




# Structural basis of DNA synthesis opposite 8-oxoguanine by human PrimPol primase-polymerase

Olga Rechkoblit <sup>1</sup>✉, Robert E. Johnson<sup>2</sup>, Yogesh K. Gupta <sup>1,3</sup>, Louise Prakash<sup>2</sup>, Satya Prakash<sup>2</sup> & Aneel K. Aggarwal <sup>1</sup>✉

PrimPol is a human DNA polymerase-primase that localizes to mitochondria and nucleus and bypasses the major oxidative lesion 7,8-dihydro-8-oxoguanine (oxoG) via translesion synthesis, in mostly error-free manner. We present structures of PrimPol insertion complexes with a DNA template-primer and correct dCTP or erroneous dATP opposite the lesion, as well as extension complexes with C or A as a 3′-terminal primer base. We show that during the insertion of C and extension from it, the active site is unperturbed, reflecting the readiness of PrimPol to accommodate oxoG(*anti*). The misinsertion of A opposite oxoG(*syn*) also does not alter the active site, and is likely less favorable due to lower thermodynamic stability of the oxoG(*syn*)•A base-pair. During the extension step, oxoG(*syn*) induces an opening of its base-pair with A or misalignment of the 3′-A primer terminus. Together, the structures show how PrimPol accurately synthesizes DNA opposite oxidatively damaged DNA in human cells.

<sup>1</sup>Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. <sup>2</sup>Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA. <sup>3</sup>Present address: Greehey Children's Cancer Research Institute & Department of Biochemistry and Structural Biology, University of Texas Health at San Antonio, San Antonio, TX, USA. ✉email: [olga.rechkoblit@mssm.edu](mailto:olga.rechkoblit@mssm.edu); [aneel.agarwal@mssm.edu](mailto:aneel.agarwal@mssm.edu)

**7,8-d**ihydro-8-oxoguanine (oxoG) is the major oxidative DNA damage lesion in human cells. The oxidative damage is induced by reactive oxygen species (ROS) arising due to normal metabolic processes as well as environmental exposures such as ionizing radiation and the UVA band of sunlight<sup>1–4</sup>. It is estimated that each cell in a human body develops ~2,800 oxoG lesions per day<sup>5</sup>.

OxoG lesion is highly mutagenic because nuclear and mitochondrial (mt) replicative DNA polymerases effectively pair an oxoG template residue not only with the correct cytosine base but also mispair it with an adenine base and preferentially extend from the misinserted primer termini<sup>6–13</sup>. This induces a G to T transversion mutation upon the next round of replication. Since mitochondria produce substantial amounts of ROS in the cell<sup>14</sup>, oxidative damage to the mtDNA is deemed to be especially detrimental<sup>15</sup>. Unexpectedly, recent sequencing studies reveal that the mutations observed in mtDNA are transitions rather than transversions and, thus, are unlikely to be associated with the oxoG-driven mutagenesis<sup>16–20</sup>.

PrimPol is the only human DNA polymerase that is known to localize to mitochondria and is able to efficiently and mostly accurately bypass an oxoG via translesion synthesis (TLS)<sup>21–25</sup>. Consequently, it is expected that PrimPol aids the limited DNA repair processes operating in mitochondria<sup>26,27</sup> to protect mtDNA from the ROS-induced mutations.

PrimPol is important for both nuclear and mtDNA replication and belongs to the archaeo-eukaryotic primase-polymerase family that differs from other DNA polymerases by possessing a primase activity<sup>21,24,28–32</sup>. Thus, in addition to the catalytic core, PrimPol contains a zinc finger module that is required for the primase function<sup>28,31,33</sup>. The dual TLS polymerase and primase activities of PrimPol allow it to bypass DNA lesions via TLS as well as to ‘skip’ DNA lesions and initiate DNA synthesis de novo downstream of the lesion site. Indeed, PrimPol-primed DNA synthesis re-initiation after multiple cisplatin therapy doses in BRCA1-deficient cancer cells has emerged as a major chemoresistance pathway<sup>34</sup>. The crystal structure of the catalytic core of PrimPol demonstrates that the active site of PrimPol is relatively constrained and is able to accommodate smaller lesions, such as oxoG, and process them via TLS, but is not likely to do so with bulkier DNA adducts, such as UV-induced cis-syn T-T dimer, (6–4) T-T photoproduct or cisplatin-DNA adducts<sup>35</sup>.

DNA polymerases can manipulate the *anti-syn* conformational equilibrium of the glycosidic torsion angle of the oxoG residue and thus governing its pairing either with a C base via its Watson–Crick edge (when in the *anti* conformation) or with an A via its Hoogsteen edge (when in the *syn* conformation). Predominantly error-free oxoG bypass polymerases such as human X-family Pol $\lambda$ <sup>36</sup> and TLS Y-family archaea Dpo4<sup>37</sup> are able to favorably accommodate a local template distortion needed to resolve an otherwise close contact between the ‘extra’ O8 atom and the oxygen atoms of the sugar-phosphate backbone of the oxoG residue. Alternatively, human or yeast TLS Y-family Pol $\eta$ <sup>38</sup> is preconfigured in the oxoG (*anti*)-permissive conformation even on the unmodified template<sup>39,40</sup>. In contrast, within the constrained active sites of high-fidelity A-family DNA Pol I fragment<sup>10</sup> and Pol T7<sup>9</sup>, the oxoG (*syn*)•A mismatch mimics a cognate base pair and allows for misinsertion and further extension, while the conformational adjustments associated with the correct oxoG(*anti*)•C pair disrupt the active site and stall the polymerase. Mitochondrial replicative Poly also belongs to the A-family<sup>41</sup>, and is strongly inhibited by the oxoG lesion and predominantly extends from the oxoG•A base pair<sup>11,42</sup>.

PrimPol also stands out among DNA polymerases due to its unusual architecture and minimal contacts with template-primer DNA duplex<sup>32,35</sup>. Instead of the classical fingers-palm-thumb

domains observed in all other DNA polymerases, PrimPol catalytic core has two modules: an N-terminal ModN and a C-terminal ModC<sup>35</sup>. ModN makes limited interactions primarily with the template DNA strand and the templating base, whereas ModC contains the active site and interacts with the nascent base pair. Due to the absence of the ‘thumb’ domain, there are almost no contacts to the DNA primer strand. There is also no polymerase-associated domain (PAD, also termed ‘little finger domain’) that provides an extensive interface with the major groove of the template-primer duplex in all Y-family TLS polymerases<sup>32</sup>. Altogether, it is unclear how PrimPol governs up to six-fold preferable insertion of a C over A opposite the oxoG and a two-fold preferable extension from the oxoG•C over the oxoG•A base pair<sup>22</sup>.

Here we present four crystal structures of PrimPol that capture the insertion of a C and an A opposite the oxoG lesion, as well as the subsequent extension from the inserted bases. We show that during the insertion of C and extension from it, the overall conformation of the active site remains the same as in the undamaged complex; thus, demonstrating the readiness of PrimPol to accommodate the oxoG(*anti*) during both steps of the lesion bypass without considerable adjustment to the template DNA backbone. Our results also suggest that the misinsertion of an A opposite the oxoG is likely impaired by the lower thermodynamic stability of the oxoG(*syn*)•A(*anti*) base-pair rather than structural alterations. However, during the next extension step, the 3′-A primer terminus opposite the oxoG(*syn*) lesion can assume either productive or non-catalytic alignment, and thereby reducing the reaction rate. Together, the structures define the basis of PrimPol’s ability to perform predominantly accurate synthesis on oxidative-damaged DNAs in human cells.

## Results

**Structure determination.** To produce the insertion ternary PrimPol complexes with the correct C incoming nucleotide triphosphate opposite the oxoG lesion, we crystallized the catalytic core of human PrimPol (residues 1–354) with a 17-nucleotide (nt) DNA template (5′-CA(oxoG)CGCTACCACACCCC-3′) and a 2′-deoxy-3′-terminated 12 nt DNA primer (5′-GGTGTGGTAGCG-3′) in the presence of dCTP. The last two 5′-CC-3′ nucleotides of the template strand from its 3′-end are intentionally left unpaired with the primer strand to promote crystal-packing interactions. The conditions for crystal growth include only Ca<sup>2+</sup> ions, which do not support catalysis by PrimPol<sup>22</sup>. To produce crystals in the presence of dATP, we employ the aforementioned 17 nt template and 2′-deoxy- or 2′,3′-dideoxy-3′-terminated 13 nt DNA primer, namely 5′-GGGTGTGGTAGCG-3′ or 5′-GGGTGTGGTAGCG<sub>dd</sub>-3′. The 13 nt primer produces a single C nt 3′ template base overhang. To capture the extension step we used a 17 nt DNA template (5′-CAT(oxoG)CCTACCACACCCC-3′), where the oxoG lesion is moved one base downstream compared to its location in the insertion complex, and 13 nt DNA primers with either a C or an A 2′-deoxy-3′-terminal bases (5′-GGGTGTGGTAGGX-3′, where X is C or A), and the next correct incoming base dATP. The structures were solved by molecular replacement using the unmodified ternary complex with T-dATP nascent base pair as a search model (PDB ID: 5L2X)<sup>35</sup> and are refined to 2.60–2.07 Å resolution range. All structures have two ternary complexes (molecules A and B) in the crystallographic asymmetric unit (AU) that are very similar to each other, except for the N-helix (residues 1–17), which is disordered in the complex B molecules. The crystal data, data collection statistics, and refinement statistics for all complexes are summarized in Table 1.

**Correct C insertion opposite oxoG.** The overall structure of the PrimPol complex depicting insertion of the correct C opposite

**Table 1 X-ray data collection and refinement statistics.**

Complex name	oxoG•dCTP	oxoG•dATP	oxoG•dATP	oxoG•C	oxoG•A
Complex type	insertion	insertion	insertion	extension	extension
Primer strand 3'	2'-deoxy	2',3'-dideoxy	2'-deoxy	2'-deoxy	2'-deoxy
PDB ID	7JK1	7JKL	7JKP	7JL8	7JLG
<i>Data collection</i>					
Space group	P1	P1	P1	P1	P1
Cell dimensions:					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	52.6 65.7 75.2	52.5 65.6 74.3	52.6 65.7 75.2	50.7 65.3 72.0	51.0 65.2 73.1
$\alpha$ , $\beta$ , $\gamma$ (°)	69.7, 82.8, 88.5	69.7, 82.8, 88.5	69.7, 82.8, 88.5	68.5, 85.7, 86.9	71.4, 84.3, 88.6
Resolution range (Å) <sup>a</sup>	70.0–2.62 (2.74–2.62)	69.2–2.37 (2.46–2.37)	69.2–2.58 (2.69–2.58)	28.2–2.10 (2.16–2.10)	27.0–2.07 (2.13–2.07)
<i>R</i> <sub>merge</sub> (%)	11.2 (122)	9.2 (91.7)	8.8 (95.5)	5.5 (54.3)	6.1 (108)
<i>I</i> / $\sigma$ <i>I</i>	6.1 (0.9)	7.0 (0.8)	7.6 (0.7)	11.6 (1.7)	8.8 (0.8)
Completeness (%)	98.0 (93.9)	91.2 (64.0)	94.0 (82.1)	97.1 (93.6)	97.0 (95.5)
Redundancy	2.1 (2.1)	3.5 (2.9)	3.4 (2.7)	3.4 (3.3)	3.4 (3.3)
CC <sub>1/2</sub> (%)	98.8 (36.3)	99.3 (44.9)	99.5 (52.3)	99.8 (78.0)	99.8 (33.6)
<i>Refinement</i>					
Resolution range (Å)	39.5–2.62 (2.71–2.62)	69.2–2.38 (2.45–2.38)	69.2–2.58 (2.68–2.58)	28.2–2.10 (2.15–2.10)	27.0–2.07 (2.12–2.07)
No. reflections	27,566	32,954	26,648	48,566	52,522
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	23.9/ 28.6 (36.9/ 38.5)	22.0/ 27.7 (38.6/39.0)	24.8/ 30.7 (48.1/ 50.0)	18.7/ 21.3 (29.2/ 32.2)	20.1/ 21.9 (37.6/ 40.2)
No. of non-hydrogen atoms	5,440	5,483	5,291	5,171	5,351
Protein	4,122	4,159	4,053	4,085	4,116
DNA	1,143	1,143	1,145	807	1,021
Ligand (dNTP)	56	60	60	60	60
Ligand (other)	25	12	6	55	19
Ion (Ca <sup>2+</sup> )	2	3	3	3	4
Water	92	106	24	161	131
<i>B</i> -factors					
Protein	64.4	66.0	76.6	52.2	61.1
DNA	115.3	111.1	145.9	114.5	124.0
Ligand (dNTP)	53.9	57.0	70.5	43.3	56.2
Ligand (other)	56.9	62.7	76.9	70.6	74.3
Ion (Ca <sup>2+</sup> )	51.7	74.4	74.6	52.1	73.4
Water	54.2	60.7	63.6	48.9	54.7
<i>R.m.s. deviations</i>					
Bond length (Å)	0.011	0.005	0.011	0.008	0.003
Bond angles (°)	0.681	0.497	0.495	1.011	0.647

<sup>a</sup>Values in parentheses are for highest-resolution shell.

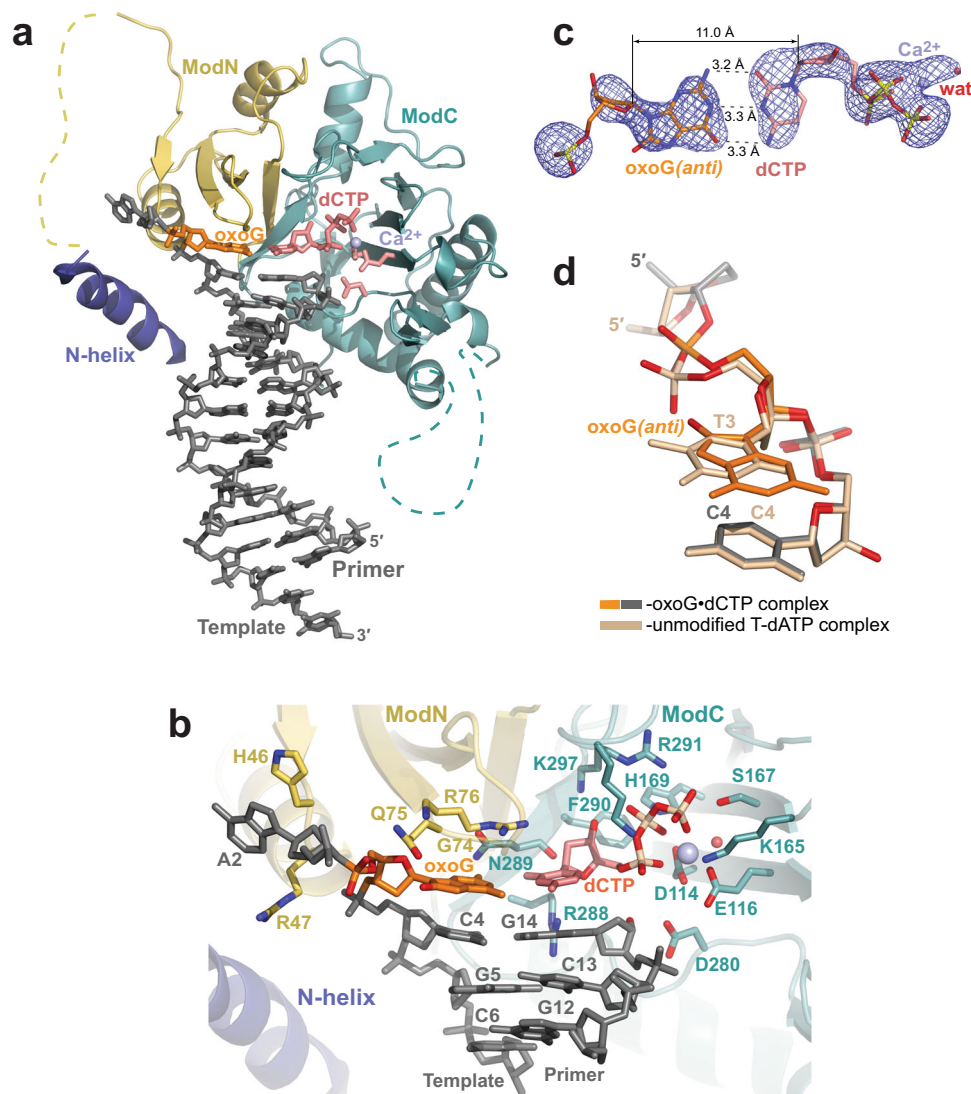
the template oxoG residue (oxoG•dCTP) is refined to 2.60 Å resolution and is similar to the structure of the unmodified complex (T-dATP) (root mean square deviation (rmsd) of ~0.35 Å over 249 Cas of molecule A complexes) (Fig. 1a and Supplementary Fig. 1). The PrimPol catalytic core clasps the template-primer by two  $\alpha/\beta$  modules, ModN and ModC (Fig. 1b). ModN (residues 35–105) interacts primarily with the template DNA strand and the template oxoG base, whereas ModC (residues 108–200 and 261–348) carries the catalytic residues Asp114, Asp116, and Glu280 and interacts with the incoming dCTP nucleotide: functions normally associated with the finger and palm domains of DNA polymerases, respectively. In the absence of the ‘thumb’ domain, there are almost no contacts to the DNA primer strand.

Only one ‘metal B’ Ca<sup>2+</sup> ion, chelated by the oxygen atoms of the triphosphate tail of dCTP, is observed in the oxoG•dCTP complex despite the intact 3'-OH primer terminus (Fig. 1a, b). A single ‘B’ ion is also present in the unmodified complex with a dideoxy-terminated primer terminus (Supplementary Fig. 1a).

Somewhat different from the unmodified complex, the N-terminal helix (N-helix; residues 1–17) is slightly shifted (by ~2.5–3 Å) and rotated (by ~3°) away from the major groove of the DNA duplex (Supplementary Fig. 1a), allowing for the formation of a salt bridge between the side chains of Arg3 of the N-helix and Glu105, the last C-terminal residue of ModN (Supplementary Fig. 1b). The side chain of Lys10, the only N-helix residue interacting with an oxygen atom of the sugar-phosphate DNA backbone (template residue G5) in the unmodified complex is disordered in the electron density map in the oxoG•dCTP complex. Thus, in addition to the flexibility of the long linker (residues 18 to 34, disordered electron map density) connecting

ModN to N-helix, the alignment of the N-helix is itself variable in PrimPol. The N-helix bears some resemblance to the PAD (or the little finger domain) in Y-family polymerases<sup>32</sup> by way of its location and placement in the DNA major groove, but it lacks the multiple contacts that the PAD forms with the backbone of the DNA. Further, we observe ordered electron density for the PrimPol-unbound end of the DNA duplex, which is engaged in end-to-end DNA–DNA interactions via two stacking two-tier triplexes that promote crystal formation (Supplementary Note 1 and Supplementary Fig. 2).

The template oxoG base assumes the *anti* conformation about the glycosidic bond and forms Watson–Crick (WC) interactions with the incoming dCTP (Figs. 1b, c). The C1'–C1' distance in the nascent oxoG(*anti*)-dCTP base pair is 11 Å (Fig. 1c), which is slightly longer than the C1'–C1' distance in the T-dATP pair of the unmodified PrimPol complex (10.4 Å) and in a standard WC pair (10.5 Å). The phosphate group of the oxoG(*anti*) shifts by ~1.5 Å to avoid a close contact between the O8 and the O5' atoms of the oxoG residue (Fig. 1d), and thereby resolving the intrinsic steric clash of the oxoG residue in the *anti* conformation. That is, this oxygen-to-oxygen atom distance would be too close (at ~2.6 Å) if the placement of the phosphate group was preserved as it is for the template base in the unmodified PrimPol complex (Fig. 1d). Thus, the O8 and the O5' of the oxoG are now located ~3.5 Å from each other. Since the phosphate group of the template residue in either complex does not make any interactions with PrimPol protein residues, the shift of the phosphate group has little implication for catalytic alignment. This is reflected in the relatively modest reduction (~26%) in the efficiency of incorporation of dCTP opposite the oxoG compare to the unmodified G template base<sup>22</sup>.

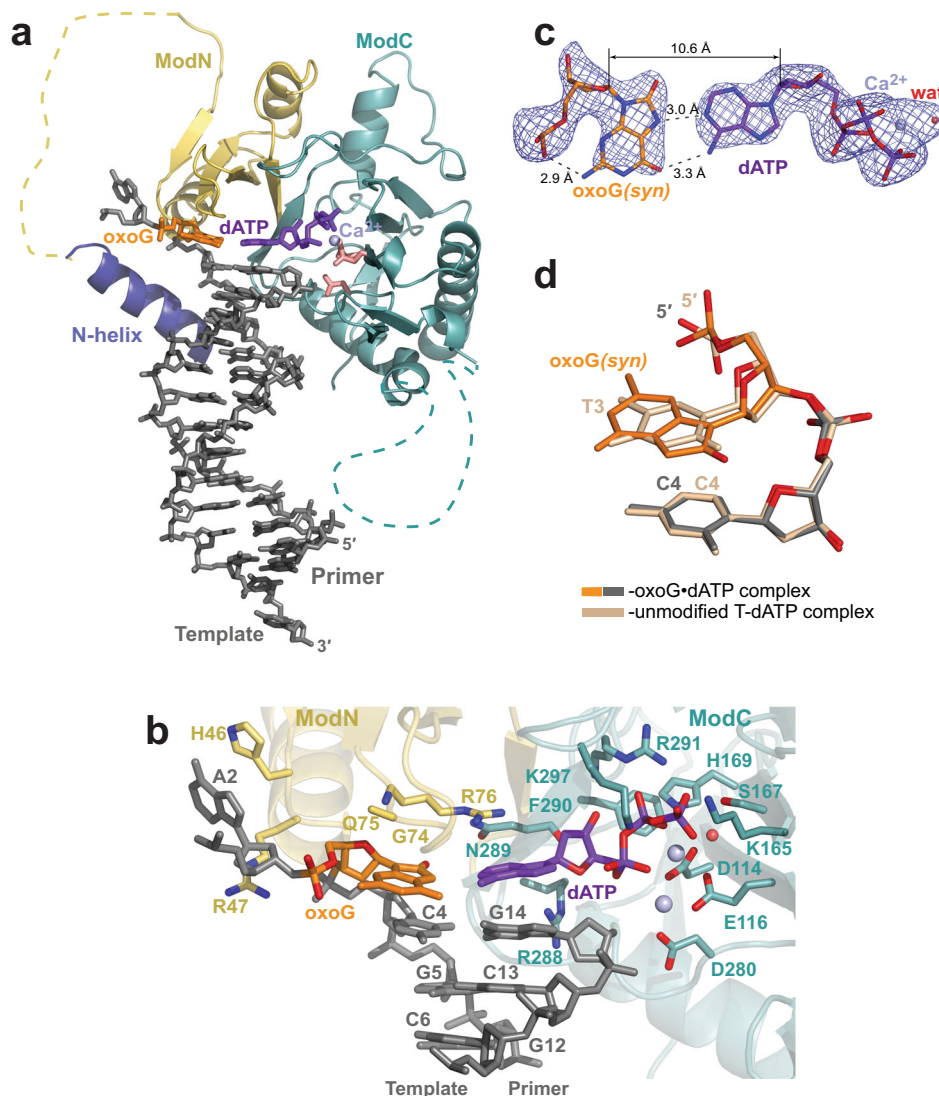


**Fig. 1** Structure of human PrimPol inserting correct dCTP opposite the oxoG template base. **a** The overall structure of the ternary PrimPol oxoG•dCTP complex (molecule A complex in the asymmetric unit (AU) with the ordered N-helix). The N-helix and modules ModN and ModC are shown in cartoon representation in dark blue, yellow, and cyan, respectively. The DNA is shown as gray sticks, and the  $\text{Ca}^{2+}$  ion ('metal B') is shown as a light blue sphere. The oxoG residue (orange) is forming a nascent base pair with the incoming dCTP (red). The unstructured regions in the ModN and ModC modules are shown as dashed yellow and cyan lines, respectively. The side chains of key catalytic active-site residues Asp114, Glu116, and Asp280 are shown as red sticks. **b** A close-up view of the oxoG template base and the dCTP in the PrimPol active site. The key catalytic residues (Asp114, Glu116, and Asp280), the residues contacting the incoming dCTP (Lys165, Ser167, His169, Arg288, Asn289, Phe290, Arg291, and Lys297), the template oxoG base (Gly74, Gln75, and Arg76), and the rest of the template strand (Lys10, His46, and Arg47) are shown in sticks and with oxygen atoms in bright red and nitrogen atoms in blue. **c** A simulated annealing  $F_o - F_c$  omit map (contoured at  $2.5\sigma$ -level at 2.60 Å resolution and colored in blue) showing the clear electron density for the entire oxoG residue in the *anti* conformation, its partner incoming dCTP, and the  $\text{Ca}^{2+}$  ion. Hydrogen bonds are indicated by black dashes. **d** Comparison of the phosphate backbone conformations of C2-oxoG3(*anti*)-C4 (gray and orange) and unmodified C2-T3-C4 (beige) template strand segments within the PrimPol active site of their respective ternary complexes. The complexes are superimposed by ModN and ModC of PrimPol protein. The oxygen atoms of the sugar-phosphate backbone and the O8 of the oxoG residue are colored in red. The phosphate group of the oxoG3(*anti*) residue is shifted by  $-1.5$  Å compared to the position occupied by the phosphate group of T3 residue.

**Mutagenic A insertion opposite oxoG.** We obtained two PrimPol complexes depicting the insertion of dATP opposite the oxoG template residue. The first complex has a 2',3'-dideoxy-3'-terminated primer and is solved to higher 2.4 Å resolution, and the second complex contains a primer with an intact 3'-OH end and is solved to lower 2.6 Å resolution. The overall structures of both complexes are very similar to the unmodified structure (Fig. 2a and Supplementary Fig. 3a) and to each other (Supplementary Fig. 3b), including the placement of the N-helix. The end-to-end DNA–DNA duplex packing interactions, producing a three-tier triple helix, are well-resolved (Supplementary Note 2 and Supplementary Fig. 4).

There are two  $\text{Ca}^{2+}$  ions in molecule B of the oxoG•dATP 3'-OH-primer-containing complex (Fig. 2b), corresponding to 'metal A' and 'metal B' in DNA polymerases, and indicating the readiness of the active site to proceed with the dATP insertion. The template oxoG base assumes the *syn* conformation about the glycosidic bond and uses its Hoogsteen edge to pair with the incoming dATP (Fig. 2b, c). Such arrangement is evident from the well-ordered electron density map for both residues of the pair (Fig. 2c) and is observed in both types of the PrimPol oxoG•dATP complexes. Thus, the protonated N7 atom of the oxoG residue forms a hydrogen bond with the N1 of the dATP





**Fig. 2** Structure of human PrimPol inserting incorrect dATP opposite the oxoG template base. **a** The overall structure of the ternary PrimPol oxoG•dATP complex with the 2',3'-dideoxy 3'-terminated primer (molecule A of the AU). The N-helix and modules ModN and ModC are shown in cartoon representation in dark blue, yellow, and cyan, respectively. The DNA is shown as gray sticks, the  $\text{Ca}^{2+}$  ion ('metal B') is shown as a light blue sphere, and the water molecule coordinated by the  $\text{Ca}^{2+}$  ion is shown as a small red sphere. The oxoG residue (orange) is forming a nascent base pair with the incoming dATP (purple). The unstructured regions in the ModN and ModC modules are shown as dashed yellow and cyan lines, respectively. The side chains of key catalytic active-site residues Asp114, Glu116, and Asp280 are shown as red sticks. **b** A close-up view of the oxoG template base and the dATP in the PrimPol active site of the 2'-deoxy-3'-terminated primer (intact 3'-OH group)-containing complex (molecule B of the AU). There are two  $\text{Ca}^{2+}$  ions (metals 'A' and 'B') in this complex. The key catalytic residues (Asp114, Glu116, and Asp280), the residues contacting the incoming dCTP (Lys165, Ser167, His169, Arg288, Asn289, Phe290, Arg291, and Lys297), the template oxoG base (Gly74, Gln75, and Arg76), and the rest of the template strand (Lys10, His46, and Arg47) are shown in sticks and with oxygen atoms in bright red and nitrogen atoms in blue. The side chains of Gly74 and Gln75 are disordered in this complex and, thus, not shown. **c** A simulated annealing  $F_o - F_c$  omit map (contoured at  $2.5\sigma$ -level at 2.40 Å resolution and colored in blue) showing the clear electron density for the entire oxoG residue in the *syn* conformation, its partner incoming dATP and the  $\text{Ca}^{2+}$  'metal B' ion observed in the complex with the 2',3'-dideoxy 3'-terminated primer. Hydrogen bonds are indicated by black dashes. Such oxoG(*syn*)•dATP base pair arrangement is observed in all mentioned PrimPol oxoG•dATP complexes. **d** Similar phosphate backbone conformations of C2-oxoG3(*syn*)-C4 (gray and orange) and unmodified C2-T3-C4 (beige) template strand segments within the PrimPol active site of their respective ternary complexes. The complexes are superimposed by ModN and ModC of PrimPol protein.

(acceptor) and the  $O^6$  atom of the oxoG (acceptor) interacts with the  $N^6$  of the dATP (donor). The  $C1'-C1'$  distance in this oxoG(*syn*)•dATP base pair is 10.6 Å, allowing for a perfect fit into the active site of PrimPol. There is also a stabilizing intramolecular bond within the oxoG(*syn*) residue, between the  $N^2$  atom of the base and the oxygen atom of its phosphate group.

While the oxoG(*syn*) residue fits the PrimPol active site without altering any active site residues (Fig. 2b) or template

strand backbone conformation (Fig. 2c), the efficiency of dATP incorporation opposite the oxoG is markedly reduced compared to that of dCTP, namely 5.7-fold less in the presence of  $\text{Mg}^{2+}$  and 1.8-fold less in the presence of  $\text{Mn}^{2+}$ <sup>22</sup>. From the structure, this reduction in incorporation of dATP versus dCTP is not caused by any structural interference per se; instead, it may reflect the reduced thermodynamic stability of an oxoG•dATP Hoogsteen base-pair compared to an oxoG•dCTP WC base-pair<sup>43</sup>.

**Extension from the correct C opposite oxoG.** The 2.1 Å resolution structure of PrimPol, extending a correct primer C opposite the oxoG, depicts the next step in accurate bypass of this oxidative lesion (Fig. 3a). The structure of the extension complex is almost identical to the unmodified complex (rmsd of ~0.46 Å over 249 Cas of molecule A, for example; Supplementary Fig. 5).

In the oxoG•C extension complex, PrimPol forms the next correct nascent base pair (Fig. 3a, b). Thus, the T3 template base, which is 5'-adjacent to the oxoG residue, is now paired with the incoming dATP. Both 'metal A' and 'B' Ca<sup>2+</sup> ions are observed in molecule B of the complex (Fig. 3b).

The oxoG(*anti*) at the extension position of the active site (Fig. 3b) is in a Watson–Crick base pair configuration with the 2'-deoxy-3'-C14 terminal primer base (Fig. 3c). Surprisingly, the conformation of the phosphate group of the oxoG is similar to that of the phosphate of an unmodified residue at the same position of the PrimPol complex (the C template strand residue in the undamaged complex) (Fig. 3d). The steric clash of the O8 atom with the sugar-phosphate backbone that arises with oxoG (*anti*) is alleviated by adjustments of the backbone torsion angles of the phosphate group and a slight rotation of the sugar ring compared to the values observed for B-DNA. Thus, the conformation of the sugar-phosphate backbone of the template base at the extension position of the PrimPol active site is preset to accommodate the oxoG(*anti*). Consequently, the main-chain nitrogen atom of Gln48 forms a hydrogen bond with the phosphate group of either the oxoG(*anti*) or an unmodified residue in a similar manner (Fig. 3d). The guanidinium group of Arg47 of ModN, however, repositions to contact a neighboring protein chain in the crystal (main-chain oxygen atom of Gly124 of a symmetry-related molecule, not shown).

Overall, the oxoG(*anti*) template strand residue is accommodated at the extension position without significant adjustments to the active site or to the position of its partner primer terminal C residue, compared to the unmodified structure. This correlates with the near-normal efficiency extension from oxoG•C pair by PrimPol (~75–85% of that observed for undamaged DNA)<sup>22</sup>.

**Extension from the incorrect A opposite oxoG.** The structure of PrimPol extending from an incorrect primer A base opposite the oxoG is solved to 2.07 Å resolution (Fig. 4a and Supplementary Fig. 6). The oxoG residue is in the *syn* conformation and uses its Hoogsteen edge to form a base pair with the 3'-terminal primer base A14 (Fig. 4b). In molecule B complex in the AU, the oxoG•A mispair makes both hydrogen bonds expected for a Hoogsteen base-pair (Fig. 4c), while in molecule A the mispair is more open with the O<sup>6</sup> atom of the oxoG too far to interact with the N<sup>6</sup> of A14 (Fig. 4d). Interestingly, in both molecules, there is no direct intermolecular hydrogen bond between the N<sup>2</sup> of the oxoG(*syn*) base and its nearest backbone oxygen atom (Fig. 4c, d). This is in contrast to the oxoG(*syn*) residue in the oxoG•dATP insertion complex (Fig. 2c), where these two atoms are closer due to a different backbone conformation. Nevertheless, there is a water molecule bridging the N<sup>2</sup> and the oxygen of the phosphate group of the oxoG(*syn*) in the molecule B complex (Fig. 4c). The overall conformation of the sugar-phosphate backbone of the oxoG(*syn*) template base at the extension position is very similar to the one in the unmodified complex and does not disrupt PrimPol active site (Fig. 4e).

Intriguingly, while 'metal A' and 'metal B' Ca<sup>2+</sup> ions are in a similar configuration in molecules A and B of the complex, the position of the 3'-OH of the primer terminus varies (Fig. 5a, b, top panels). In molecule A, even though the oxoG(*syn*) mispair with A14 is somewhat open and held by only a single hydrogen bond (Figs. 4d and 5a, bottom panel), the 3'-OH of A14 is directly

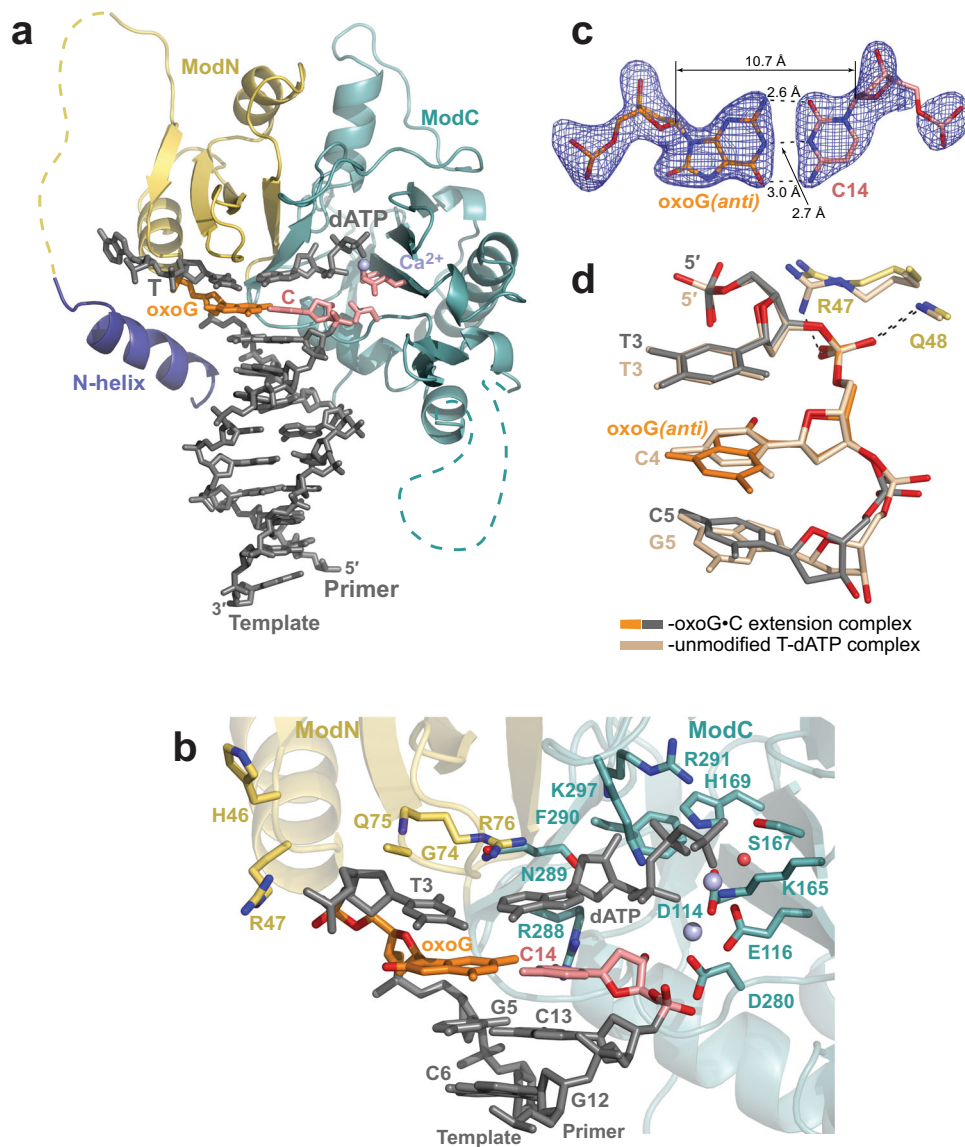
below the α-phosphate of the dATP, in an orientation favorable for an in-line reaction (Fig. 5a, top and bottom panels). By contrast, the more stable oxoG(*syn*)•A mispair with two hydrogen bonds in molecule B (Figs. 4c and 5b, bottom panel), is achieved by rotation of the primer A14 residue clockwise relative to the protein active site, so the A14 is able to fulfill both hydrogen bonds with oxoG(*syn*) (Fig. 5a, b, bottom panels). The rotation of the primer terminal A14 is evident by an increase of a twist angle between the base of A14 relative to the base of the incoming dATP that is positioned similarly in both molecules, from 25° in molecule A complex to 37° in the molecule B complex. Consequently, the rotation causes the 3'-OH of A14 in molecule B complex to shift away from the optimal position observed in molecule A. A new water molecule coordinated by the 'A' ion, labeled 'wat 1' on Fig. 5b top panel, takes nearly the same place as the 3'-OH in the molecule A. The displaced 3'-OH in molecule B is further away from the catalytic Ca<sup>2+</sup> ion 'A' and thus is poorly positioned for metal ion 'A'-supported deprotonation needed to initiate the nucleotidyl-transfer reaction<sup>44</sup>. Taken together, the protrusion of the oxoG(*syn*) into the major groove causes either the decreased stability of the oxoG(*syn*)•A mispair or the misalignment of the 3'-OH primer terminus of the partner A residue. These factors appear to be responsible for 1.8- (Mg<sup>2+</sup>-reaction condition) to 2.1-fold (Mn<sup>2+</sup>-reaction condition) reduction of extension efficiency from an A compare to a C base opposite the oxoG lesion<sup>22</sup>.

## Discussion

7,8-dihydro-8-oxoguanine (oxoG) is a ubiquitous DNA damage lesion produced due to interactions of reactive oxygen species (ROS) with nuclear and mitochondrial DNA in human cells. Here, we provide the structural basis for human archaeo-eukaryotic primase-polymerase PrimPol's ability to efficiently bypass the oxoG lesion in a predominantly error-free manner. We show that PrimPol accommodates an oxoG lesion in the *anti* conformation at both the insertion and the extension steps of lesion bypass, without altering a catalytically-ready alignment of the active site. This allows PrimPol to conduct a proficient error-free replication through the lesion site by insertion of a correct C base opposite the oxoG(*anti*) and extension from it. Overall, the 'minimal' architecture of PrimPol whereby it is devoid of contacts with the major and minor grooves of the DNA as well as the sugar-phosphate backbone of the templating nucleotide<sup>35</sup> is beneficial for nearly unobstructed and correct lesion bypass of oxoG.

We also find that the mutation-inducing insertion of dATP opposite the oxoG(*syn*) does not cause any significant structural hindrance within the PrimPol active site. Why then is the efficiency of dATP incorporation opposite the oxoG reduced ~6-fold compared to dCTP<sup>22</sup>? One factor may be the lower thermodynamic stability of an oxoG•dATP nascent base-pair relative to that of an oxoG•dCTP base-pair<sup>43</sup>. Also, it is possible that a conformational *anti-syn* equilibrium of an unpaired oxoG template residue in a binary complex might influence a selection of an incoming nucleotide. For example, the template oxoG can exist in both conformations within the active site of X-family Pol β, but only the oxoG(*syn*) conformation is stabilized via hydrogen bonding interactions, and thereby promoting the binding and insertion of dATP<sup>45</sup>. Since PrimPol is devoid of amino acids interacting with the templating base, oxoG(*anti*) may be favored over oxoG(*syn*) due to stacking interactions with the neighboring 3'-template base, leading to the preferred incorporation of dCTP over dATP.

PrimPol is also reduced in its efficiency (~2-fold) in extending DNA synthesis from an A compared to a C base opposite the oxoG lesion<sup>22</sup>. We show here that during the extension step,



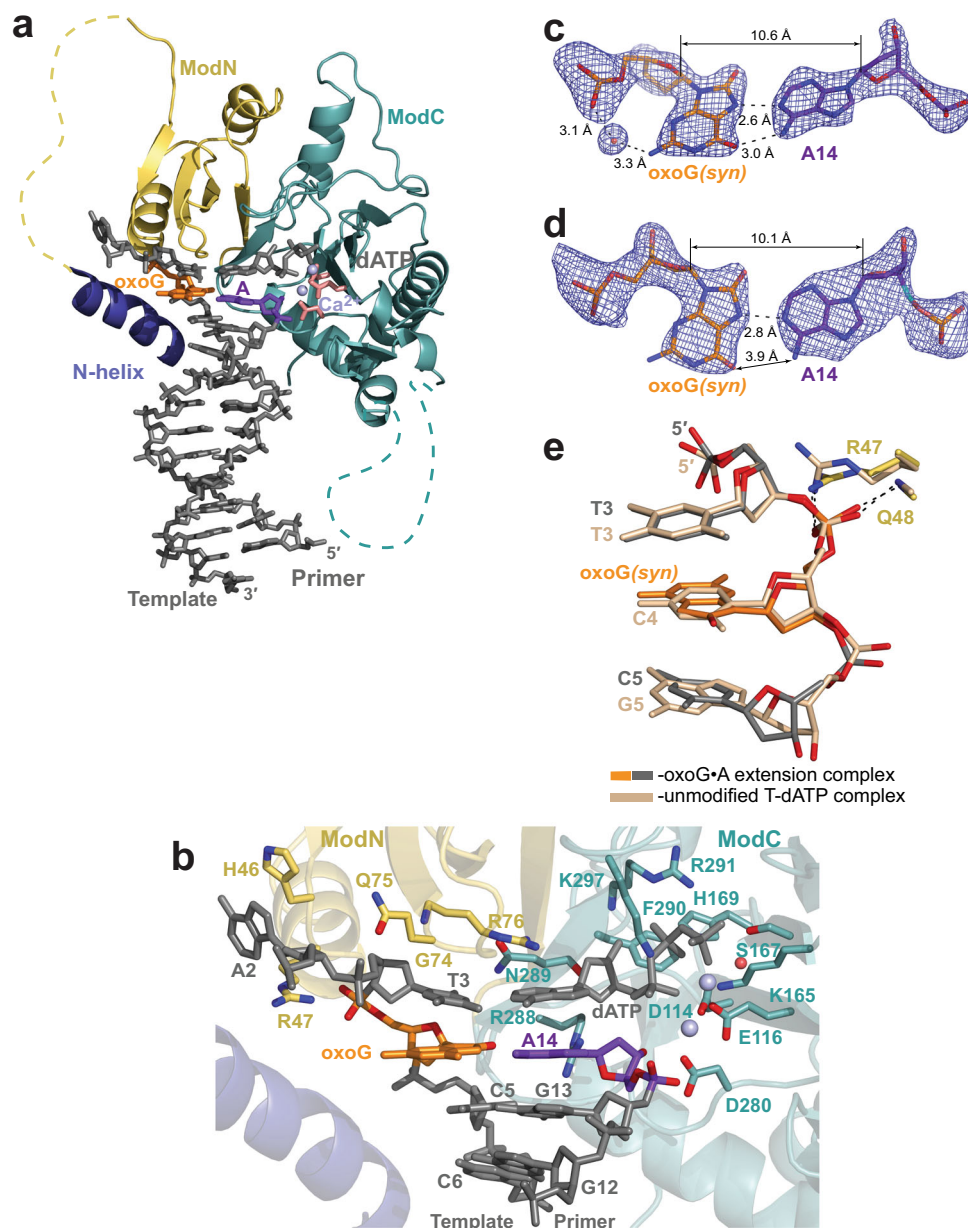
**Fig. 3** Structure of human PrimPol extending from a correct primer C base opposite the oxoG template base. **a** The overall structure of the ternary PrimPol oxoG•C extension complex (molecule A of the asymmetric unit, AU). The N-helix and modules ModN and ModC are shown in cartoon representation in dark blue, yellow, and cyan, respectively. The DNA is shown as gray sticks, the  $\text{Ca}^{2+}$  ion ('metal B') is shown as a light blue sphere. The oxoG residue (orange) is forming a base pair with the primer terminal C base (red). The incoming dATP forms a nascent base pair with a template T base, which is 5'-adjacent to the oxoG residue. The unstructured regions in the ModN and ModC modules are shown as dashed yellow and cyan lines, respectively. The side chains of key catalytic active-site residues Asp114, Glu116, and Asp280 are shown as red sticks. **b** A close-up view of the oxoG template base and the 3'-deoxy primer C14 base in the PrimPol active site (molecule B in the AU). There are two  $\text{Ca}^{2+}$  ions (metal 'A' and 'B') in this complex shown as light blue spheres, and the water molecule coordinated by the  $\text{Ca}^{2+}$  ion is shown as a small red sphere. The key catalytic residues (Asp114, Glu116, and Asp280), the residues contacting the incoming dATP (Lys165, Ser167, His169, Arg288, Asn289, Phe290, Arg291, and Lys297), the templating base (Gly74, Gln75, and Arg76), and the rest of the template strand (Lys10, His46, and Arg47) are shown in sticks and with oxygen atoms in bright red and nitrogen atoms in blue. **c** A simulated annealing  $F_o - F_c$  omit map (contoured at  $2.5\sigma$ -level at 2.10 Å resolution and colored in blue) showing the clear electron density for the entire oxoG residue in the *anti* conformation and its partner C14 residue. Hydrogen bonds are indicated by black dashes. **d** Similar phosphate backbone conformations of T3-oxoG4(*anti*)-C5 (gray and orange) and unmodified T3-C4-G5 (beige) template strand segments within the PrimPol active site of their respective ternary complexes. The complexes are superimposed by ModN and ModC of PrimPol protein. The main-chain nitrogen atom of Gln48 contacts the oxygen atom of the phosphate group of the oxoG(*anti*), similarly to its contact with the phosphate group of the C4 in the unmodified complex. The guanidinium group of Arg47 of ModN, however, repositions to contact a neighboring protein chain in the crystal (main-chain oxygen atom of Gly124 of a symmetry-related molecule, not shown).

compared to C opposite oxoG, incorrect A leads to significant deviations from the optimum alignment in the active site. That is, the protrusion of oxoG(*syn*) into the major groove in the extension complex causes either the opening of the oxoG(*syn*)-A pair with a loss of one hydrogen bond between the bases or the misalignment of the 3'-OH primer terminus when the partner A

residue rotates to reestablish both hydrogen bonds within the pair. Together, these structural deviations may underlie the reduced efficiency of PrimPol in extending DNA synthesis from a 3'-terminal A base opposite an oxoG.

The PrimPol structures also shed additional light on the active-site geometry. In the unmodified PrimPol structure<sup>35</sup> we observed



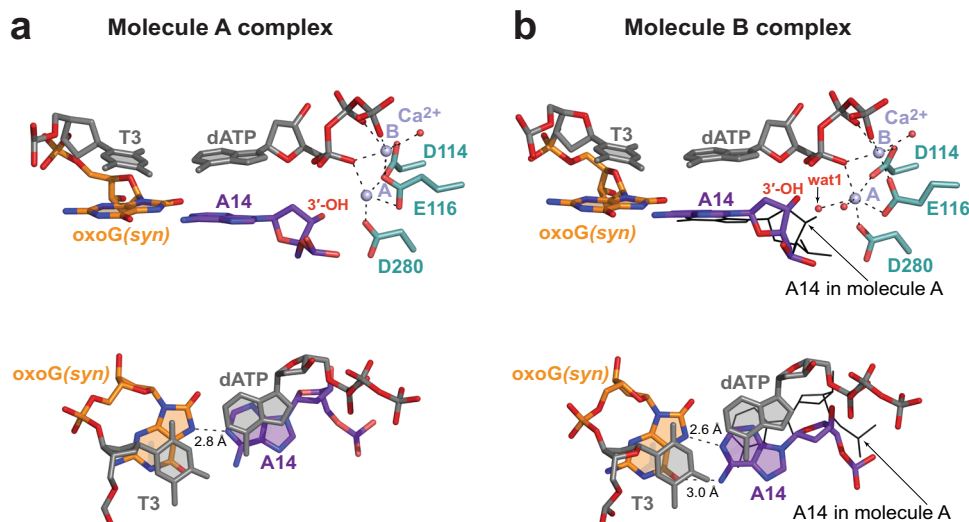


**Fig. 4 Structure of human PrimPol extending from an incorrect primer A base opposite the oxoG template base.** **a** The overall structure of the ternary PrimPol oxoG•A extension complex (molecule A of the asymmetric unit, AU). The N-helix and modules ModN and ModC are shown in cartoon representation in dark blue, yellow, and cyan, respectively. The DNA is shown as gray sticks, the two  $\text{Ca}^{2+}$  ions (metals 'A' and 'B') are shown as light blue spheres. The oxoG residue (orange) is forming a base pair with the primer terminal A base (purple). The incoming dATP forms a nascent base pair with a template T base, which is 5'-adjacent to the oxoG residue. The unstructured regions in the ModN and ModC modules are shown as dashed yellow and cyan lines, respectively. The side chains of key catalytic active-site residues Asp114, Glu116, and Asp280 are shown as red sticks. **b** A close-up view of the oxoG template base and the 3'-deoxy primer C14 base in the PrimPol active site (molecule A in the AU). Two  $\text{Ca}^{2+}$  ions (metals 'A' and 'B') in this complex shown as light blue spheres, and the water molecule coordinated by the  $\text{Ca}^{2+}$  ion is shown as a small red sphere. **c** A simulated annealing Fo–Fc omit map (contoured at  $2.5\sigma$ -level at 2.07 Å resolution and colored in blue) showing the clear electron density for the oxoG residue in the *syn* conformation and its partner A14 residue in molecule B of the AU. The water molecule that facilitates the intramolecular bond between the  $\text{N}^2$  and the phosphate group of the oxoG residue is observed only in molecule B. **d** A simulated annealing Fo–Fc omit map (contoured at  $2.5\sigma$ -level) showing the electron density for the oxoG (*syn*) and its partner A14 residues in molecule A of the AU. There is only one hydrogen bond in this base pair. **e** Similar phosphate backbone conformations of T3-oxoG4(*syn*)-C5 (gray and orange) and unmodified T3-C4-G5 (beige) template strand segments within the PrimPol active site of their respective ternary complexes. The main-chain nitrogen atom of Gln48 and the side chain of Arg47 contact the phosphate group of the oxoG(*syn*) similarly to the protein contacts with the phosphate group of C4 in the unmodified complex.

only one  $\text{Ca}^{2+}$  'ion B' chelated by the oxygen atoms of the triphosphate tail of the incoming dATP as assumed due to the use of 2',3'-dideoxy primer terminus. Interestingly, in the current study, with the 'native' 3'-OH-2'-deoxy primers, a single  $\text{Ca}^{2+}$  'ion B' is detected in both complex molecules of the asymmetric

unit (AU) of the oxoG•dCTP insertion complex and in molecules 'A' of the oxoG•dATP insertion and the oxoG•C extension complexes. Thus, binding of dNTP together with 'ion B' to PrimPol-DNA appears to be the first step in ternary PrimPol complex assembly, similar to that observed for Y-family human





**Fig. 5** Correct and misaligned placement of the 3'-OH primer terminus in molecule A and B complexes in the asymmetric unit of the oxoG•A extension complex. **a** Molecule A complex. Top panel: the side view parallel to the DNA duplex axis. Two Ca<sup>2+</sup> ions (metals 'A' and 'B') shown as light blue spheres, and the water molecule coordinated by the Ca<sup>2+</sup> ion is shown as a small red sphere. Bottom panel: the view perpendicular to the DNA duplex axis, highlighting the relative positions of dATP and the 3'-primer terminal A14. **b** Molecule B complex. Top and Bottom panels as described above. The placement of A14 residue of molecule A complex is shown in black lines. The A and B molecules are superimposed by PrimPol protein. The oxygen atoms of the sugar-phosphate backbone are colored in red.

Pol  $\eta$  by time-resolved X-ray crystallography<sup>46,47</sup>. In the PrimPol-DNA-dNTP-'ion B' complexes, the distances of primer 3'-OH to  $\alpha$ -phosphate (Pa) of the incoming dNTP vary from 4.29 to 4.70 Å. Unexpectedly, the entrance of the second Ca<sup>2+</sup> 'catalytic ion A' detected in the extension oxoG•C molecule B complex and oxoG•A molecule A complex change the 3'-OH to Pa distances only slightly to 3.99 and 4.11 Å, respectively. Such 3'-OH to Pa distances are at least  $\sim 0.7$  Å too long for phosphoryl transfer to occur. This is different compared to Pol  $\eta$ , where the ternary complex with a single Ca<sup>2+</sup> 'ion B' has the 3'-OH to Pa distance of 4.20 Å, which upon the binding of the second ion (Mg<sup>2+</sup>) at the catalytic 'site A' and exchange of 'ion B' to Mg<sup>2+</sup> decreases to  $\sim 3.30$  Å and then is further reduced by 0.5 Å during the nucleotidyl-transfer reaction<sup>46</sup>. Thus, the active site of PrimPol, even after the binding of both ions, appears to be less tightly organized than Pol  $\eta$ 's and possibly more dependent on slow thermal motion to align the catalytic groups. The slow catalytic tuning might represent the 'rate-limiting conformational transition step' detected for PrimPol during a pre-steady-state kinetics study<sup>48</sup>.

High amounts of ROS<sup>14</sup> and limited repair capabilities<sup>26,27</sup> make mitochondrial DNA especially vulnerable to oxidative damage<sup>15</sup>. Mitochondrial replicative Pol  $\gamma$  is strongly inhibited by the oxoG lesion and prefers to extend from the oxoG•A base pair<sup>11,42</sup>, and thus is unlikely to be involved in error-free oxoG bypass in vivo. PrimPol, may partake in that role. It is the only TLS DNA polymerase known to localize to mitochondria and is able to efficiently and mostly accurately bypass the oxoG lesions<sup>21–25</sup>. While PrimPol does not appear to rescue in vitro reconstituted mitochondrial replisome stalled at the oxoG lesion site<sup>42</sup>, it is likely that additional accessory proteins and different biochemical conditions within mitochondria in vivo might facilitate the bypass efficiency.

Taken together, PrimPol has emerged as a unique primase-polymerase that can partake in direct TLS of small DNA lesions such as oxoG. We provide here the structural underpinnings for how PrimPol is able to conduct proficient and mostly error-free oxoG lesion bypass.

## Methods

**Protein purification.** The catalytic domain of human PrimPol (residues 1–354) was expressed in *Saccharomyces cerevisiae* as an N-terminally tagged glutathione S-transferase (GST) fusion protein and then purified as described previously<sup>35</sup>. The yeast cell pellets were resuspended in lysis buffer (50 mM TRIS-HCl (tris(hydroxymethyl)aminomethane) pH 8.0, 10% sucrose, 150 mM NaCl, 175 mM ammonia sulfate, 1 mM EDTA, 10 mM  $\beta$ -mercaptoethanol supplemented with protease inhibitors, 1 mg/mL leupeptin, 5 mg/mL chymostatin, 1 mg/mL aprotinin, 1 mg/mL pepstatin, and 0.078 mg/mL benzamidin) and lysed with a bead beater. Total cellular proteins were precipitated by ammonium sulfate, resuspended, and loaded onto a glutathione-Sepharose gravity flow column in buffer A (50 mM TRIS-HCl pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM DTT). The GST-PrimPol fusion protein bound to the resin was washed with buffer A, buffer A supplemented with a total of 1 M NaCl, and again with buffer A. PrimPol protein (carrying an extra GPGGDPH peptide on its N terminus due to the expression construct design) was eluted with buffer A after on-column overnight cleavage of the GST tag with PreScission protease. PrimPol protein was then further purified on HiTrap Heparin column (GE Healthcare). The binding buffer for the Heparin column contained 25 mM TRIS-HCl (pH 7.5), 150 mM NaCl, and 2 mM DTT, and the elution buffer consisted of 25 mM TRIS-HCl (pH 7.5), 1 M NaCl, and 2 mM DTT. Finally, PrimPol protein was purified by gel-filtration on Superdex 75 column (GE Healthcare) in 25 mM TRIS (pH 7.5), 250 mM NaCl, and 2 mM tris(2-carboxyethyl) phosphate (TCEP), concentrated to  $\sim 35$  mg/ml, and stored in aliquots at  $-80$  °C.

**Crystallization.** The crystals of the ternary complex with the correct C incoming nucleotide triphosphate (dCTP) opposite the template oxoG residue were obtained by incubating the human PrimPol catalytic core with a 17-nucleotide (nt) DNA template (5'-CA(oxoG)CGCTACCACACCCC-3') (TriLink Biotechnologies Inc.) annealed to a 2'-deoxy-3' terminated 12 nt DNA primer (5'-GGTGTGGTAGCG-3') (Glen Research, Inc) (Supplementary Table 1) in a presence of dCTP. The crystal growth and harvesting were performed as described in our study of PrimPol structure with unmodified DNA<sup>35</sup>. Annealed template-primer DNA was mixed with PrimPol in a 1.2:1 molar ratio to a final concentration of 0.3 mM in 25 mM TRIS-HCl (pH 7.5), 155 mM NaCl, 1.5 mM TCEP, 12 mM CaCl<sub>2</sub>, and 7 mM dCTP. The complexes were incubated at room temperature for 5 min and then centrifuged at 9,300  $\times$ g for 10 min at 4 °C. The crystals were grown at 20 °C by a hanging drop method against a reservoir solution containing 225–250 mM CaCl<sub>2</sub> and 16–19% PEG 3350. To produce bigger crystals (rod-like, up to  $\sim 0.2$  mm in length), a round of microseeding was performed. The crystals were cryoprotected stepwise in reservoir solution supplemented with 24% PEG 3350 and 10% and then 20% glycerol and flash-frozen in liquid nitrogen for X-ray data collection.

The crystals of the PrimPol ternary complex with incoming dATP opposite the template oxoG were produced with the aforementioned 17 nt template and 2'-deoxy- or 2',3'-dideoxy-3' terminated 13 nt DNA primer, namely 5'-GGGTGTGGTAGCG-3' (Glen Research, Inc) or 5'-GGGTGTGGTAGCG<sub>ddT</sub>-3' (W.M. Keck Foundation,

Yale). The extension ternary complexes were obtained with a 17 nt DNA template (5'-CAT(oxoG)CCTACCACACCCC-3'), a 13 nt DNA primers with either a C or an A 2'-deoxy 3'-terminal bases (5'-GGGTGTGGTAGGX-3', where X is C or A), and the next correct incoming base dATP. The crystals were obtained as described above for the oxoG•C complex.

**Structure determination and refinement.** The X-ray diffraction data were collected at the 24-ID NE-CAT beamline at Advanced Photon Source in Chicago and were processed by the RAPD pipeline (<http://necat.chem.cornell.edu/>). We solved the structure of the PrimPol ternary complex with the correct incoming dCTP opposite the template oxoG base (oxoG•dCTP) by the molecular replacement (MR) method (Phaser)<sup>49</sup> in the Phenix program package<sup>50</sup> using the PrimPol ternary complex structure with the unmodified DNA (PDB ID: 5L2X)<sup>35</sup> as a search model. The model building was finished manually in Coot<sup>51</sup> based on the electron density maps calculated in PHENIX Refine<sup>50</sup>. The final model of the oxoG•dCTP complex was refined in PHENIX Refine to 2.6 Å resolution and belongs to P1 space group with unit cell dimensions of  $a = 52.6$  Å,  $b = 65.7$  Å,  $c = 75.2$  Å,  $\alpha = 69.7^\circ$ ,  $\beta = 82.8^\circ$ , and  $\gamma = 88.5^\circ$ . The structure is refined to  $R_{\text{free}}$  of 28.6% and  $R_{\text{work}}$  of 23.9% and consists of two PrimPol molecules, two DNA template-primer molecules, two dCTP, two  $\text{Ca}^{2+}$  ions, and a total of 92 solvent molecules. The placement and conformation of the oxoG and dCTP residues was verified using simulated annealing omit maps calculated in PHENIX<sup>50</sup> with the residues omitted from the model before heating to 2,000 K and then slowly cooling. The average B-factor for the portion of the DNA interacting with PrimPol protein via the sugar-phosphate backbone (residues oxoG3-C4-G5-C6 of the template and G14 of the primer strands) is lower (68.9 Å<sup>2</sup>) than the overall average B-factor for the DNA (115.3 Å<sup>2</sup>) and is similar to the average B-factor of the protein (64.4 Å<sup>2</sup>). The relatively high B-factors for 'unbound' portions of the DNA reflect their conformational flexibility in the absence of direct contact with PrimPol.

The crystals with the incoming dATP opposite the template oxoG and containing 2',3'-dideoxy-3' terminated 13 nt DNA primer, diffracted to 2.37 Å resolution and belong to P1 space group with the unit cell parameters similar to the oxoG•C complex. We have used the oxoG•dCTP complex as a search model to solve the structure by MR. The oxoG•dATP structure is refined to  $R_{\text{free}}$  of 27.7% and  $R_{\text{work}}$  of 22.0% and consists of two PrimPol molecules, two DNA template-primer molecules, two dATPs, three  $\text{Ca}^{2+}$  ions, and a total of 106 solvent molecules.

The structure of the complex with the incoming dATP opposite the template oxoG and containing 3'-deoxy-3' terminated 13 nt DNA primer, was refined to 2.58 Å resolution, also in P1 space group with similar unit cell dimensions. The structure is refined to  $R_{\text{free}}$  of 30.7% and  $R_{\text{work}}$  of 24.8% and consists of two PrimPol molecules, two DNA template-primer molecules, two dATPs, three  $\text{Ca}^{2+}$  ions, and 24 water molecules. Similar to the oxoG•dCTP complex, in the oxoG•dATP structures, the average B-factors for the portions of the DNA interacting with PrimPol protein are comparable to that of the protein, whereas B-factors for the regions of the DNA that do not interact with PrimPol are higher, reflecting their greater conformational freedom.

The oxoG•C PrimPol extension complex structure with a correct primer C opposite the template oxoG and the next correct incoming nucleotide (dATP) was also determined by MR. Because the unit cell dimensions of the complex were isomorphous to the unit cell dimensions of the parent structure with undamaged DNA (5L2X), we imported the R-free set from 5L2X using the Phenix reflection file editor. We also conducted simulated annealing on the coordinates after MR, and rebuilt the model manually in Coot<sup>51</sup> using the electron density maps calculated in PHENIX Refine<sup>50</sup>. The oxoG•C structure was refined to 2.1 Å resolution with several cycles of TLS refinement conducted at the end of the refinement process. After TLS refinement,  $R_{\text{free}}$  and  $R_{\text{work}}$  are 21.3 and 18.7%, respectively, and the final model consists of two PrimPol molecules, two DNA template-primer molecules, two dATPs, four  $\text{Ca}^{2+}$  ions, and a total of 162 solvent molecules. The oxoG•A PrimPol extension complex structure with an incorrect primer A opposite the template oxoG and the next incoming nucleotide (dATP) was similarly determined by MR with 5L2X as a search model. Again, because the unit cell dimensions of the complex were very close to the unit cell dimensions of 5L2X, we imported the R-free set from the parent structure and followed the refinement process described above for the oxoG•C extension complex, to a resolution of 2.07 Å. After TLS refinement, the  $R_{\text{free}}$  and  $R_{\text{work}}$  are 21.9 and 20.1%, respectively. The final model consists of two PrimPol, two DNA template-primer molecules, two dATP, four  $\text{Ca}^{2+}$  ions, and a total of 131 solvent molecules. In the extension complexes, the B-factors for the protein-bound DNA nucleotides are lower than for the DNA overall, and are closer in values to the protein residues.

Curiously, the validation reports calculated upon the deposition of data to Protein Data Bank (PDB), indicate that the extension complex structures have a somewhat high number of RSRZ outlier protein residues (14.0% in the oxoG•C and 13.0% in the oxoG•A). However, most of the flagged residues have an unequivocally good fit to the electron density map. For example, in the oxoG•C complex structure, residues Phe290 in chain A flagged with RSRZ = 4.5 and in chain B flagged with RSRZ = 4.8, have well-defined densities that are equivalent in quality to the densities for the not-flagged Phe166A and Phe166B residues (Supplementary Fig. 7). In addition, Phenix comprehensive validation, indicates

that both Phe290A and Phe290B have a real space correlation coefficient (CC) of 0.98 and 0.99, respectively.

The crystal data, together with the data collection and refinement statistics, are summarized in Table 1. Figures were prepared using PyMOL<sup>52</sup> (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession codes 7JK1, 7JKL, 7JKP, 7JL8 and 7JLG. Other data are available from the corresponding authors upon reasonable request.

Received: 30 July 2020; Accepted: 14 June 2021;

Published online: 29 June 2021

## References

- Cadet, J., Douki, T., Gasparutto, D. & Ravanat, J. L. Oxidative damage to DNA: formation, measurement and biochemical features. *Mutat. Res.* **531**, 5–23 (2003).
- Kozmin, S. et al. UVA radiation is highly mutagenic in cells that are unable to repair 7,8-dihydro-8-oxoguanine in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **102**, 13538–13543 (2005).
- Neeley, W. L. & Essigmann, J. M. Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem. Res. Toxicol.* **19**, 491–505 (2006).
- Mouret, S., Forestier, A. & Douki, T. The specificity of UVA-induced DNA damage in human melanocytes. *Photochem. Photobiol. Sci.* **11**, 155–162 (2012).
- Tubbs, A. & Nussenzweig, A. Endogenous DNA damage as a source of genomic instability in cancer. *Cell* **168**, 644–656 (2017).
- Shibutani, S., Takeshita, M. & Grollman, A. P. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**, 431–434 (1991).
- Einolf, H. J., Schnetz-Boutaud, N. & Guengerich, F. P. Steady-state and pre-steady-state kinetic analysis of 8-oxo-7,8-dihydroguanosine triphosphate incorporation and extension by replicative and repair DNA polymerases. *Biochemistry* **37**, 13300–13312 (1998).
- Einolf, H. J. & Guengerich, F. P. Fidelity of nucleotide insertion at 8-Oxo-7,8-dihydroguanine by mammalian DNA polymerase delta-Steady-state and pre-steady-state kinetic analysis. *J. Biol. Chem.* **276**, 3764–3771 (2001).
- Briebe, L. G. et al. Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *EMBO J.* **23**, 3452–3461 (2004).
- Hsu, G. W., Ober, M., Carell, T. & Beese, L. S. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* **431**, 217–221 (2004).
- Hanes, J. W., Thal, D. M. & Johnson, K. A. Incorporation and replication of 8-oxo-deoxyguanosine by the human mitochondrial DNA polymerase. *J. Biol. Chem.* **281**, 36241–36248 (2006).
- Maga, G. et al. 8-oxo-guanine bypass by human DNA polymerases in the presence of auxiliary proteins. *Nature* **447**, 606–608 (2007).
- Markkanen, E., Castrec, B., Villani, G. & Hubscher, U. A switch between DNA polymerases delta and lambda promotes error-free bypass of 8-oxo-G lesions. *Proc. Natl Acad. Sci. USA* **109**, 20401–20406 (2012).
- Zorov, D. B., Juhaszova, M. & Sollott, S. J. Mitochondrial reactive oxygen species (ROS) and ROS-induced ROS release. *Physiol. Rev.* **94**, 909–950 (2014).
- Kauppila, T. E. S., Kauppila, J. H. K. & Larsson, N. G. Mammalian mitochondria and aging: an update. *Cell Metab.* **25**, 57–71 (2017).
- Kennedy, S. R., Salk, J. J., Schmitt, M. W. & Loeb, L. A. Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. *PLoS Genet.* **9**, e1003794 (2013).
- Itsara, L. S. et al. Oxidative stress is not a major contributor to somatic mitochondrial DNA mutations. *PLoS Genet.* **10**, e1003974 (2014).
- Vermulst, M. et al. Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat. Genet.* **39**, 540–543 (2007).
- Ameur, A. et al. Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genet.* **7**, e1002028 (2011).
- Kauppila, J. H. K. et al. Base-excision repair deficiency alone or combined with increased oxidative stress does not increase mtDNA point mutations in mice. *Nucleic Acids Res.* **46**, 6642–6669 (2018).

21. Garcia-Gomez, S. et al. PrimPol, an archaic primase/polymerase operating in human cells. *Mol. Cell* **52**, 541–553 (2013).
22. Zafar, M. K., Ketkar, A., Lodeiro, M. F., Cameron, C. E. & Eoff, R. L. Kinetic analysis of human PrimPol DNA polymerase activity reveals a generally error-prone enzyme capable of accurately bypassing 7,8-Dihydro-8-oxo-2'-deoxyguanosine. *Biochemistry* **53**, 6584–6594 (2014).
23. Krasich, R. & Copeland, W. C. DNA polymerases in the mitochondria: a critical review of the evidence. *Front. Biosci.* **22**, 692–709 (2017).
24. Bailey, L. J., Bianchi, J. & Doherty, A. J. PrimPol is required for the maintenance of efficient nuclear and mitochondrial DNA replication in human cells. *Nucleic Acids Res.* **47**, 4026–4038 (2019).
25. Wisnovsky, S., Sack, T., Pagliarini, D. J., Laposa, R. R. & Kelley, S. O. DNA polymerase theta increases mutational rates in mitochondrial DNA. *ACS Chem. Biol.* **13**, 900–908 (2018).
26. Quiros, P. M., Mottis, A. & Auwerx, J. Mitonuclear communication in homeostasis and stress. *Nat. Rev. Mol. Cell Biol.* **17**, 213–226 (2016).
27. Guilliam, T. A., Bailey, L. J., Brissett, N. C. & Doherty, A. J. PolDIP2 interacts with human PrimPol and enhances its DNA polymerase activities. *Nucleic Acids Res.* **44**, 3317–3329 (2016).
28. Mouron, S. et al. Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nat. Struct. Mol. Biol.* **20**, 1383–1389 (2013).
29. Wan, L. et al. hPrimPol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. *Embo Rep.* **14**, 1104–1112 (2013).
30. Bianchi, J. et al. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol. Cell* **52**, 566–573 (2013).
31. Martinez-Jimenez, M. I., Calvo, P. A., Garcia-Gomez, S., Guerra-Gonzalez, S. & Blanco, L. The Zn-finger domain of human PrimPol is required to stabilize the initiating nucleotide during DNA priming. *Nucleic Acids Res.* **46**, 4138–4151 (2018).
32. Jain, R., Aggarwal, A. K. & Rechkoblit, O. Eukaryotic DNA polymerases. *Curr. Opin. Struct. Biol.* **53**, 77–87 (2018).
33. Keen, B. A., Jozwiakowski, S. K., Bailey, L. J., Bianchi, J. & Doherty, A. J. Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic Acids Res.* **42**, 5830–5845 (2014).
34. Quinet, A. et al. PRIMPOL-mediated adaptive response suppresses replication fork reversal in BRCA-deficient cells. *Mol. Cell* <https://doi.org/10.1016/j.molcel.2019.10.008> (2019).
35. Rechkoblit, O. et al. Structure and mechanism of human PrimPol, a DNA polymerase with primase activity. *Sci. Adv.* **2**, e1601317 (2016).
36. Burak, M. J., Guja, K. E., Hambardjieva, E., Derkunt, B. & Garcia-Diaz, M. A fidelity mechanism in DNA polymerase lambda promotes error-free bypass of 8-oxo-dG. *Embo J.* **35**, 2045–2059 (2016).
37. Rechkoblit, O. et al. Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion. *PLoS Biol.* **4**, 25–42 (2006).
38. Haracska, L., Yu, S. L., Johnson, R. E., Prakash, L. & Prakash, S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase  $\epsilon$ . *Nat. Genet.* **25**, 458–461 (2000).
39. Silverstein, T. D. et al. Structural basis for error-free replication of oxidatively damaged DNA by yeast DNA polymerase  $\epsilon$ . *Structure* **18**, 1463–1470 (2010).
40. Patra, A. et al. Kinetics, structure, and mechanism of 8-Oxo-7,8-dihydro-2'-deoxyguanosine bypass by human DNA polymerase  $\epsilon$ . *J. Biol. Chem.* **289**, 16867–16882 (2014).
41. Szymanski, M. R. et al. Structural basis for processivity and antiviral drug toxicity in human mitochondrial DNA replicase. *EMBO J.* **34**, 1959–1970 (2015).
42. Stojkovic, G. et al. Oxidative DNA damage stalls the human mitochondrial replisome. *Sci. Rep.* **6**, 28942 (2016).
43. McCauley, M. J. et al. Quantifying the stability of oxidatively damaged DNA by single-molecule DNA stretching. *Nucleic Acids Res.* **46**, 4033–4043 (2018).
44. Steitz, T. A. Structural biology—a mechanism for all polymerases. *Nature* **391**, 231–232 (1998).
45. Batra, V. K., Shock, D. D., Beard, W. A., McKenna, C. E. & Wilson, S. H. Binary complex crystal structure of DNA polymerase beta reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc. Natl Acad. Sci. USA* **109**, 113–118 (2012).
46. Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. J. & Yang, W. Watching DNA polymerase  $\epsilon$  make a phosphodiester bond. *Nature* **487**, 196–201 (2012).
47. Gao, Y. & Yang, W. Capture of a third Mg(2)(+) is essential for catalyzing DNA synthesis. *Science* **352**, 1334–1337 (2016).
48. Xu, W., Zhao, W., Morehouse, N., Tree, M. O. & Zhao, L. Divalent cations alter the rate-limiting step of PrimPol-catalyzed DNA elongation. *J. Mol. Biol.* **431**, 673–686 (2019).
49. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
50. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
51. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
52. DeLano, W. L. & Lam, J. W. PyMOL: a communications tool for computational models. *Abstr. Pap. Am. Chem. Soc.* **230**, U1371–U1372 (2005).

## Acknowledgements

We thank the staff at Northeastern Collaborative Access Team (NECAT) beam line at the Argonne National Laboratory for facilitating X-ray data collection. This work was partially supported by NIH grants R35-GM131780 (A. K. A.) and R01 GM126087 (L. P.). The NECAT beamline was funded by the National Institute of General Medical Sciences (NIGMS) from the NIH (P41 GM103403). The Pilatus 6 M detector on 24-ID-C beamline was funded by an NIH–Office of Research Infrastructure Programs High-End Instrumentation grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. Y. K. G. is supported by the Max and Minnie Tomerlin Voelcker Foundation and the Cancer Prevention Research Institute of Texas (RP190534).

## Author contributions

O.R. and A.K.A. designed the experiments. R.E.J. performed protein expression under the supervision of S.P. and L.P.; O.R. performed protein purification, crystallization, X-ray data collection, and refinement; Y.K.G. assisted with structure refinement of the oxoG•dCTP complex. O.R. and A.K.A. wrote the paper with contributions from L.P. and S.P. All authors reviewed the paper.

## Competing interests

Y.K.G. is the founder of Atomic Therapeutics. This affiliation does not affect the authors' impartiality, adherence to journal standards and policies, or availability of data. The other authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-24317-z>.

**Correspondence** and requests for materials should be addressed to O.R. or A.K.A.

**Peer review information** *Nature Communications* thanks Katarzyna Bebenek and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021