would give a 20mer fragment (5'-AGCTTCAGTGCTAGCCTGCA-3') on 20% denature PAGE gel if incorporation of nucleotides opposite L-nucleoside restored the *Pst*I recognition sequence. On the contrary, no band was present corresponding to 20mer fragment due to disrupting *Pst*I recognition sequence. To take a concrete example, replication products (5'-CTGCAG-3') were found to be further digested by *Pst*I when dA was incorporated opposite L-dG within G13 (5'-CG^LGCAG-3') during DNA replication. However, the corresponding product harboring an L-dG → dC, dA or dG mutation at the lesion site, i.e. 5'-CCGCAG-3', 5'-CAGCAG-3' or 5'-CGGCAG-3', could not be digested by *Pst*I. Similarly, for G14 (5'-CTG^LCAG-3'), G15 (5'-CTGG^LAG-3') and G16 (5'-CTGCG^LG-3'), restriction digestion would occur when only dC, dG and T were incorporated opposite L-dG, respectively. Thus, with the combination of the two enzyme digestion procedures, we were able to distinguish unequivocally the four potential types of replication products, and by monitoring the products from the strand of 20mer ODNs (5'-AGCTTCAGTGCTAGCCTGCA-3'), we could quantify the bypass efficiency and mutation frequencies.

The bypass efficiencies of the L-nucleosides were then calculated from the ratio of the combined intensities of bands observed for the 20mer products from the lesion-containing genome over the intensity of the 20mer product from the control genome STD. With the use of this method, we were able to determine quantitatively the degrees to which the L-dC, L-dG and L-dA lesions inhibit DNA replication and induce mutations in *E. coli* cells. Although *Pst*I-digestion can be conducted for pUC19 self-ligation plasmid, only 16-mer band on PAGE gel will not interfere assessment of incorporation frequency and bypass efficiency.

Name	Sequence
C13	5'-AATTCTCC <u>C</u> ^L GCAGGCTAGCACTGA-3'
C14	5'-AATTCTCCT <u>C</u> ^L CAGGCTAGCACTGA-3'
C15	5'-AATTCTCCTG <u>C</u> ^L AGGCTAGCACTGA-3'
C16	5'-AATTCTCCTGC <u>C</u> ^L GGCTAGCACTGA-3'
50	5'-AATTCTCC <u>CG</u> CAGGCTAGCACTGA-3'
51	5'-AATTCTCCT <u>CC</u> CAGGCTAGCACTGA-3'
STD	5'-AATTCTCCTG <u>C</u> AGGCTAGCACTGA-3'
53	5'-AATTCTCCTGC <u>CC</u> GGCTAGCACTGA-3'
G13	5'-AATTCTCC <u>G</u> ^L GCAGGCTAGCACTGA-3'
G14	5'-AATTCTCCT <u>G</u> ^L CAGGCTAGCACTGA-3'
G15	5'-AATTCTCCTGG <u>G</u> ^L AGGCTAGCACTGA-3'
G16	5'-AATTCTCCTGCG <u>G</u> ^L GGCTAGCACTGA-3'
54	5'-AATTCTCC <u>GG</u> CAGGCTAGCACTGA-3'
56	5'-AATTCTCCTGG <u>AG</u> GCTAGCACTGA-3'
57	5'-AATTCTCCTGCG <u>GG</u> GCTAGCACTGA-3'
A13	5'-AATTCTCC <u>A</u> ^L GCAGGCTAGCACTGA-3'
A14	5'-AATTCTCCT <u>A</u> ^L CAGGCTAGCACTGA-3'
A15	5'-AATTCTCCTGA <u>A</u> ^L AGGCTAGCACTGA-3'
A16	5'-AATTCTCCTGCA <u>A</u> ^L GGCTAGCACTGA-3'
58	5'-AATTCTCC <u>AG</u> CAGGCTAGCACTGA-3'
59	5'-AATTCTCCT <u>AC</u> CAGGCTAGCACTGA-3'
60	5'-AATTCTCCTGA <u>AG</u> GCTAGCACTGA-3'
COM20	5'-AGCTTCAGTGCTAGCCTGCA <u>C</u> GAG-3'
C12F	5'-TTAGCTCACTCATTAGGCACCC-3'
C12R	5'-GGGCCTCTTCGCTATTACGC-3'

Table 1. Single-strand oligonucleotides used in this study. Underlined sequences : nucleoside (red site) of *Pst*I recognition site was replaced by L-nucleosides and natural nucleosides, respectively.

The bypass efficiencies and mutation frequency were assessed by our previous method²¹. Simply, the bypass efficiencies for the L-nucleoside-carrying genomes were then normalized against that for the control genome. Then, the bypass efficiencies of the L-nucleoside were calculated by comparing intensities of 20mer bands observed from the L-nucleoside-containing genome over the intensity of the control genome. The mutation frequency of L-nucleosides was obtained based on above bypass efficiencies, calculating by ratio each bypass efficiency (i.e. L-dG → T) to total bypass efficiency, i.e. L-dG → T, dG, dC and dA. Thus, by monitoring the products and evaluating bypass efficiency and mutation frequency from the replication of L-nucleosides-containing strand, we were able to study the effect of L-nucleosides on DNA replication in vitro and in cells^{20,21}.

E. coli cells are the most widely studied bacteria that live in the intestines of people and animals and contain five types of DNA polymerases contributing DNA replication. DNA polymerase I is the most abundant and high-fidelity polymerase. Certainly, DNA polymerase III (family C) is also the major replicase²². And DNA polymerase II has 3'-5' nucleic acid exonuclease activity and restarts replication after it has been halted by DNA strand

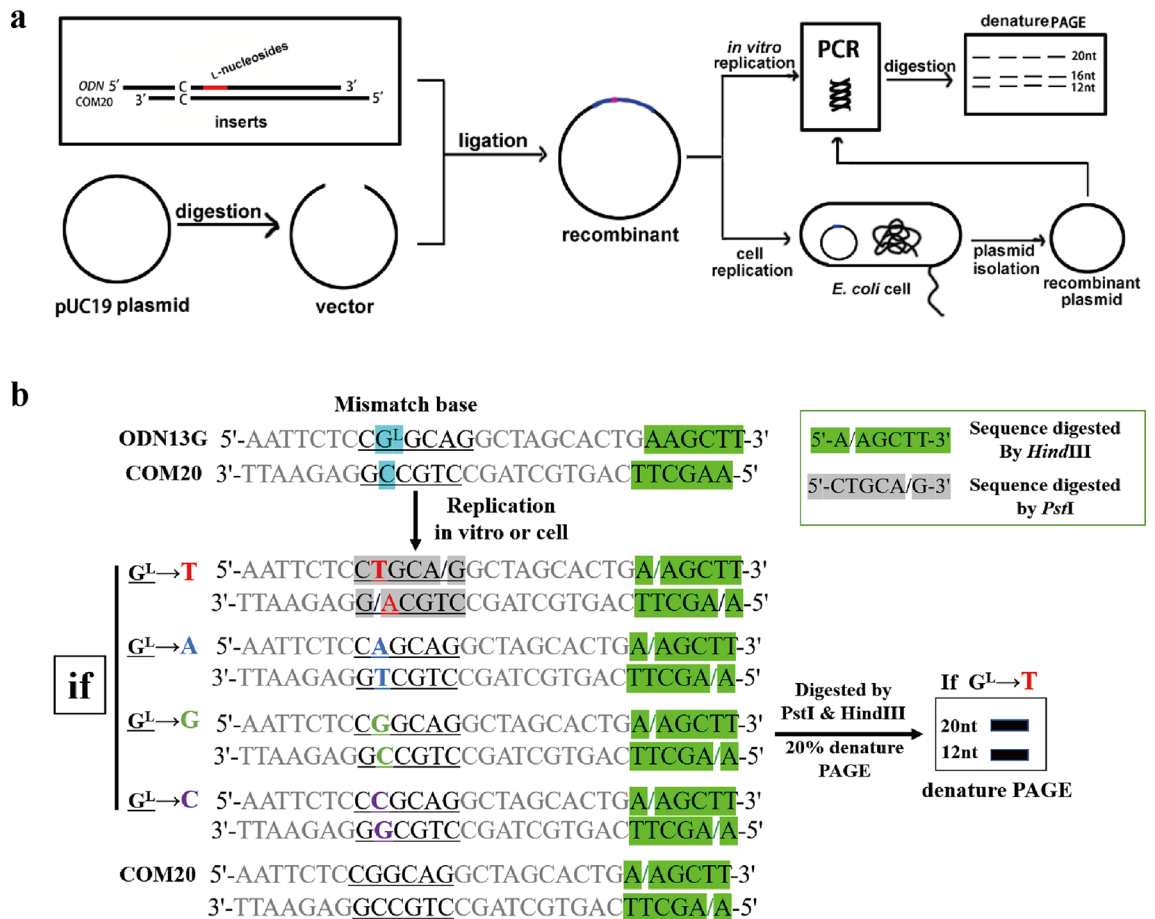


Figure 1. Schematic diagram of restriction endonuclease mediated assay for study on replication of DNA containing lesion.

damage. DNA polymerase I and DNA polymerase II belongs to A and B family polymerases, respectively²³. It was reported that A family and B family polymerases had significantly different effect during DNA replication¹⁹ which inspired us to investigate replicative bypass of L-deoxyribonucleosides using A and B family polymerases. Due to commercial availability, Taq DNA polymerase and Vent (exo⁻) DNA polymerase, which respectively represented A and B family polymerases, were used for in vitro DNA replication.

Previously, we constructed four L-T-containing recombinants using T13 (5'-CT^LGCAG-3'), T14 (5'-CTT^LCAG-3'), T15 (5'-CTG^LAG-3') and T16 (5'-CTGCT^LG-3') to assess the bypass efficiencies and mutation frequencies of the L-T lesions in vitro and in cells. By a restriction enzyme-mediated assay provided by us, we showed the L-T was bypassed in *E. coli* cells (99%), and the bypass efficiency in cells was higher than Taq DNA polymerases (71%) and Vent (exo⁻) DNA polymerases (47%). In addition, the cell replication across L-T lead to L-T → dA (13%), T (22%), dG (19%) and dC (46%) mutations²¹. Next, we hope to gain a comprehensive understanding of the impact of L-deoxynucleotides on DNA replication in vitro and in *E. coli* cells.

Impacts of L-dC on the bypass efficiency and mutation frequency of DNA replication in vitro and in *E. coli* cells. To understand the impact of L-dC lesions on DNA replication in vitro and in *E. coli* cells, we constructed four L-dC-containing recombinants using C13 (5'-CC^LGCAG-3'), C14 (5'-CTC^LCAG-3'), C15 (5'-CTGC^LAG-3') and C16 (5'-CTGCC^LG-3') to assess the bypass efficiencies and mutation frequencies of the L-dC lesions. In this assay, DNA containing L-nucleosides was directly replicated by Taq DNA polymerase and Vent (exo⁻) DNA polymerases for in vitro study. For study of cell replication, the recombinants containing L-nucleosides were transformed into *E. coli* cells, following by PAGE analysis for digested PCR products of plasmid genome.

Two restriction enzymes, *HindIII* and *EcoRI*, were employed to digest replication products in vitro and in cells from four lesion-containing genomes, only 24mer bands was observed on denature PAGE gel, which indicated successful construction of L-C-bearing plasmids (Fig. 2). In addition, PCR products from recombinants containing L-nucleosides were digested with *EcoRI* and *HindIII*, no other bands (>24-mer or <24-mer ODNs) were observed on PAGE gel, which indicated no nucleotide insertion and deletions either in vitro or in *E. coli* cells.

The DNA fragments corresponding to 16-mer restriction products digested by *HindIII* and *PstI* were detectable, likely due to fragment generation by the self-ligated pUC19 plasmid. As expected, PCR products of STD

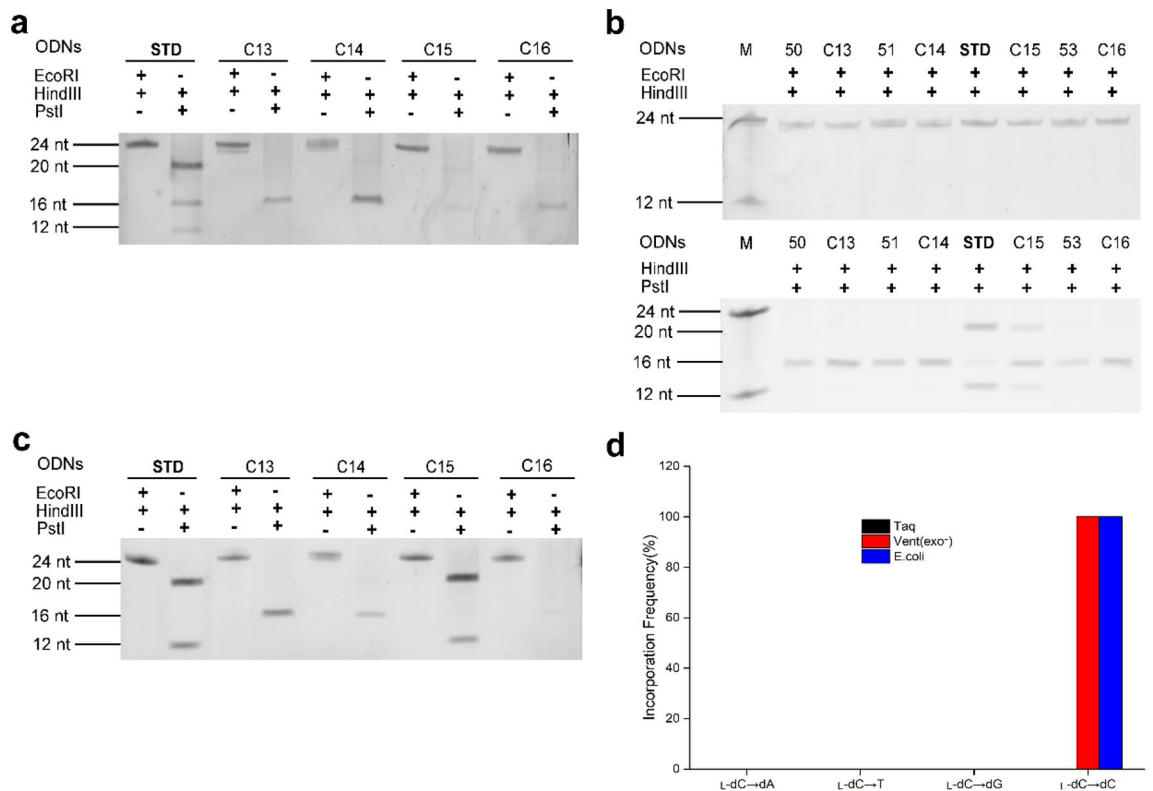


Figure 2. PAGE analysis of restriction fragments of replication products arising from the L-dC lesion-bearing genome (**a**) Taq DNA polymerase (**b**) Vent (exo⁻) DNA polymerase (**c**) *E. coli* cells and (**d**) its mutation frequency.

(Fig. 2a (Fig. S1), lane 2; Fig. 2b (Fig. S2), lane 6 and Fig. 2c (Fig. S3), lane 2) were digested by *PstI* and *HindIII* to produce 20mer fragments. Because the *PstI* digestion recognition sequence within insert sequence was not scrambled, so it can be digested by *PstI* and *HindIII* to generate products corresponding to 20mer fragments. In contrast, *PstI* recognition sequence with one nucleoside substituted (50, 51 and 53) will not be digested by *PstI*, which can't produce 20mer ODNs (Fig. 2b (Fig. S2), lane 2, 4 and 8) on PAGE gel due to disorder of its *PstI* recognition sequence.

Figure 2a (Fig. S1) showed that replication products from recombinants, of which T, dG, dC and dA on STD (5'-CTGCAG-3') was replaced by L-dC, was digested by *PstI* and *HindIII* under catalysis of Taq DNA polymerase. If dA, dC, dG and T were incorporated opposite L-dC of C13, C14, C15 and C16, respectively, its replication products would be digested by *PstI* due to restores of its recognition sequences, and thus 20mer band on PAGE gel was observed. Actually, no 20mer fragment was detectable on gel for all the above experimental groups (Fig. 2a (Fig. S1), lane 4, 6, 8 and 10), showing dA, dC, dG and T was not incorporated opposite L-dC during DNA replication. These results indicated that the replication of DNA containing L-dC was inhibited by Taq DNA polymerase.

Different from what happened on Taq DNA polymerase, a 20mer oligonucleotide was observed for C15 (Fig. 2b (Fig. S2), lane 7) when catalyzed by Vent (exo⁻) DNA polymerase, which indicated L-dC paired with dG in replication products. No bands corresponding to 20mer fragments were present for C13, C14 and C16 (Fig. 2b (Fig. S2), lane 3, 5 and 9), showing no dA, dC and T was incorporated opposite L-dC during DNA replication. Prior to normalizing band intensities of the 20mer fragment to that of the 24-mer fragment, the bypass efficiencies of the L-dC lesions were then calculated from the ratio of the combined intensities of bands observed for the 20mer products from the lesion-containing genome over the intensity of the 20mer product from the control genome (STD). And incorporation frequencies were calculated from the percentage of a mutagenic product relative to all possible products from the control genome. The bypass efficiency of 39% was obtained (Supplementary Table S1 and S2) and L-dC induced no mutagenic replication catalysed by Vent (exo⁻) DNA polymerase.

For replication in *E. coli* cells, the result was same as in vitro replication catalysed by Vent (exo⁻) DNA polymerase. Only one band corresponding to 20mer fragment was obtained for C15 (Fig. 2c (Fig. S3), lane 8), which indicated L-dC: dG pairing formed in *E. coli* cells, with bypass efficiency being 81% (Supplementary Table S1 and S2). No bands corresponding to 20mer fragment were present for C13, C14 and C16 (Fig. 2c (Fig. S3), lane 4, 6 and 10), showing no dA, dC and T was incorporated opposite L-dC in cell replication.

Hence, these results revealed that L-dC completely block the DNA replication machinery under catalysis of Taq DNA polymerases, whereas L-dC can be bypassed by Vent (exo⁻) DNA polymerases or in *E. coli* cells. The bypass efficiency in *E. coli* cells (81%) was higher than that obtained under catalysis of Vent (exo⁻) DNA polymerases (39%). What's more, no mutagenic replication was found for L-dC-containing DNA in vitro and in *E. coli* cells.

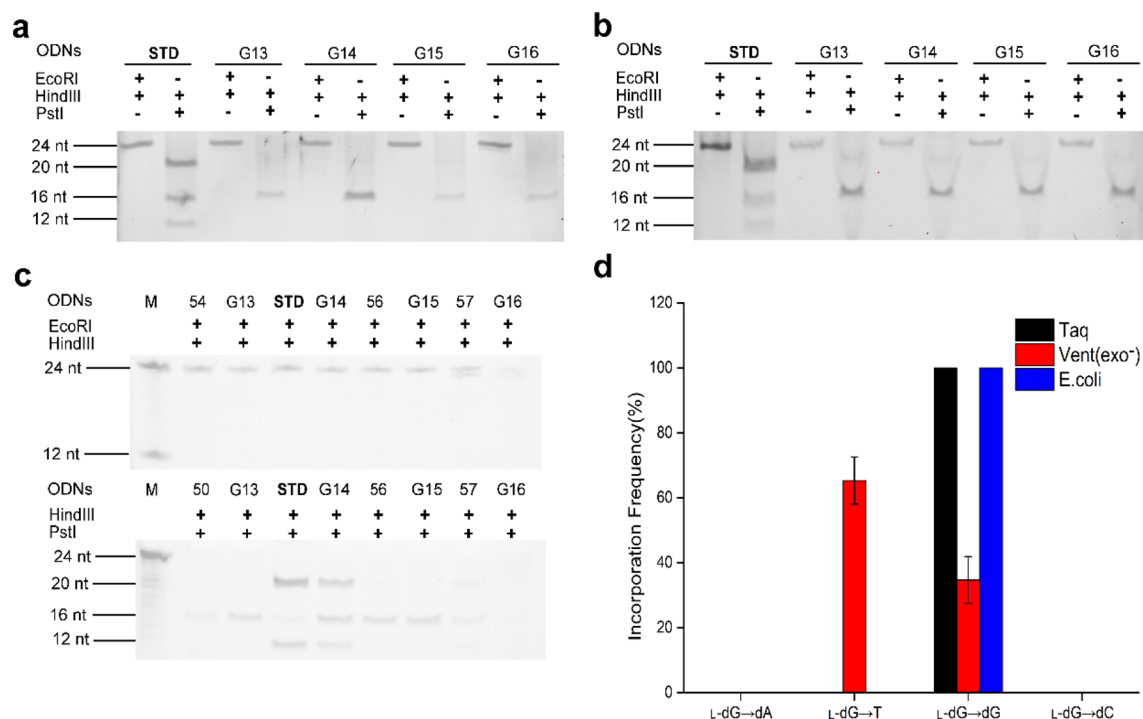


Figure 3. PAGE analysis of restriction fragments of replication products arising from the L-dG lesion-bearing genome (a) Taq DNA polymerase (b) Vent (exo⁻) DNA polymerase (c) *E. coli* cells and (d) its mutation frequency.

Impacts of L-dG on the bypass efficiency and mutation frequency of DNA replication in vitro and in *E. coli* cells. To assess the bypass efficiencies and mutation frequencies of the L-dG lesions, four L-G-containing recombinants were parallelly constructed by incorporation of G13 (5'-CG^LGCAG-3'), G14 (5'-CTG^LCAG-3'), G15 (5'-CTGG^LAG-3') and G16 (5'-CTGCG^LG-3').

As shown in Fig. 3, 24mer bands on denature PAGE gel corresponding to the digested products by *HindIII* and *EcoRI* were present, indicating successful construction of L-G-bearing plasmids. In addition, undesirable 16-mer ODNs from *PstI* and *HindIII* digestion was observed, suggesting there was self-ligated pUC19 plasmid. For positive restriction endonuclease assay, 20mer fragments from *PstI* and *HindIII* digested PCR products of STD (Fig. 3a (Fig. S4), lane 2; Fig. 3b (Fig. S5), lane 2 and Fig. 3c (Fig. S6), lane 4) were obtained, as expected.

However, for the reaction system catalysed by Taq DNA polymerases, a 20-mer oligonucleotide band was found for G14 (Fig. 3a (Fig. S4), lane 6) when the replication products were incubated with *PstI* and *HindIII*, which suggested that replication of G14 with replacing G with L-dG resulted in L-dG: dC pairing. Moreover, L-dG induced L-dG → dG substitution with bypass efficiency of 19% (Supplementary Table s S1 and S2). By comparison, 20mer band was absent on gel for G13, G15 and G16 (Fig. 3a (Fig. S4), lane 4, 8 and 10), in which L-dG was substituted for T, C and A of 5'-CTGCAG-3', respectively. The result indicated that no dA, dG and T was incorporated opposite L-dG during DNA replication. That is, L-dG did not result in a mutation in the reaction system catalysed by Taq DNA polymerases.

In the reaction system catalysed by Vent (exo⁻) DNA polymerases, 20-mer DNA fragments of digested replication products can be detected in recombinant plasmids G14, where L-dG substituted for G of 5'-CTGCAG-3' (Fig. 3b (Fig. S5), lane 6), which indicated that dC was incorporated opposite L-dG. Additionally, 20-mer DNA fragments was observed for G13 (Fig. 3b (Fig. S5), lane 4), indicated that dA was incorporated opposite L-dG of 5'-CG^LGCAG-3', that is, L-dG induced L-dG → T transversion mutation besides L-dG → G. The bypass efficiency of L-dG was 12% and the mutation frequency of L-dG → T was 65% (Supplementary Table s S1 and S2). In contrast, no bands corresponding to 20mer fragments was found for G15 and G16 (Fig. 3b (Fig. S5), lane 8 and 10) proved that dG and T was not incorporated opposite L-dG during DNA replication catalysed by Vent (exo⁻) DNA polymerase.

For replication investigation in *E. coli* cells, only in the case of G14 was a 20-mer oligonucleotide band found on gel (Fig. 3a (Fig. S4), lane 6) when the replication products were incubated with *PstI* and *HindIII*. This indicated that replication of DNA containing L-dG resulted in L-dG: dC pairing, with bypass efficiency being 82% (Fig. 3c (Fig. S6), lane 5; Supplementary Tables S1 and S2), whereas dA, dG and T was not incorporated opposite L-dG during DNA replication (Fig. 3c (Fig. S6), lane 3, 7 and 9).

Hence, for study on L-dG-containing DNA replication, Taq DNA polymerase and Vent (exo⁻) DNA polymerase were inhibited by L-dG, and the bypass efficiencies were decreased to 19% and 12%, respectively. But, in cell replication assay resulted in apparent increase in bypass efficiency (82%) relative to in vitro replication. Mutagenic replication was found in vitro but not in cells, which may attribute to repair machinery and related proteins in cells.

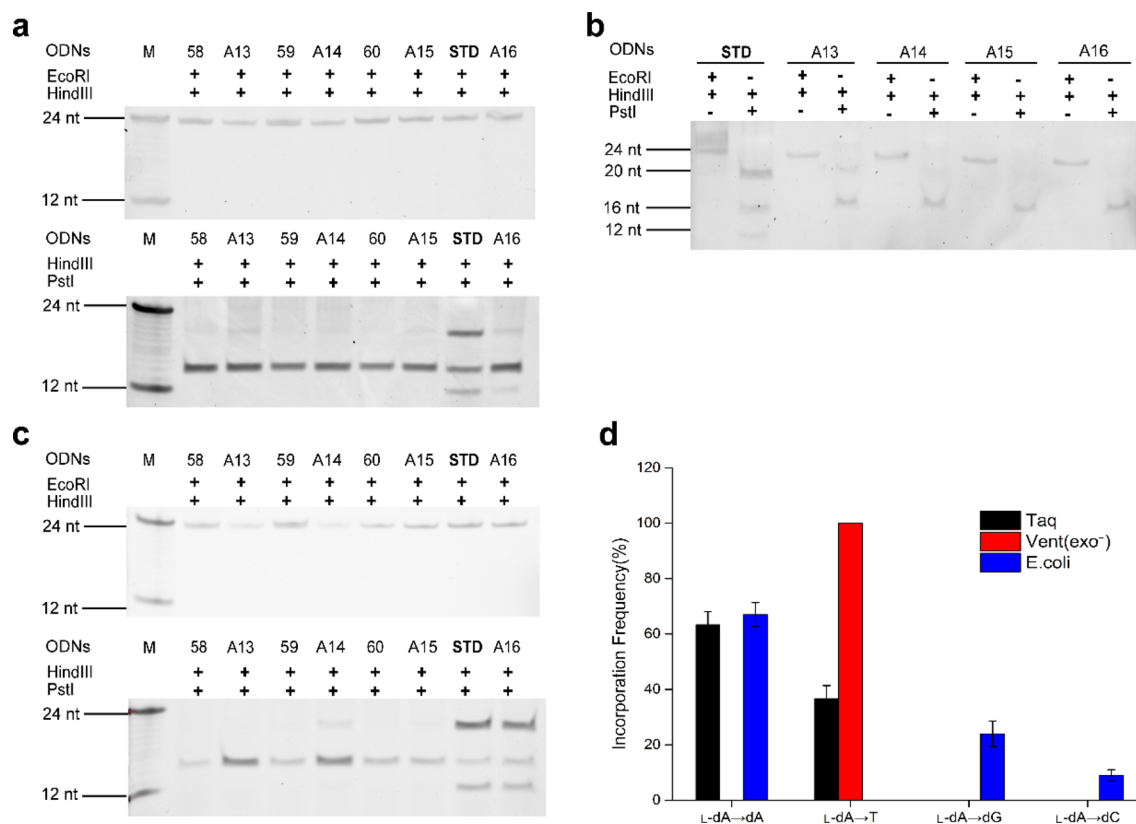


Figure 4. PAGE analysis of restriction fragments of replication products arising from the L-dA lesion-bearing genome (a) Taq DNA polymerase (b) Vent (exo⁻) DNA polymerase (c) *E. coli* cells and (d) its mutation frequency.

Impacts of L-dA on the bypass efficiency and mutation frequency of DNA replication in vitro and in *E. coli* cells. We constructed four L-dA-containing recombinants using A13 (5'-CA^LGCAG-3'), A14 (5'-CTA^LCAG-3'), A15 (5'-CTGA^LAG-3') and A16 (5'-CTGCA^LG-3') to assess the bypass efficiencies and mutation frequencies of the L-dA lesions. Figure 4 indicated successful L-A-bearing plasmids and self-ligated pUC19 plasmid. As expected, *PstI* and *HindIII* digestion products corresponding to 20mer ODNs was obtained for STD (Fig. 4a (Fig. S7), lane 8; Fig. 4b (Fig. S8), lane 2 and Fig. 4c (Fig. S9), lane 8) revealed that the *PstI* digestion recognition sequence within insert sequence was not destroyed. In contrast, no 20-mer ODNs was observed for 58, 59 and 60 (Fig. 4a (Fig. S7) and 4c (Fig. S9), lane 2, 4 and 6) which resulted from destroyed *PstI* recognition sequence.

The effect of L-dA on DNA replicative bypass and fidelity was investigated by Taq and Vent (exo⁻) DNA polymerases as well as in cells. Figure 4a (Fig. S7) showed experimental results in the case of Taq polymerase. Replication products of A13 and A16 were digested by restriction endonuclease *PstI* and *HindIII* to produce 20-mer fragments (Fig. 4a (Fig. S7), lane 3 and lane 9), whereas no 20mer restriction-digested PCR fragments of A14 and A15 were detected (Fig. 4a (Fig. S7), lane 5 and lane 7). This indicated that dA and T rather than dC and dG were incorporated opposite L-dA during DNA replication. Furthermore, Taq DNA polymerase bypassed L-dA with a low bypass efficiency of 16%. L-dA paired with T with frequency being 63% and induced L-dA → T transversion at a frequency of 37% (Supplementary Tables S1 and Table S2).

Different from Taq DNA polymerase, only one band corresponding to digested 20-mer DNA fragment of PCR product was detected in recombinant plasmids A13 containing L-dA lesion when catalysed by Vent (exo⁻) DNA polymerases (Fig. 4b (Fig. S8), lane 4), indicate that L-dA was strongly miscoding, and it induced completely L-dA → T transversion mutation catalysed by Vent (exo⁻) DNA polymerases, with bypass efficiency being 20% (Supplementary Tables S1 and Table S2).

In the study of the replication of DNA containing L-dA in cells, PCR products of A14, A15, and A16 were digested by restriction endonuclease *PstI* and *HindIII* to produce 20-mer fragments (Fig. 4c (Fig. S9), lane 5, lane 7 and lane 9), verifying the L-A → dA, dG, and dC mutation occurred during cell replication, and L-dA did not induce L-dA → T mutation (Fig. 4c (Fig. S9), lane 3). In addition, the results showed that bypass efficiency of L-dA in *E. coli* cells was up to 74%, which was 3–5 times as high as that for Taq DNA polymerases (16%) and Vent (exo⁻) DNA polymerases (20%). Also, L-A → dA, dG and dC with mutation frequency being 67%, 24% and 9% (see Supplementary Table S1 and Table S2 online).

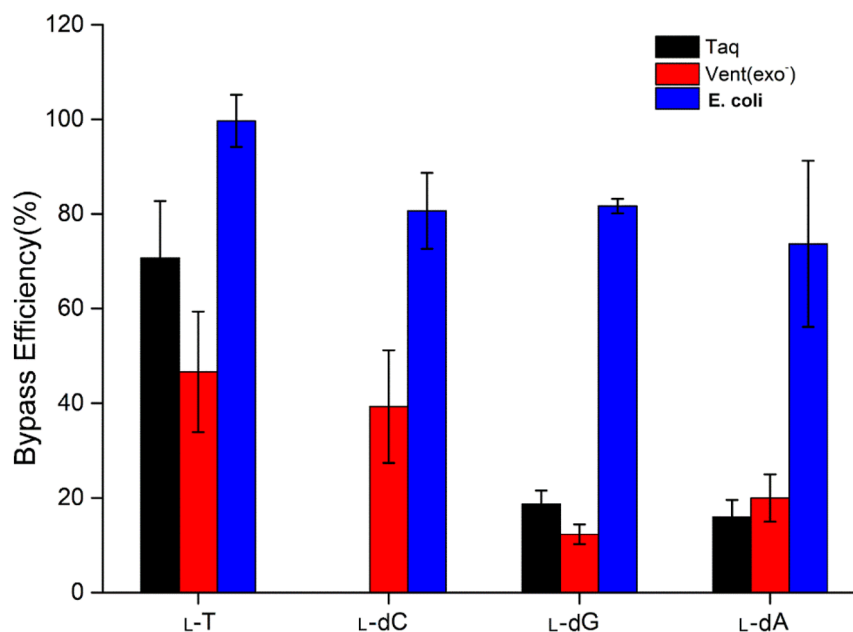


Figure 5. The bypass efficiency of replication of DNA containing L-nucleosites in cells and in vitro solution.

Discussion

We examined previously the cytotoxic and mutagenic properties of L-T in *E. coli* as well as several DNA polymerases^{20,21}, which revealed distinctly different bypass efficiencies and mutagenic properties in different replication systems. This study was systematic interrogation about the impact of the L-nucleosides on the bypass efficiency and fidelity of DNA replication. Our results revealed that L-nucleosides could be readily bypassed by DNA replication machinery in *E. coli* cells (see Fig. 5 and Supplementary Table S2 online). Moreover, relative to Taq DNA polymerase and Vent (exo⁻) DNA polymerases, the relatively high efficiencies in bypassing L-nucleosides were observed in *E. coli* cells, where bypass efficiencies of L-dC, L-dG and L-dA being 81%, 82% and 74%, respectively. This finding paralleled the observations made in *E. coli* cells for L-T, which showed elevated bypass efficiency of L-T (to 99%). On the other hand, the study displayed the markedly decreased efficiency in bypassing L-nucleosides by Taq DNA polymerase and Vent (exo⁻) DNA polymerase for in vitro study and even completely inhibition for L-dC by Taq DNA polymerase, suggesting that L-nucleosides exerted a strong blockage effect on in vitro DNA replication. This difference of bypass efficiency in vitro and in cells could be associated with differences of *E. coli* polymerases and tested DNA polymerases, or cellular factors increased the bypass efficiency of L-nucleotides. *E. coli* replication system involved in five kinds of DNA polymerases (I, II, III, IV, V), and three of five DNA polymerases had high expression levels by the SOS DNA-damage response. Pol II was responsible for bacterial DNA repair and replication restart after DNA damage²⁴. Pol III is an asymmetric, dimeric enzyme, which carries out the high-speed replication of the two DNA strands²². Pol IV and Pol V, which are Y-family DNA polymerases²⁴, have low-processivity error-prone ability to bypass lesions via translesion synthesis. Similar to Y family polymerases in *E. coli*, some cellular factors, such as UmuD and UmuC, could involve in bypassing distinct template lesions²⁵. Ogt, an alkyltransferase in *E. coli*, involved in repairing O4-alkyl dT with small straight-chain alkyl functionality²⁶. What's more, many enzymes showed synergism with each other²⁷. Therefore, a tested DNA polymerase used in in vitro experiments has a special function, whereas the replication system in *E. coli* is so complicated that bypass efficiency of L-nucleotides on the template can be increased in *E. coli*.

For in vitro replication, Taq DNA polymerase and Vent (exo⁻) DNA polymerase exhibited similar effects on bypassing different L-nucleosides except L-dC. The observation underscored that L-dC completely inhibited Taq DNA polymerase, whereas it can be bypassed by Vent (exo⁻) DNA polymerases (Fig. 5). Thereby we performed molecular dynamics (MD) simulation to understand about the impact of L-nucleosides on DNA replication and to uncover the roles of DNA polymerases in bypassing these lesions. Considering that PDB crystallographic data for Vent (exo⁻) DNA polymerase was not available, we replaced Vent (exo⁻) DNA polymerase with Deep vent DNA polymerase, since they belonged to the B family of DNA polymerases and had the similar active domain.

For Taq DNA polymerase assemblies, root mean square deviation (RMSD) values of the complex containing L-dC, L-dG and L-dA modified duplexes as well as the normal DNA duplexes obtained from the final 10 ns simulations were presented in Supplementary Fig. S10a–c. L-dA and L-dG-contained duplexes showed RMSD values within the range of 0–5 Å, which were quite close to normal duplex (0–5.5). The L-dC-modified duplexes, by contrast, showed a high RMSD value (0–4.5) relative to normal conformation. Moreover, the results showed that the pseudorotation angles of L-dA and L-dG exhibited the same sugar conformation to regular conformation, associated to C2'-exo and C3'-endo conformation, whereas L-dC associated to O4'-exo-C1'-endo conformation type, resulting in a twist structure (see Supplementary Fig. S10d, e, f online). Together, these results reflected that L-dG and L-dA were bypassed by Taq DNA polymerase whereas L-dC was not.

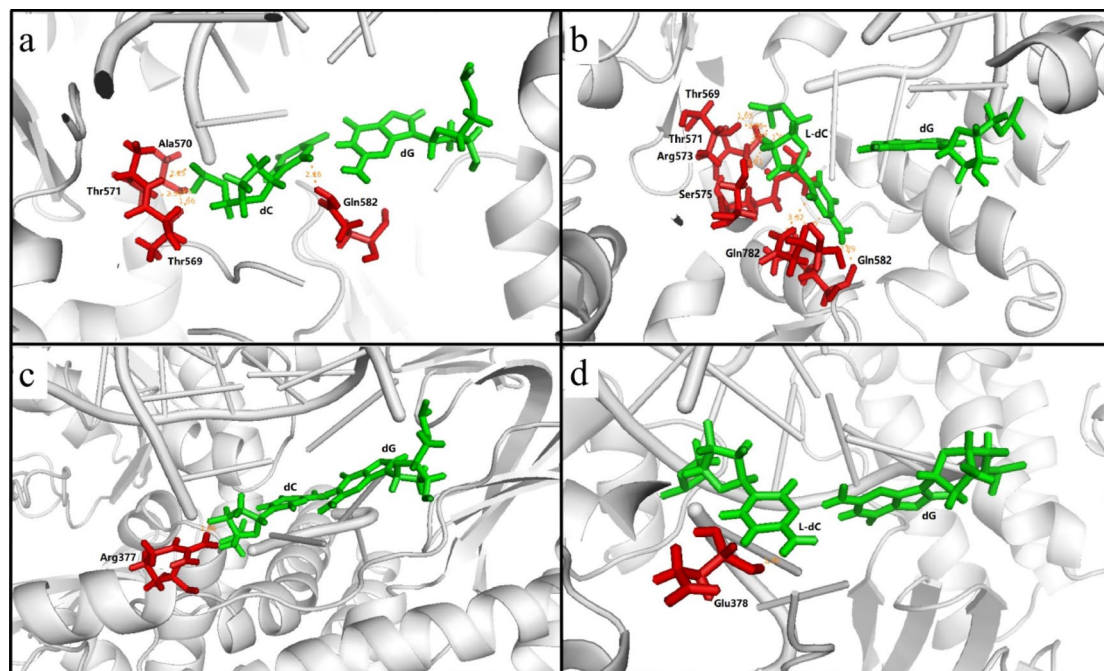


Figure 6. The cognate nucleotide insertion opposite to L-dC and dC by Taq DNA polymerase and Vent DNA polymerases. Taq DNA polymerase can bypass dC (a) but not bypass L-dC (b). Vent DNA polymerase can bypass not only dC (c) but also L-dC (d). Nucleotide pairs were shown in green, neighbouring amino acid residues were shown in red.

In contrast to Taq DNA polymerase, experiment showed that all of the L-nucleosides were bypassed by Vent (exo⁻) DNA polymerase, although L-nucleosides moderately blocked DNA replication. This can be verified by the fact that RMSD values has no obvious difference in assemblies between L-dG, L-dA and L-dC and the corresponding D-nucleosides using Deep vent DNA polymerase (see Supplementary Fig. S11 online).

On the basis of these results, we reasoned that the polymerase played a key role in bypassing L-nucleoside lesions. It was known that family B DNA polymerases (Deep vent, Vent (exo⁻), and Therminator) were more efficient at incorporating chemical modifications in their substrates than the DNA polymerases from family A, such as Taq DNA polymerase^{19,28–30}, because of striking difference derives from several key catalytic amino residues of O-helix of the finger domain³¹. As shown in Fig. 6, the nucleotide selection in Taq DNA polymerase was coordinated by this highly conserved residues Thr569 (O–H...O–P), Ala570 (N–H...O–P), Thr571 (O–H...O–P), Gln582 (N–H...O–C) in active domain (Fig. 6a). Nevertheless, when the template dC were replaced with L-dC, the key residues were arranged to residues Thr569 (O–H...O–P), Thr571 (O–H...O–P), Arg573 (N–H...O–P), Ser575 (O–H...O–C), Gln782 (N–H...O–C), Gln582 (N–H...O–C) (Fig. 6b), hydrogen bonded with L-dC to drag nucleotide at pre-translocation configuration, where hydrogen bonds on deoxyribose produce the main effect, such as Arg573 (N–H...O–P), Ser575 (O–H...O–C), Gln782 (N–H...O–C), converting favourable 3′-endo-sugar conformation to unfavourable O4′-exo-C1′-endo conformation, resulting in enzyme's failure to bypass of the lesion. For Deep vent DNA polymerase assemblies, the critical catalytic residues may be rearranged from Arg377 (N–H...O–C) to Glu378 (N–H...O–C) (Fig. 6c, d) when the template base switched from dC to L-dC. Obviously, the sterical clash with the side chain of amino acid residue in Deep vent DNA polymerase can be much reduced relative to Taq DNA polymerase, without the conformation and hydrogen bonding force being changed, thereby Deep vent DNA polymerase was also capable of incorporating nucleotides opposite to L-dC. In addition, when the template base changed from dC to L-dC in the complexes of Taq polymerase, and the binding energy changes from –341.6501 to –312.2146 (kcal/mol), producing an energy difference of 29.4355 (kcal/mol). In the Deep vent DNA polymerase, the binding energy changes from –43.5818 to –36.5313 (kcal/mol), producing an energy difference of 7.0505 (kcal/mol). Thus, the difference between L-dC and dC with respect to binding energy of the complexes of Taq polymerase was more pronounced than those of Deep vent DNA polymerase (see Supplementary Table S3 online).

Another major goal of the study was to examine the impact of the L-nucleoside lesions on fidelity of DNA replication. Although some DNA polymerases can employ damaged DNA as a template, nucleotide insertions mediated by these polymerases in cells were sometimes inaccurate, thereby leading to mutations³². Previous results revealed that L–T was found not to be mutagenic by a single DNA polymerase¹⁹, whereas it displayed mutagenic effects in *E. coli* cells, where TLS polymerases may have different mutagenic affects in extending the nascent strand from L–T-containing strand²¹. In the study, L-dC induced no mutation on DNA replication in vitro solution and in cells (Fig. 2d). Interestingly, L-dG lesion was no mutagenic in cells, whereas L-dG → T and L-dG → dG conversions catalysed by Vent (exo⁻) DNA polymerase was observed, where L-dG → T frequency (65%) was two times higher than L-dG → dG frequency (see Fig. 3d and Supplementary Table S1 online). Our

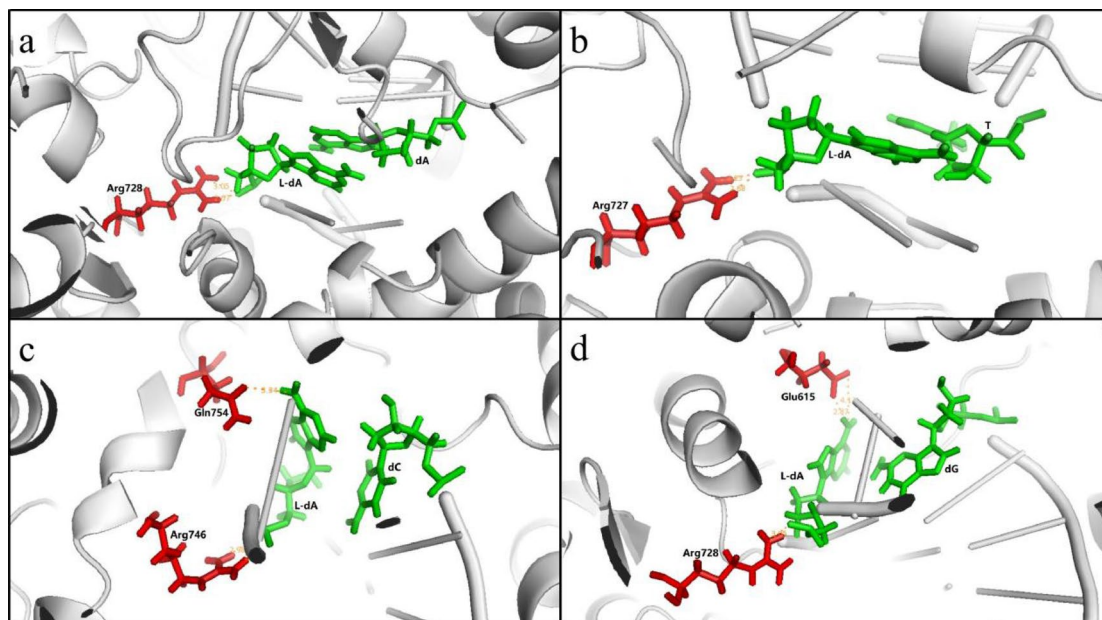


Figure 7. The cognate and non-cognate nucleotide insertion opposite to L-dA by Taq DNA polymerase. **(a)** L-dA:dA, **(b)** L-dA:T, **(c)** L-dA:dC, **(d)** L-dA:dG. Nucleotide pairs were shown in green, neighbouring amino acid residues were shown in red.

results also demonstrated that L-dA exhibited the L-dA → T (37%) and L-dA → dA (63%) conversions for Taq DNA polymerase and L-dA → T (100%) conversions for Vent (exo⁻) DNA polymerase. Differently, L-dA also induced more types of conversions in cell replication, i.e. L-dA → dA (67%), L-dA → dG (24%) and L-dA → dC (9%) (see Fig. 4d and Supplementary Table S1 online). Among the four L-nucleoside lesions, L-T exhibited the highest frequency of nucleoside misincorporation in *E. coli* cells, with L-T → dA (13%), L-T → dG (19%) and L-T → dC (46%) mutations (Supplementary Table S1 and Fig. S12 online). The difference in mutagenic frequencies of L-nucleosides bypass *in vivo* and *in vitro* may be attributed to functional interaction between TLS polymerases and other related proteins^{33–35}.

Except that DNA polymerases have relative low error rate on replication, the mutagenic effect may attribute to change of the stereochemistry of the templating nucleotide from D to L, which was expected to alter its interactions with the incoming nucleotide and active site residues and influence nucleotide binding and incorporation³⁶. Therefore, we performed molecular simulations to investigate how these lesions elicited mutations. We performed MD simulation for insertion of the cognate and non-cognate nucleotides opposite to L-nucleosides by using a template strand containing an L-deoxynucleoside to pair with the primer strand with the insertion of a dNTP in the presence of DNA polymerase. The results revealed that whether L-nucleosides elicited mutations depended on the interaction of the base pair formed from L-nucleosides and the incoming nucleotide with active site residues despite of ribosyl alteration from D to L. Figure 7 and Supplementary Fig. S13 showed this vein related to Taq DNA polymerase involved in nucleotide incorporation opposite L-dA. In the normal case, the residue Arg728 of O-helix interacts with dA the template (Fig. 7a) to operate dNTP by conformational changes from an “open” complex to a catalytically competent “closed” complex³⁷. When L-dA was introduced into the template, the key catalytic residues Arg727 (N-H...O-P), Arg728 (N-H...O-P), hydrogen bonded to L-dA, can pack against the nascent base pair L-dA:T and L-dA:dA, respectively, forming closed ternary complexes, similar to the key residues of dA:T (Fig. 7a, b and Supplementary Fig. S13). Differently, L-dA in the nascent base pair L-dA:dC and L-dA:dG was confined to two key residues Arg746 (N-H...O-P) and Gln754 (N-H...O-C) for L-dA:dC, Arg728 (N-H...O-P) and Gln615 (N-H...O-C) for L-dA:dG by hydrogen bonds, displayed a twist structure deviated from Watson–Crick base pairs. When the L-dA alignment is dC or dG, L-dA is constrained by hydrogen bonds between the hydroxyl and phosphate oxygen. As a result, the closed conformations can be not stabilized by these interactions. Thus, Taq polymerase has a tendency to incorporate dATP and dTTP, but no tendency to insert dGTP and dCTP (see Fig. 7 and Supplementary Table S1 online). Besides, the topological difference of more DNA polymerases would be needed to reveal whether this was the case by analysis of the X-ray crystal structure of their complexes in the future.

Conclusions

In conclusions, our systematic study based on shuttle vector that containing L-nucleosides provided important novel insights about effect of L-nucleosides on efficiency and fidelity of DNA replication *in vitro* and in *E. coli* cells. Our results indicated that L-nucleosides exhibited pronounced replicative efficiency in bypassing the lesions in *E. coli* cells, whereas they exerted a strong blockage effect on DNA replication by Taq DNA polymerases or Vent (exo⁻) DNA polymerases, and especially L-dC completely inhibited DNA replication by Taq DNA polymerases. Additionally, we found that L-dA and L-T stimulated various types of mismatch in *E. coli* cells, including purine → pyrimidine or pyrimidine → purine conversion and purine → purine or pyrimidine → pyrimidine

transversion. MD simulations were performed to clarify the change in structures, dynamics, stability of DNA-polymerase assemblies with incorporation of L-nucleosides, and help to understand how L-nucleosides impair the efficiency and fidelity of DNA replication. Further experiments will be needed to investigate the repair pathway and mechanism due to the effect of multiple family DNA polymerases. Our study provides a new understanding of the effect of L-2'-deoxynucleosides on DNA replication in vitro solution and in cells.

Material and methods

Plasmid pUC19, DH5 α competent cells, Taq DNA polymerase, X-gal, and dNTPs were purchased from Tiangen Biotec Co., Ltd. (Beijing, China). Vent (exo⁻) DNA polymerase and restriction enzymes (*Hind*III, *Eco*RI, and *Pst*I) were obtained from New England Biolabs Inc (Ipswich, MA, USA). SYBR[®] gold nucleic acid gel stain was purchased from Invitrogen, Ltd. (Paisley, UK). T4 DNA ligase and T4 polynucleotide kinase were purchased from Beyotime (Shanghai, China). Plasmid small-extraction kit and 40% acrylamide (acrylamide:bis-acrylamide 19:1) were from Beijing Solarbio Science & Technology Co., Ltd. (Beijing, China). The E.Z.N.A.[®] Cycle-Pure Kit was obtained from Omega Biotek Inc (Georgia, USA). All oligonucleotides used in this study were purchased from Suzhou Beixin Biotechnology Co., Ltd. Ultrapure water was prepared using double distilled water (UPW, 18 M Ω .cm).

The following equipments were used: PAGE electrophoresis system (Junyi JY200C, Beijing, China), Micro high-speed centrifuge (Xiangyi TG16-W, Changsha, China), Microvolume UV-Vis spectrophotometer (NanoDrop2000 Thermo Fisher, Boston, Massachusetts, USA), ChemiDoc[™] MP Imaging System (BioRad, Hercules, California, USA), MJ MiNi Thermal Cycler (Bio-Rad PTC-1148, Hercules, California, USA), and High-speed refrigerated centrifuge (Xiangyi TGL-16, Changsha, China).

Preparation of the L-nucleoside-carrying 24-mer ODNs. 20 μ L of reactions that contains 2 μ L of reaction buffer A (10 Å —), an equimolar amount of L-nucleoside-carrying ODNs (i.e. G13 containing L-dG, 5'-AATTCTCCG¹GCAGGCTAGCACTGA-3') and its complementary strand (COM20, 5'-AGCTTCAGT GCTAGCCTGCACGAG-3'), 1 μ L of 0.1 mM ATP, 1 μ L of T4 polynucleotide kinase (10 U/ μ L), and nuclease-free ionized water were mixed thoroughly and 5'-phosphorylated at 37 $^{\circ}$ C for 30 min, followed by incubation at 95 $^{\circ}$ C for 10 min and cooling for 10 min at room temperature.

Construction of L-nucleosides-containing genomes. According to the previously reported procedures²¹, 300 ng of pUC19 plasmids were digestion with 20 μ L *Eco*RI (20000 U mL⁻¹) and 20 μ L *Hin*-dIII (20000 U mL⁻¹). The resulting linearized vector was isolated by E.Z.N.A.[®] Cycle-Pure Kit, following the manufacturer's instructions. The ligation reactions in 20 μ L of reaction containing the linearized vector and double-stranded insert at a molar ratio of vector to insert of 1:3, 2 μ L of 10 \times T4 DNA ligase buffer, 0.5 μ L (1000 U μ L) of T4 DNA ligase, and distilled water, were performed at 16 $^{\circ}$ C for overnight.

Transfection of L-nucleosides-containing genomes. Prior to plasmid-containing cell culture, 10 μ L of ligation mixture was transformed into 50 μ L chemically competent DH5 α *E. coli* cells by heat shock at 42 $^{\circ}$ C for 90 s. Then, 400 μ L LB liquid medium was added into above cells, followed by cultivation at 37 $^{\circ}$ C with shaking (225 rpm) for 1 h. lastly, 300 μ L cultured bacterial can be added into another 5 mL LB liquid medium containing AMP (ampicillin) at concentration of 100 mg/L and cultivated at 37 $^{\circ}$ C with shaking (225 rpm) for 14–16 h. The progeny genomes were isolated using plasmid small-extraction kit.

Quantification of bypass efficiencies and mutation frequencies. The version of restriction enzyme-mediated assay developed by us was employed to assess the bypass efficiencies and mutation frequency of the DNA lesion in vitro and in cells²⁰ (Fig. 1). The L-nucleoside-carrying recombinants from ligation reaction and progeny genomes arising from cellular replication, representing as in vitro and in cell replication study, were amplified by PCR with the use of Taq DNA polymerase, respectively. The two primers were C12F 5'-TTAGCT CACTCATTAGGCACCC-3' and C12R 5'-GGGCCTCTTCGCTATTACGC-3', and the PCR amplifications initiated at 95 $^{\circ}$ C for 3 min, followed by 39 cycles at 95 $^{\circ}$ C for 30 s, 52.2 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 1 min, a final extension at 72 $^{\circ}$ C for 5 min, and hold at 4 $^{\circ}$ C. The PCR products were purified using Cycle Pure Kit (Omega) and stored at -20 $^{\circ}$ C until use.

For the analysis of bypass efficiency, a portion of the PCR products was treated with *Eco*RI/*Hind*III and *Pst*I/*Hind*III at 37 $^{\circ}$ C in 5 μ L of NEB CutSmart buffer for 4 h, respectively, referring to NEB restriction digest protocol. The reaction mixture was quenching with 3 (for *Eco*RI/*Hind*III digestion products) and 1 (for *Pst*I/*Hind*III digestion products) time volume of 2 \times denaturing DNA loading buffer containing bromophenol blue dyes. The mixture was loaded onto 20% denature polyacrylamide gel (acrylamide : bis-acrylamide = 19:1) and electrophoresis was done at 150 V for 1.5 h. The gel was stained with SYBR Gold Nucleic Acid Gel Stain and photographed using a ChemiDoc[™] MP Imaging System (Bio-Rad). The quantification of band intensities was conducted using Image Lab software version 6.0 (Bio-Rad).

The effects of L-nucleosides on DNA replication were determined by the relative bypass efficiency (RBE) and mutation frequency values, respectively. The mutation frequency was calculated from the percentage of mutagenic product (i.e. the product with L-G \rightarrow T mutation) among the sum of all products arising from replication of the lesion-containing genome. The bypass efficiency was sum of all relative incorporation efficiency of all nucleotides opposite L-nucleosides.

Molecular dynamics simulations. The crystal structures of the pre-translocation (closed) state of Taq DNA polymerase (3KTQ)³⁸ and the pre-translocation (closed) state of Deep vent DNA polymerase (5H12)³⁹ both obtained using X-ray diffraction, were downloaded from the PDB crystal library. The Taq DNA polymerase was in the pre-translocation state, where the finger structure domain was in the closed state and the C-terminal position of the O-helix was in the "Pre" state. Deep vent DNA polymerase with structureless DNA requires software to bind to the DNA duplex under study and then pre-process the two complex structures accordingly.

This part of the study used the PMEMD software⁴⁰ in Amber 16 software on a GPU platform to run the kinetic simulations. Before the kinetic simulations, the complex system was hydrogenated using the leap program. The force field used for the protein was the ff14SB force field⁴¹, and the force field used for DNA was the DNA-OL15⁴² force field. The complexes were then placed in a 12 Å positive charged octahedral box of aqueous solvent, and the water molecules were made electrically neutral by using TIP3P⁴³, which added counterbalancing ions (Na⁺, Cl⁻) to the entire complex system.

After processing the complex system using the leap procedure, the whole system was optimized. After optimization, two ramp-up processes were carried out for the complexes within the NVT system in seven steps, from 0 to 310 K and 345.15 K. After optimization and ramp-up, molecular dynamics simulations were carried out for each complex for 10 ns, with a time step of 2 fs and a frame saved at 1 ps intervals. The SHAKE algorithm was used to control the stretching of the chemical bonds containing the nitrogen atoms and the truncation value of the non-bonded interactions was set to 10 Å.

The cpptraj program in Amber 16 was used for the analysis of trajectories, the analysis of RMSD, hydrogen bonding analysis, and distance analysis. This part of study used the MM/GBSA method to calculate the interaction energy of DNA with proteins. PYMOL⁴⁴, Origin and MATLAB were used for graphing and data analysis.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 12 August 2022; Accepted: 21 November 2022

Published online: 07 December 2022

References

- Wang, Z., Xu, W., Liu, L. & Zhu, T. F. A synthetic molecular system capable of mirror-image genetic replication and transcription. *Nat. Chem.* **8**, 698–704. <https://doi.org/10.1038/nchem.2517> (2016).
- Kielkowski, P., Fanfrlik, J. & Hocek, M. 7-Aryl-7-deazaadenine 2'-deoxyribonucleoside triphosphates (dNTPs): Better substrates for DNA polymerases than dATP in competitive incorporations. *Angew. Chem. Int. Ed. Engl.* **53**, 7552–7555. <https://doi.org/10.1002/anie.201404742> (2014).
- Balintova, J., Welter, M. & Marx, A. Antibody-nucleotide conjugate as a substrate for DNA polymerases. *Chem. Sci.* **9**, 7122–7125. <https://doi.org/10.1039/c8sc01839a> (2018).
- Capek, P. *et al.* An efficient method for the construction of functionalized DNA bearing amino acid groups through cross-coupling reactions of nucleoside triphosphates followed by primer extension or PCR. *Chem. Eur. J.* **13**, 6196–6203. <https://doi.org/10.1002/chem.200700220> (2007).
- Kempeneers, V., Renders, M., Froeyen, M. & Herdewijn, P. Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. *Nucleic Acids Res.* **33**, 3828–3836. <https://doi.org/10.1093/nar/gki695> (2005).
- Jiang, C. W. *et al.* Synthesis and recognition of novel isonucleoside triphosphates by DNA polymerases. *Bioorgan. Med. Chem.* **15**, 3019–3025. <https://doi.org/10.1016/j.bmc.2007.02.003> (2007).
- Schultz, H. J. *et al.* Taq DNA polymerase mutants and 2'-modified sugar recognition. *Biochemistry* **54**, 5999–6008. <https://doi.org/10.1021/acs.biochem.5b00689> (2015).
- Spadari, S. *et al.* L-Thymidine is phosphorylated by herpes simplex virus type 1 thymidine kinase and inhibits viral growth. *J. Med. Chem.* **35**, 4214–4220. <https://doi.org/10.1021/jm00100a029> (1992).
- Bryant, M. L. *et al.* Antiviral beta-L-nucleosides specific for hepatitis B virus infection. *Front. Viral Hepat.* **45**, 245–261. <https://doi.org/10.1016/b978-044450986-4/50072-2> (2003).
- De Clercq, E. & Field, H. J. Antiviral prodrugs - the development of successful prodrug strategies for antiviral chemotherapy. *Brit. J. Pharmacol.* **147**, 1–11. <https://doi.org/10.1038/sj.bjp.0706446> (2006).
- Chatterjee, N. & Walker, G. C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **58**, 235–263. <https://doi.org/10.1002/em.22087> (2017).
- Torgovnick, A. & Schumacher, B. DNA repair mechanisms in cancer development and therapy. *Front. Genet.* **6**, 1. <https://doi.org/10.3389/fgene.2015.00157> (2015).
- Stallons, L. J. & McGregor, W. G. Translesion synthesis polymerases in the prevention and promotion of carcinogenesis. *J. Nucleic Acids.* **2010**, 1–10. <https://doi.org/10.4061/2010/643857> (2010).
- Neeley, W. L., Delaney, J. C., Henderson, P. T. & Essigmann, J. M. *In vivo* bypass efficiencies and mutational signatures of the guanine oxidation products 2-aminoimidazole and 5-guanidino-4-nitroimidazole. *J. Biol. Chem.* **279**, 43568–43573. <https://doi.org/10.1074/jbc.M407117200> (2004).
- Henderson, P. T. *et al.* The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors in vivo. *Biochemistry-US* **42**, 9257–9262. <https://doi.org/10.1021/bi0347252> (2003).
- Delaney, J. C. & Essigmann, J. M. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine and 3-methylthymine, in alkB *Escherichia coli*. *P. Natl. Acad. Sci. USA* **101**, 14051–14056. <https://doi.org/10.1073/pnas.0403489101> (2004).
- Bailey, E. A., Iyer, R. S., Stone, M. P., Harris, T. M. & Essigmann, J. M. Mutational properties of the primary aflatoxin B-1-DNA adduct. *P. Natl. Acad. Sci. USA* **93**, 1535–1539. <https://doi.org/10.1073/pnas.93.4.1535> (1996).
- Wu, J. *et al.* Translesion synthesis of O-4-alkylthymidine lesions in human cells. *Nucleic Acids Res.* **44**, 9256–9265. <https://doi.org/10.1093/nar/gkw662> (2016).
- Xiao, Y. *et al.* Mirror-image thymidine discriminates against incorporation of dNTP into DNA and repairs itself by DNA polymerases. *Bioconjugate Chem.* **28**, 2125–2134. <https://doi.org/10.1021/acs.bioconjchem.7b00301> (2017).
- Kan, Y., Wu, L. & He, Y. Mutation analysis of L-thymidine-induced replication products using a restriction enzyme-mediated assay. *Curr. Protoc. Nucleic Acid Chem.* **83**, 21. <https://doi.org/10.1002/cpnc.121> (2020).

21. Kan, Y. *et al.* Replication of DNA containing mirror-image thymidine in *E. coli* cells. *Chem. Res. Toxicol.* **33**, 2276–2285. <https://doi.org/10.1021/acs.chemrestox.9b00502> (2020).
22. Fijalkowska, I. J., Schaaper, R. M. & Jonczyk, P. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol. Rev.* **36**, 1105–1121. <https://doi.org/10.1111/j.1574-6976.2012.00338.x> (2012).
23. Catherine, M. J., Judith, L. C. Encyclopedia of biological chemistry, 720–728 (2004).
24. Hastings, P. J. *et al.* Competition of *Escherichia coli* DNA polymerases I, II and III with DNA Pol IV in stressed cells. *PLoS One* **5**, 10862. <https://doi.org/10.1371/journal.pone.0010862> (2010).
25. Beuning, P. J., Simon, S. M., Godoy, V. G., Jarosz, D. F. & Walker, G. C. Characterization of *Escherichia coli* translesion synthesis polymerases and their accessory factors. *Methods Enzymol.* **408**, 318–340. [https://doi.org/10.1016/s0076-6879\(06\)08020-7](https://doi.org/10.1016/s0076-6879(06)08020-7) (2006).
26. Wang, P., Amato, N. J., Zhai, Q. & Wang, Y. Cytotoxic and mutagenic properties of O4-alkylthymidine lesions in *Escherichia coli* cells. *Nucleic Acids Res.* **43**, 10795–10803. <https://doi.org/10.1093/nar/gkv941> (2015).
27. Lindahl, T. New class of enzymes acting on damaged DNA. *Nature* **259**, 64–66. <https://doi.org/10.1038/259064a0> (1976).
28. Baccaro, A., Steck, A. L. & Marx, A. Barcoded nucleotides. *Angew. Chem.* **51**, 254–257. <https://doi.org/10.1002/anie.201105717> (2012).
29. Veedu, R. N., Vester, B. & Wengel, J. Efficient enzymatic synthesis of LNA-modified DNA duplexes using KOD DNA polymerase. *Org. Biomol. Chem.* **7**, 1404–1409. <https://doi.org/10.1039/b819946a> (2009).
30. Ikonen, S., Macickova-Cahova, H., Pohl, R., Sanda, M. & Hocek, M. Synthesis of nucleoside and nucleotide conjugates of bile acids, and polymerase construction of bile acid-functionalized DNA. *Org. Biomol. Chem.* **8**, 1194–1201. <https://doi.org/10.1039/b924072a> (2010).
31. Gardner, A. E., Joyce, C. M. & Jack, W. E. Comparative kinetics of nucleotide analog incorporation by vent DNA polymerase*. *J. Biol. Chem.* **279**, 11834–11842. <https://doi.org/10.1074/jbc.M308286200> (2004).
32. Du, H., Leng, J. P., Wang, P. C., Li, L. & Wang, Y. S. Impact of tobacco-specific nitrosamine-derived DNA adducts on the efficiency and fidelity of DNA replication in human cells. *J. Biol. Chem.* **293**, 11100–11108. <https://doi.org/10.1074/jbc.RA118.003477> (2018).
33. Yoon, J. H., Bhatia, G., Prakash, S. & Hurwitz, P. J. Error-free replicative bypass of thymine glycol by the combined action of DNA polymerases κ and ζ in human cells. *P. Natl. Acad. Sci. USA* **107**, 14116–14121. <https://doi.org/10.1073/pnas.1007795107> (2010).
34. Wu, J., Du, H., Li, L., Price, N. E. & Wang, Y. The impact of minor-groove N2-Alkyl-2'-deoxyguanosine lesions on DNA replication in human cells. *ACS Chem. Biol.* **14**, 1708–1716. <https://doi.org/10.1021/acschembio.9b00129> (2019).
35. Shachar, S. *et al.* Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals (vol 28, pg 383, 2009). *EMBO J.* **28**, 992–992. <https://doi.org/10.1038/emboj.2009.72> (2009).
36. Gaur, V. *et al.* Structural and kinetic insights into binding and incorporation of L-nucleotide analogs by a Y-family DNA polymerase. *Nucleic Acids Res.* **42**, 9984–9995. <https://doi.org/10.1093/nar/gku709> (2014).
37. Wu, S., Wang, J., Pu, X., Li, L. & Li, Q. T7 RNA polymerase discriminates correct and incorrect nucleoside triphosphates by free energy. *Biophys. J.* **114**, 1755–1761. <https://doi.org/10.1016/j.bpj.2018.02.033> (2018).
38. Li, Y., Korolev, S. & Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of thermus aquaticus DNA polymerase I structural basis for nucleotide incorporation. *EMBO J.* **17**, 7514–7525. <https://doi.org/10.1093/emboj/17.24.7514> (1998).
39. Hikida, Y., Kimoto, M., Hirao, I. & Yokoyama, S. Crystal structure of deep vent DNA polymerase. *Biochem. Biophys. Res. Commun.* **483**, 52–57. <https://doi.org/10.1016/j.bbrc.2017.01.007> (2017).
40. Kaus, J. W., Pierce, L. T., Walker, R. C. & McCammont, J. A. Improving the efficiency of free energy calculations in the amber molecular dynamics package. *J. Chem. Theory Comput.* **9**, 4131–4139. <https://doi.org/10.1021/ct400340s> (2013).
41. Maier, J. A. *et al.* ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255> (2015).
42. Galindo-Murillo, R. *et al.* Assessing the current state of amber force field modifications for DNA. *J. Chem. Theory Comput.* **12**, 4114–4127. <https://doi.org/10.1021/acs.jctc.6b00186> (2016).
43. Peng, Z. W., Ewig, C. S., Hwang, M. J., Waldman, M. & Hagler, A. T. Derivation of class II force fields. 4. van der waals parameters of alkali metal cations and halide anions. *J. Phys. Chem. A.* **101**, 7243–7252. <https://doi.org/10.1021/jp964080y> (1997).
44. Rigby, R. E. & Parker, A. B. Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem. Mol. Biol. Edu.* **44**, 433–437. <https://doi.org/10.1002/bmb.20966> (2016).

Acknowledgements

This work was supported by Beijing Natural Science Foundation (grant number 2192058), National Natural Science Foundation of China (grant number 21778054, 51972302), National Key Research and Development Program of China (grant number 2016YFF0203703), State Key Laboratory of Natural and Biomimetic Drugs (grant number K20180202), Fundamental Research Funds for the Central Universities and High Level Introduction of Talent Research Start-up Fund in Inner Mongolia University Funding Project (grant number 10000-21311201/170).

Author contributions

Y.K. performed the experiments and wrote the manuscript; L.W. and Y.H. designed the research; L.W. and Z.J. revised the manuscript; Z.J. performed molecular dynamics simulations; Y.K. helped to compute and simulate dynamics data; D.L. and L.Y. analyzed the experimental data. All authors have reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-24802-5>.

Correspondence and requests for materials should be addressed to L.W. or Y.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022