

ARTICLE

DOI: 10.1038/s41467-017-01013-5

OPEN

Mechanism of error-free replication across benzo [a]pyrene stereoisomers by Rev1 DNA polymerase

Olga Rechkoblit¹, Alexander Kolbanovskiy², Hannah Landes¹, Nicholas E. Geacintov² & Aneel K. Aggarwal¹

Benzo[a]pyrene (BP) is a carcinogen in cigarette smoke which, after metabolic activation, can react with the exocyclic N^2 amino group of guanine to generate four stereoisomeric BP- N^2 -dG adducts. Rev1 is unique among translesion synthesis DNA polymerases in employing a protein-template-directed mechanism of DNA synthesis opposite undamaged and damaged guanine. Here we report high-resolution structures of yeast Rev1 with three BP- N^2 -dG adducts, namely the 10S (+)-*trans*-BP- N^2 -dG, 10R (+)-*cis*-BP- N^2 -dG, and 10S (-)-*cis*-BP- N^2 -dG. Surprisingly, in all three structures, the bulky and hydrophobic BP pyrenyl residue is entirely solvent-exposed in the major groove of the DNA. This is very different from the adduct alignments hitherto observed in free or protein-bound DNA. All complexes are well poised for dCTP insertion. Our structures provide a view of *cis*-BP- N^2 -dG adducts in a DNA polymerase active site, and offer a basis for understanding error-free replication of the BP-derived stereoisomeric guanine adducts.

¹Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, Box 1677, 1425 Madison Avenue, New York, NY 10029, USA.

²Department of Chemistry, New York University, 31 Washington Place, New York, NY 10003-5180, USA. Correspondence and requests for materials should be addressed to O.R. (email: olga.rechkoblit@mssm.edu)

Tobacco smoking is directly associated with a majority of lung cancer cases in the United States and is among the few firmly established links between the etiology and the manifestation of this disease^{1–4}. Benzo[*a*]pyrene (BP) is one of the most potent and extensively studied carcinogens in cigarette smoke⁵. It is also ubiquitous in the human environment since BP is generated as a result of incomplete combustion of organic matter, such as fossil fuels and wood, and is present in automobile exhaust fumes and charcoal-grilled foods^{6–8}. Exposure of animals^{9,10} and human cells¹¹ to BP gives rise to G to T transversion mutations, which are also observed with high frequencies in smoking-related lung cancers^{2,3} and other malignancies⁴.

The carcinogenicity of BP derives from its metabolic activation by the cytochrome P450 pathway that generates reactive diol epoxide intermediates that react covalently with DNA¹². More specifically, the metabolic activation of BP in human cells gives rise to a pair of mirror image BP diol epoxides, the (+)-*anti*-BPDE and (–)-*anti*-BPDE enantiomers (Fig. 1a), each of which reacts predominantly with the exocyclic *N*² amino group of guanine by *trans*- and *cis*- epoxide ring opening. This leads to four possible stereoisomeric BP-*N*²-dG adducts (Fig. 1b). A minor fraction of adducts are formed by reactions of the diol epoxides with the *N*⁶-amino groups of adenine¹³. The highly mutagenic and carcinogenic 10*S* (+)-*trans*-BP-*N*²-dG adduct is the predominant stereoisomer found in human cells treated with BP (~90%)¹³, and is more resistant to removal by the nucleotide excision repair (NER) pathway than the (+)-*cis*-adducts and (–)-*cis*-adducts^{14,15}.

Unrepaired BP-*N*²-dG adducts severely impede high-fidelity DNA polymerases that replicate genomic DNA^{16,17}. Nevertheless, experiments in mammalian cells with site-specifically modified (+)-*trans*-BP-*N*²-dG oligonucleotides embedded in gapped plasmid vectors demonstrate that the adduct is bypassed with ~50% efficiency^{18,19}. Such replication obstacles can be handled by specialized, lower fidelity translesion DNA synthesis (TLS) polymerases in the S phase of the cell cycle to ensure continuous progression of the replisome²⁰. Alternatively, the

damaged DNA site can be skipped to leave a single-stranded DNA gap to be filled in later in the G2 phase of the cell cycle.

Humans possess four Y-family TLS polymerases (Pol η, Pol ι, Pol κ, and Rev1) while yeast has two (Pol η and Rev1)²¹. Pol κ and Rev1, can efficiently and preferentially insert the correct cytosine base opposite the BP-*N*²-dG adducts in vitro^{22–26}, while Pol η is proficient in incorporating mutation-inducing adenine base^{23,27}, and Pol ι is blocked by the adducts²³. In addition, the B-family TLS Pol ζ, a “universal extender”, is able to elongate from either the correct or incorrect bases inserted by other polymerases (including Rev1) across the lesion site^{28–30}.

In cellular studies, mouse cells deficient in Pol κ have reduced cell survival and accumulate more mutations after exposure to BPDE³¹. In addition, Pol κ was required for recovery from BPDE-induced S-phase checkpoints³². Rev1 was found to accumulate in nuclear foci upon exposure to BPDE in human cells³³, while Pol ζ's importance was documented in human Nalm-6-MSH+ cells³⁴. Quantitative assessments of the impact of TLS polymerases on the bypass of (+)-*trans*-BP-*N*²-dG-containing gapped plasmids were studied by siRNA-induced knockdown of both Pol κ and Pol ζ in human U2OS cells. These experiments revealed that 42% of the TLS bypass events were error-free²⁹, suggesting that another TLS polymerase with the ability to insert the correct cytosine base opposite the BP-*N*²-dG adducts was most likely involved.

From in vitro biochemical and structural studies Rev1 is a plausible candidate. The structures of yeast and human Rev1 catalytic cores revealed a novel protein-template-directed mechanism of DNA synthesis^{35,36}, wherein the template dG or an *N*²-dG adduct is evicted from the DNA helix and an arginine makes specific hydrogen bonds with the Watson–Crick (W–C) edge of the incoming dCTP. Rev1 is thus unique among TLS polymerases in its protein-template-directed mechanism of DNA synthesis and its specificity in incorporating C opposite template dG. The Rev1 catalytic activity was shown to be required in vivo for TLS opposite *N*²-dG-derived lesions such as those generated by 4-nitroquinoline-1-oxide (4-NQO)³⁷, as well as 1,*N*⁶-

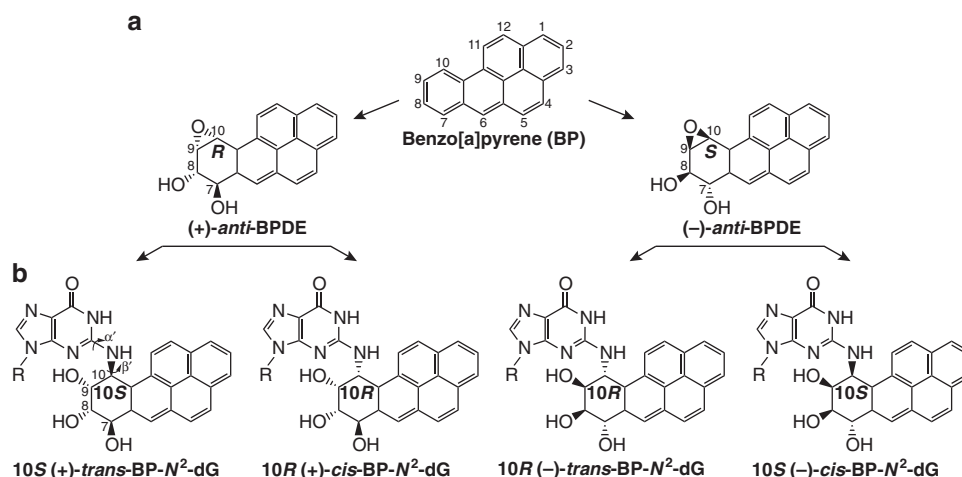


Fig. 1 Metabolic activation of benzo[*a*]pyrene (BP) and conformations of stereoisomeric BP-*N*²-dG adducts. **a** Metabolic activation of BP by the cytochrome P450 pathway to (+)-*anti*-BPDE ((+)-7*R*,8*S*-dihydroxy-9*S*,10*R*-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene), and its mirror image enantiomer (–)-*anti*-BPDE ((–)-7*S*,8*R*-dihydroxy-9*R*,10*S*-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene). **b** Stereoisomeric BP-*N*²-dG adducts. These are covalent products derived from the binding of (+)-*anti*-BPDE or (–)-*anti*-BPDE at its C10 position to the *N*²-2'-deoxyguanosine residues in DNA by *trans*- and *cis*- opening of the epoxide ring of BPDE. The (+) and (–) designations denote the mirror image relationships between the absolute configurations of the hydroxyl -OH groups at the C7, C8, and C9 chiral carbon centers. The absolute configuration of the adducts at the C10 carbon atom is 10*S* for the (+)-*trans*-adducts and (–)-*cis*-adducts, and 10*R* for the (–)-*trans*-lesions and (+)-*cis*-lesions. The linkage geometry between the BP and the *N*² amino group of guanine is defined by the torsion angles α' and β' as follows: α' , N1(dG)–C2(dG)–*N*²(dG)–C10(BP), and β' , C2(dG)–*N*²(dG)–C10(BP)–C9(BP). The relative percentage of each adduct formed in human cells is (+)-*trans*, 89%; (–)-*trans*, 7%; (+)-*cis*, 3%, and (–)-*cis*, 1%¹³

ethenoadducts³⁸ and abasic sites^{39–41}. In addition, the misincorporation frequencies of Rev1 opposite the 10S (+)-*trans*- and 10R (–)-*trans*-BP-dG adducts are low and similar to those observed with unmodified template dG ($\sim 10^{-2}$ to 10^{-4})²⁴.

Detailed NMR studies have revealed dramatically different alignments of the covalently attached carcinogen moieties in the different stereoisomeric BP-N²-dG adducts in free DNA duplexes in aqueous solution. In case of the *trans*-BP-N²-dG adducts, W–C pairing of the modified guanine with the partner cytosine base is maintained^{42, 43}, and the BP moiety is situated within the minor groove of the DNA duplex. In the case of the 10S (+)-*trans*-adduct⁴², the BP moiety points toward the 5′-end of the modified template strand⁴², while in the case of the 10R (–)-*trans*-adduct it points toward the 3′-end⁴³. In the case of the 10R (+)-*cis*-adducts and 10S (–)-*cis*-adducts, the guanine and partner cytosine bases are extruded from the helix, with the hydrophobic BP moiety intercalating within the duplex DNA in place of the displaced G:C base pair^{44, 45}. The BP moieties of the two *cis*-adducts are again oriented differently, toward the major and minor grooves of DNA, respectively. Altogether, the stereochemistry of each BPDE adduct determines how it is accommodated in a DNA polymerase active site.

Crystal structures of DNA polymerases with BP-N²-dG adducts are limited to the 10S (+)-*trans*-stereoisomeric adduct. For example, structures of both A-family Pol I from *Bacillus stearothermophilus* (BF Pol I)⁴⁶ and Y-family human Pol κ ^{47, 48} with 10S (+)-*trans*-BP-N²-dG show that the BP moiety is aligned in the minor groove of the bound DNA, in a manner similar to that anticipated from the NMR structure of the same (+)-*trans*-BP-N²-dG in free DNA⁴². By contrast, in structures of Y-family archeal Dpo4 with 10S (+)-*trans*-BP-N²-dG, the adduct is flipped out the DNA helix, resulting in a non-instructional gap in the template strand⁴⁹.

We present here high-resolution structures of yeast Rev1 with not only the 10S (+)-*trans*-BP-N²-dG, but also the 10R (+)-*cis*-BP-N²-dG and 10S (–)-*cis*-BP-N²-dG adducts as the template bases. Together, these structures provide detailed comparisons of the accommodations of the different BP-N²-dG stereoisomeric adducts within the confines of a DNA polymerase active site. Surprisingly, the bulky and hydrophobic BP pyrenyl moiety is entirely solvent-exposed in the major groove of the template–primer DNA helix. This structural alignment is fundamentally different from the ones observed in free DNA as well as previous polymerase–DNA complexes. Furthermore, the structures reveal an opposite orientation of the BP moiety in the 10S (+)-*trans*- and 10R (+)-*cis*-adducts, imposed by the opposite absolute configurations of the C10 carcinogen–DNA linkage site. We also find that the 10S (–)-*cis*- adduct does not mirror the alignment of its 10R (+)-*cis*-enantiomer, and thus breaking the “rule” of opposite orientations of the 10S and 10R stereoisomeric BP-N²-dG adducts. Notably, all three complexes are reaction-ready and well poised for dCTP insertion.

Results

Preparation of BP-N²-dG-modified DNA oligonucleotides.

Site- and stereo-specifically modified 17-mer DNA oligonucleotides with single (+)-*trans*-BP-N²-dG, (–)-*trans*-BP-N²-dG, (+)-*cis*-BP-N²-dG, or (–)-*cis*-BP-N²-dG adducts (Fig. 1b) were generated by the direct synthesis method⁵⁰ using racemic (±)-*anti*-BPDE. Briefly, a 17-mer DNA oligonucleotide with a single G base at the 5th position from the 5′-end (5′-CATCGC-TACCACACCCC-3′) was incubated with BPDE. The BP-DNA adducts were separated from the unreacted oligonucleotide and fully hydrolyzed to BP tetraols by HPLC methods (Supplementary Fig. 1a). The mixture of adducted oligonucleotides was then

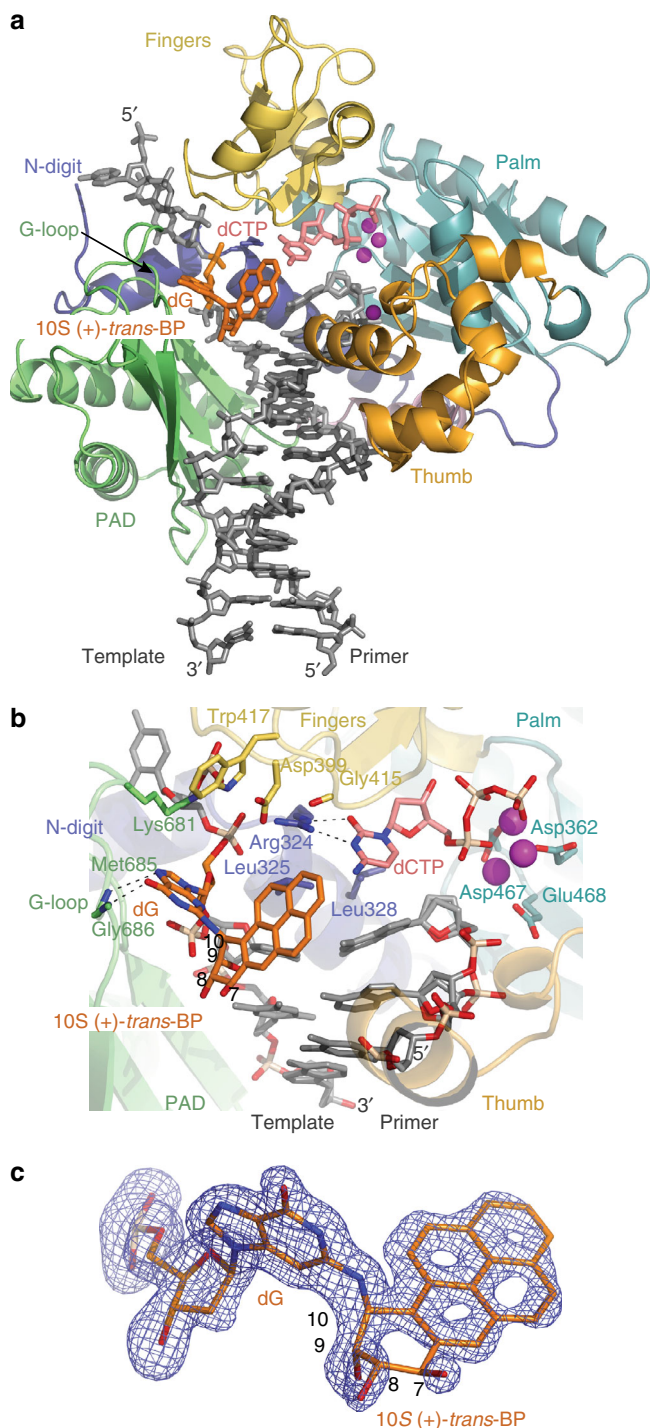
Table 1 X-ray data collection and refinement statistics

	10S (+)- <i>trans</i> -BP-dG	10R (+)- <i>cis</i> -BP-dG	10S (–)- <i>cis</i> -BP-dG
<i>Data collection</i>			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions:			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	64.0, 64.9, 130.7	62.5, 180.4, 54.6	63.3, 65.5, 131.7
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å) ^a	45–1.85 (1.88–1.85)	50–1.92 (1.95–1.92)	45–2.25 (2.29–2.25)
<i>R</i> _{merge} (%)	6.6 (45.2)	6.3 (65.0)	10.3 (72.9)
<i>I</i> / σ <i>I</i>	26.3 (2.2)	20.2 (1.4)	18.6 (2.2)
Completeness (%)	99.7 (95.5)	95.3 (92.9)	100 (100)
Redundancy	7.2 (4.2)	5.8 (4.4)	7.4 (6.9)
<i>Refinement</i>			
Resolution range (Å)	45–1.85	45–1.92	45–2.25
No. reflections	44,991	43,093	25,424
<i>R</i> _{work} / <i>R</i> _{free}	17.0/20.8	17.6/21.4	17.0/21.6
No. atoms			
Protein	3,490	3,481	3,488
DNA	610	615	527
Ligand (dCTP)	28	28	28
Ligand (other)	75	72	59
Ion (Mg ²⁺)	4	4	4
Water	340	311	161
B-factors			
Protein	24.0	31.5	37.4
DNA	32.0	38.3	51.6
Ligand (dCTP)	16.8	20.0	28.5
Ligand (other)	39.5	47.9	58.7
Ion (Mg ²⁺)	22.8	25.0	31.9
Water	35.4	40.9	41.4
R.m.s. deviations			
Bond length (Å)	0.009	0.010	0.011
Bond angles (°)	1.39	1.39	1.47

^aValues in parentheses are for highest-resolution shell

subjected to further separation of components by HPLC methods (Supplementary Fig. 1b and c). To characterize the stereochemistry of the covalently bound BP-N²-dG adducts in the different modified oligonucleotides, $\sim 34 \mu\text{g}$ of each purified sample was subjected to enzymatic hydrolysis to the BP-N²-dG nucleoside levels and the circular dichroism (CD) spectra of the hydrolyzed adducts (Supplementary Fig. 1d and e) was compared to the previously described standards⁵⁰.

Structure determination. We originally sought to co-crystallize the yeast Rev1 polymerase catalytic core with DNA templates containing each of the four stereoisomeric BP-N²-dG adducts. We succeeded in co-crystallizing the enzyme with three of the four stereoisomeric adducts, namely the 10S (+)-*trans*-BP-N²-dG, 10R (+)-*cis*-BP-N²-dG, and 10S (–)-*cis*-BP-N²-dG adducts (Fig. 1b; Supplementary Fig. 1). The 17-mer templates containing the adducts were paired with a 12-mer primer terminated with 2′,3′-dideoxyguanine (5′-GGGGTGTGGTAG^{dd}-3′), and with dCTP as the incoming nucleotide. The complex containing the 10R (–)-*trans*-BP-N²-dG adduct failed to produce crystals despite the fact that Rev1 has similar dCTP incorporation efficiency opposite both the 10S (+)-*trans*-BP-N²-dG and 10R (–)-*trans*-BP-N²-dG adducts²⁴. The structure of the 10R (+)-*cis*-BP-containing complex was solved by the molecular replacement method using the Rev1 complex with an unmodified DNA and dCTP (PDB ID: 2AQ4)³⁵ as a search model and refined at 1.92 Å resolution and *R*_{work}/*R*_{free} 17.6%/21.4%, respectively. The structure of the 10S (+)-*trans*-BP-dG-containing ternary complex was obtained by molecular replacement using the 10R (+)-*cis*-BP complex as a search model and refined to 1.85 Å resolution and *R*_{work}/*R*_{free} 17.0%/20.8%, respectively. The complex containing the 10S (–)-*cis*-BP-N²-dG was refined to 2.25 Å resolution and *R*_{work}/*R*_{free} 17.0%/21.6%. The crystal data, together with the data collection and refinement statistics, are summarized in Table 1.



Structure of the 10S (+)-*trans*-BP-dG ternary complex. Rev1 adopts a very similar conformation to that when it binds unmodified G (Fig. 2a; r.m.s. deviation of 0.76 Å for 433 C α s). That is, the enzyme embraces the template–primer DNA with its palm (residues 356–365, 438–536), fingers (residues 366–437), and thumb (537–603) domains as well as the PAD (polymerase associated domain; residues 621–738), which is unique to Y-family polymerases (Fig. 2a)³⁵. An α -helical substructure at the N terminus, an N-digit (residues 305–355) augments this embrace and makes critical interactions with the incoming dCTP and the templating base. The active site residues (Asp362, Asp467, and Glu468) are located on the palm domain to catalyze the

nucleotidyl transfer reaction (Fig. 2b). The fingers domain interacts with the incoming dCTP and the 5'-template overhang, while an extra long loop of the PAD, a G loop, interacts with the base of the modified guanine. The thumb and the PAD are positioned on opposite sides of the DNA duplex and are connected by a long helical linker (Fig. 2a). The incoming dCTP does not pair with templating 10S (+)-*trans*-BP-dG; instead, Arg324 (from the N-digit) acts as a “surrogate” residue and makes a set of complementary hydrogen bonds with the cytosine base of the incoming dCTP (Fig. 2b) as observed previously in the unmodified structure³⁵.

The modified dG is evicted from the DNA helix by Leu325 from the N-digit (Fig. 2b). The base rotates $\sim 90^\circ$ away from the minor groove so that the N² amino group now faces the major groove of the DNA duplex (instead of its usual placement in the minor groove). Consequently, the BP moiety is located in the major groove of the DNA. The stereochemistry of the BP moiety is well defined in the 1.85 Å electron density map and reveals an orientation that is almost perpendicular to the plane of the guanine base (torsion angle β' C2(dG)-N²(dG)-C10(BP)-C9(BP) is -75.9°) (Fig. 2c and Supplementary Fig. 2a). The low average B-factors for atoms of the BP and the modified dG moieties (32.6 and 20.8 Å², respectively) are close in value to the B-factors of the atoms of protein-bound portion of the DNA duplex (24.5 Å²) and indicate the well-ordered alignment of the adduct. Interestingly, the BP pyrenyl ring system points towards the 5'-direction of the modified strand (Fig. 2b and Supplementary Fig. 2a). This is made possible by the modified dG base tilting by $\sim 44^\circ$ in the 3'-direction to create room for the BP moiety to extend in the opposite direction, as shown in Fig. 3a, b. Furthermore, the BP moiety is rotated away from the W–C edge of dG, wherein the benzylic ring torsion angle α' (N1(dG)-C2(dG)-N²(dG)-C10(BP) is 163.8° (Figs. 2b, c). The hydroxyl groups at the C9 (C9-OH) and C8 (C8-OH) of the BP face away from the W–C edge of dG. The overall conformation of the BP-base linkage is in the most energetically favorable region ($\alpha' = 180 \pm 40^\circ$ and $\beta' = -90 \pm 40^\circ$) as calculated for an isolated (+)-*trans*-BP-dG nucleoside⁵¹.

Despite the tilt in the dG base to make room for the BP moiety, it maintains (as in the unmodified Rev1 complex structure)

Fig. 2 The 10S (+)-*trans*-BP-dG-modified Rev1 ternary complex. **a** Overall structure of the complex; the palm, fingers, thumb and PAD domains are shown in cartoon representation in cyan, yellow, light orange, and green, respectively. The linker joining the thumb and the PAD is shown in pink. The α -helical N-digit of Rev1 is shown in dark blue. The DNA template–primer duplex is shown in gray sticks with the template (+)-*trans*-BP-N²-dG in orange and incoming dCTP residue in red. The Mg²⁺ ions are represented as magenta spheres. **b** Alignment of the 10S (+)-*trans*-BP-N²-dG adduct in the Rev1 active site. The 10S (+)-*trans*-BP moiety is placed in the solvent-filled space between the PAD and the template–primer DNA helix. Asp362, Asp467, and Glu468 are the catalytic residues. Modified template dG and incoming dCTP partner with segments of Rev1 and not with each other. Arg324 makes hydrogen bonds with the base of dCTP, Leu325 pushes modified template dG out of the DNA helix and Leu328 is partially stacked on top of the 3'-terminal primer base. G-loop residues Met685, Gly686 and Lys681 as well as Trp417, Asp399, and Gly415 interact with the extrahelical template dG. The 3'-terminal and the adjacent residue of the primer strand have double conformations of their phosphate backbone reflecting the mobility of the 3' terminus in the Rev1 complex. **c** A simulated annealing Fo – Fc omit map (contoured at 3.0 σ -level at 1.85 Å resolution and colored in blue) showing the clear electron density for the BP moiety and the modified dG. The dG and the BP C7-OH hydroxyl group are positioned above the BP benzylic ring, while the C9-OH and C8-OH hydroxyl groups are below, thus defining the conformation and the expected stereochemistry of the 10S (+)-*trans* BP benzylic ring

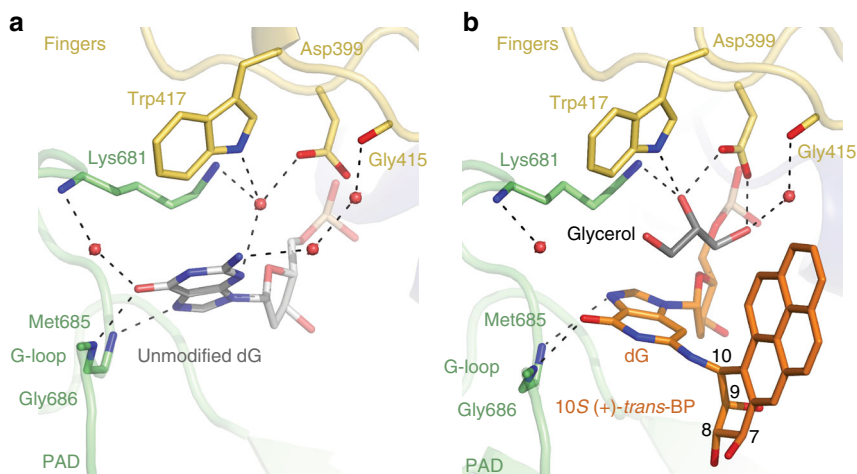


Fig. 3 A close up view of the unmodified dG and 10S (+)-*trans*-BP- N^2 -dG adduct contacts with Rev1. **a** The template dG in the unmodified complex (PDB ID: 2AQ4)³⁵. The backbone amines of Met685 and Gly686 form direct hydrogen bonds with the base of the dG. Side chains of residues Asp399, Trp417, Lys681 and Gly415 and the backbone amine of Lys681 form water-mediated contacts with the base. **b** The 10S (+)-*trans*-BP- N^2 -dG contacts with Rev1. The modified dG maintains hydrogen bonds between the N7 and O^6 atoms at its “Hoogsteen edge” and the main-chain amides of Met685 and Gly686 of the G loop. However, water-mediated contacts with Asp399, Trp417, Lys681, and Gly415 are lost. A molecule of glycerol (shown in gray) appears to take the place of some of the water molecules observed in the unmodified structure (**a**)

hydrogen bonds between the N7 and O^6 atoms at its “Hoogsteen edge” and the main-chain amides of Met685 and Gly686 of the G loop (Fig. 3a, b). However, a number of water-mediated interactions between the N3 atom and the side chains of Asp399, Trp417, and Lys681, a water bridge between the O^6 atom and the backbone NH of the Lys681, as well as a two-water molecule bridge between the N^2 group and the backbone carbonyl of Gly415. Interestingly, a molecule of glycerol (shown in gray on Fig. 3b) appears to take place of some of the water molecules observed in the unmodified structure (Fig. 3a).

Structure of the 10R (+)-*cis*-BP-dG ternary complex. In contrast to the 10S (+)-*trans*-BP-dG ternary structure, the BP moiety is directed towards the 3'-end of the template strand (Figs. 4a, b). This likely reflects the fact that the absolute configuration of the C10 DNA linkage of the 10R (+)-*cis*-BP-dG adduct is a mirror of the 10S (+)-*trans*-BP-dG adduct (Fig. 1b). The conformations of the 10R (+)-*cis*-BP and modified dG moieties are well ordered as indicated by the 1.92 Å electron density map (Fig. 4c and Supplementary Fig. 2b) and low B-factors (39.9 and 24.9 Å², respectively) that are close in value observed for the Rev1-bound portion of the DNA (33.0 Å²). The torsion angle β' (C2(dG)- N^2 (dG)-C10(BP)-C9(BP)) is 111.8° (Fig. 4c), as compared to -75.9° in the 10S (+)-*trans*-BP-dG ternary structure (Fig. 2c). Consequently, the benzylic ring of the 10R (+)-*cis*-BP residue is now above the guanine base in close proximity to the fingers domain of Rev1 (Fig. 4b). Moreover, the base of the modified dG remains untilted, and many of the direct and water-mediated hydrogen bonds observed in the structure with unmodified dG are present (Figs. 3a and 5a). This includes hydrogen bonds between the N7 and O^6 atoms at the Hoogsteen edge of the base and the main-chain amides of Met685 and Gly686 (Fig. 5a), and water-mediated bonds between the O^6 of the base and Lys681. Notably, the C9-OH displaces the highly coordinated water molecule and forms direct hydrogen bonds with the Rev1 side chains of Asp399 and Trp417 (Fig. 5a). There is also a water-mediated intermolecular bridge between the N1 of dG and the C8-OH hydroxyl of the BP moiety.

Although, the pyrenyl ring system points in opposite directions in the 10S (+)-*trans*- and 10R (+)-*cis*-BP-dG adducts, they occupy

relatively the same physical space in the Rev1 active site, between the PAD and the fingers domain (Fig. 5b). Also, in both cases the pyrenyl ring system is directed away from the W-C edge of the dG (torsion angle $\alpha' = 175.3^\circ$). However, the C9-OH and C8-OH hydroxyl groups are oriented toward the W-C edge of the dG residue due to the intrinsic stereochemistry of the *cis*-adducts that places the C9 and C8 hydroxyls on the same side of the benzylic ring as the guanine base (Fig. 4c). This conformation of the 10R (+)-*cis*-adduct is within the most energetically favorable region calculated for this nucleoside ($\alpha' = 185 \pm 35^\circ$ and $\beta' = 100 \pm 30^\circ$)⁵¹. Otherwise, the overall structures of the 10R (+)-*cis*-BP-dG-modified and 10S (+)-*trans*-BP-dG-modified complexes are similar (r.m.s. deviation = 0.97 Å for 433C α s) (Figs. 2a and 4a).

Structure of the 10S (-)-*cis*-BP-dG ternary complex. In contrast to the 10S (+)-*trans*-BP-dG and 10R (+)-*cis*-BP-dG structures described above, the electron density for the BP moiety is not well defined in the 10S (-)-*cis*-BP-dG structure (Fig. 5c). However, the electron density for the dG base is well defined and shows it in an untilted orientation, similar to the one observed in the 10R (+)-*cis*-BP complex. Consequently, there is no room for the bulky BP pyrenyl ring system to reside above the guanine base to mirror the 10R (+)-*cis*-BP adduct alignment. Thus, the 10S (-)-*cis*-BP pyrenyl ring system most likely points toward the 3'-direction of the modified strand, where there is enough space for an alignment similar to the one acquired by the 10R (+)-*cis*-BP moiety (Fig. 5a). This assumption is consistent with the residual electron density observed below the guanine base. Such an orientation would place the 10S (-)-*cis*-BP-dG into a less energetically favorable conformational region ($\alpha' = -10 \pm 40^\circ$ and $\beta' = -100 \pm 30^\circ$)⁵¹ with the dG N1 edge and benzylic BP edge closer to each other. The higher energetic cost of such orientation stems from the stereo crowding between the N1 edge of guanine and the bay region of the aromatic BP pyrenyl moiety. While computationally feasible, BP-dG conformations in this region of the potential energy surface have hitherto not been detected experimentally. Thus, this structure provides experimental evidence that the BP moieties of the 10R (+)-*cis*-BP-dG and 10S (-)-*cis*-BP-dG adduct pair can be oriented the same way and not in an approximate mirror arrangement. The observed disorder in the electron density map in the case of the 10S (-)-*cis*-BP moiety is, probably, due

to the BP residue sampling the large number of possible conformations in a search for a more favorable alignment within the Rev1 complex crystal.

Metal ions in the active site. In all three structures, there are three Mg^{2+} ions (Mg^{2+}_A , Mg^{2+}_B , and Mg^{2+}_C) in the active site (Fig. 6). Mg^{2+}_A , Mg^{2+}_B are analogous to the two metal ions present in high fidelity and TLS polymerases⁵². Mg^{2+}_C has an octahedral geometry with short ligation distances and is coordinated by the two non-bridging oxygen atoms of the α - and γ -phosphates of dCTP, a carboxylate oxygen of Asp362 and three waters. A similar Mg^{2+}_C ion has been observed in the active site of high-fidelity Pol δ ⁵³, and may facilitate dCTP binding to the active site and assist in the leaving of the pyrophosphate. This Mg^{2+}_C ion differs from the transient product-associated ion that binds between the α - and β -phosphates of dNTP during phosphodiester bond formation in Pol η ⁵⁴ and Pol β ⁵⁵ complexes. A fourth Mg^{2+} ion (Fig. 2a) is observed bridging the phosphate group of the primer base next to the primer terminus and the loop of the thumb domain. An analogous entity has been previously assigned as a water molecule in the yeast Rev1 structures^{35, 56} and as a Mg^{2+} ion in the human Rev1 complex³⁶.

Discussion

We show here that Rev1 is remarkably well-adapted for the error-free bypass of 10S (+)-*trans*-BP- N^2 -dG, 10R (+)-*cis*-BP- N^2 -dG, and 10S (-)-*cis*-BP- N^2 -dG stereoisomeric adducts. The ability of Rev1 to evict these bulky adducts from the DNA helix and to position Arg324 as the surrogate protein template residue allows for the incorporation of a correct C opposite these stereoisomers, consistent with the previous *in vitro* studies²⁴. Strikingly, the BP moiety in each case is pushed into the more capacious major groove side of the DNA, where it is solvent-exposed. This positioning of the BP moiety in the DNA major groove differs fundamentally from the configuration in free DNA. From solution NMR studies of BP lesions within B-DNA⁵⁷, the 10S (+)-*trans*-adducts and 10R (-)-*trans*-adducts reside in the minor groove, while the 10R (+)-*cis*-adducts and 10S (-)-*cis*-adducts adopt an intercalative conformation where the pyrenyl ring system displaces the dG:dC base pair and assumes an intercalated conformation^{44, 45}. Moreover, in free DNA, one face of the hydrophobic pyrenyl ring system is shielded from solvent by van der Waals interactions with the DNA backbone (10S (+)-*trans*-adducts and 10R (-)-*trans*-adducts), while both faces are shielded by stacking between adjacent bases within the DNA helix ((+)-*cis*-adducts and 10S (-)-*cis*-adducts). By contrast, in the Rev1 active site, both faces of the hydrophobic pyrenyl ring system are solvent-exposed (Fig. 2b). The Rev1 PAD and fingers domain, as well as the DNA sugar-phosphate backbone, are too far away to interact directly with the pyrenyl ring system in all three structures.

Although the 10S (+)-*trans*-BP- N^2 -dG adduct differs in how it is accommodated in the Rev1 active site versus free DNA, the torsion angles α' and β' defining the local conformation of the carcinogen-DNA linkage site (Fig. 1b) are similar in both cases and are in the most favorable range for a 10S (+)-*trans*-BP-dG-modified nucleoside, namely $\alpha' = 180 \pm 40^\circ$, $\beta' = -90 \pm 40^\circ$ ⁵¹. This similarity in α' and β' torsion angles partly underlies the fact that the BP pyrenyl ring system points towards the 5' end of the modified strand in both the Rev1 active site and in free DNA (Fig. 2b). The 10S (+)-*trans*-BP- N^2 -dG adduct in this favorable energy region has also been captured in the active sites of several other DNA polymerases, including BF Pol I⁴⁶, archeal Dpo4⁴⁹, and human Pol κ ^{47, 48}. However, the overall adduct alignments in

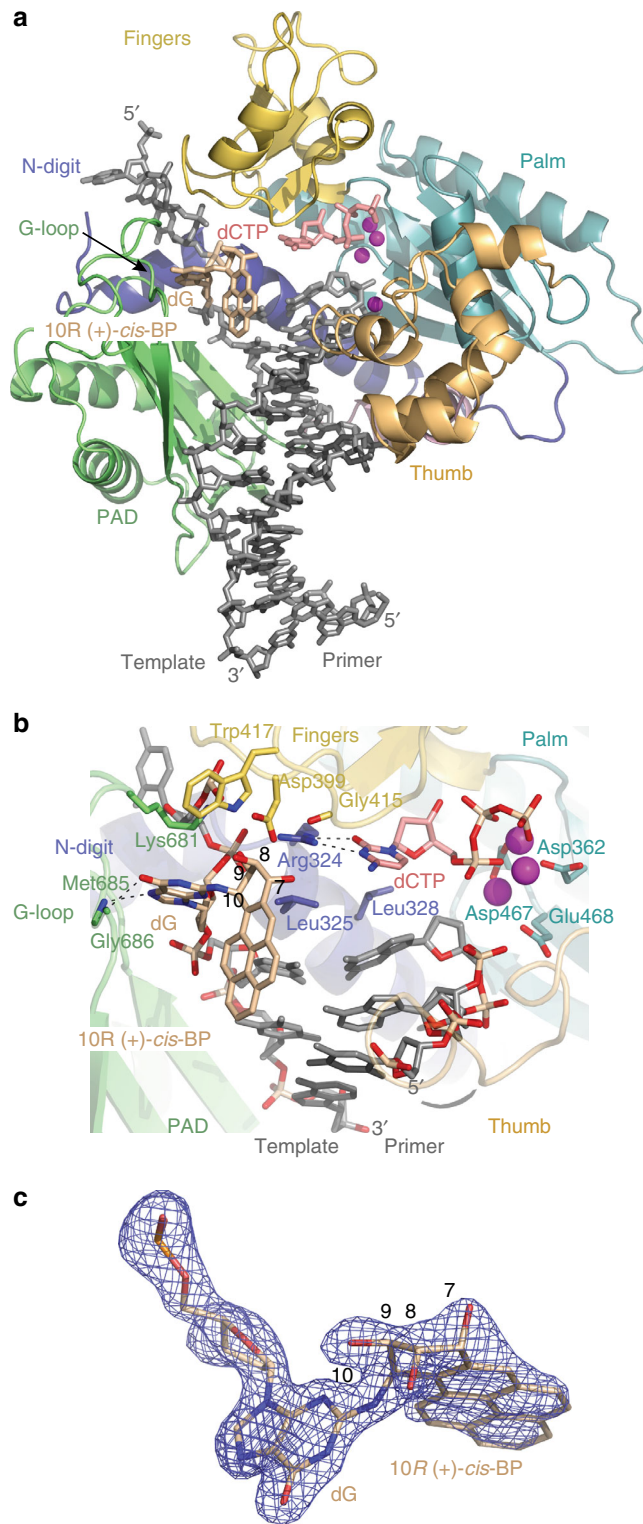


Fig. 4 The 10R (+)-*cis*-BP- N^2 -dG-modified Rev1 ternary complex. **a** Overall structure of the complex. The coloring scheme and the details shown are as in Fig. 2. The template 10R (+)-*cis*-BP- N^2 -dG is in beige. **b** Alignment of the 10R (+)-*cis*-BP- N^2 -dG adduct in the Rev1 active site. The 10S (+)-*trans*-BP moiety is positioned in the solvent-filled space between the PAD and the template-primer DNA helix. **c** A simulated annealing Fo-Fc omit map (contoured at 3.5σ at 1.92 Å resolution and colored in blue) showing clear density for the BP moiety and the modified dG. The dG as well as the C9-OH and C8-OH hydroxyl groups are oriented below the BP benzylic ring, while the C7-OH is above the ring. This confirms the expected stereochemistry of the 10R (+)-*cis* BP benzylic ring

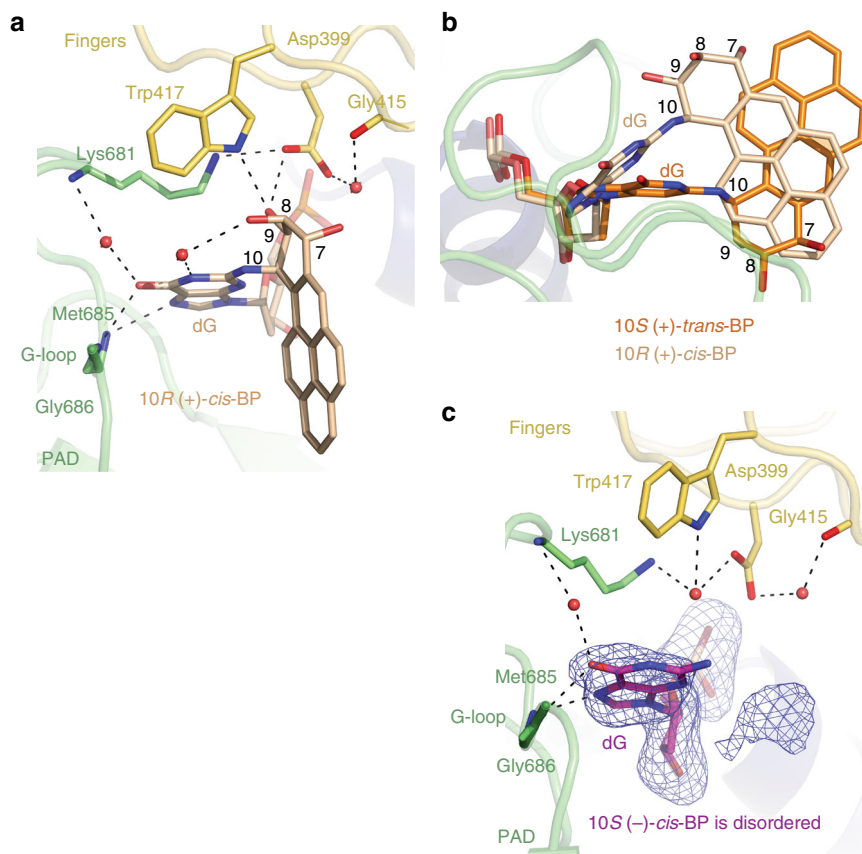


Fig. 5 A close up view of the 10R (+)- and 10S (-)-*cis*-BP- N^2 -dG adducts within Rev1 and comparisons with the 10S (+)-*trans*-BP structure. **a** Contacts of the 10S (-)-*cis*-BP- N^2 -dG adduct with the Rev1 protein. The BP-modified dG forms hydrogen bonds between the N7 and O^6 atoms at its Hoogsteen edge and the main-chain amides of Met685 and Gly686 of Rev1. There is also a water-mediated bond between the O^6 of the base and the peptide amide of Lys681. The C9-OH of the BP moiety forms hydrogen bonds with the side chains of Trp417 and Asp399 and the C8-OH group of the BP is bridged to the N1 of the dG base through a water molecule. **b** Comparison of the 10R (+)-*cis*-BP- N^2 -dG and 10S (+)-*trans*-BP adduct alignments. The 10R (+)-*cis*-BP-dG (shown in beige colored sticks) in the complex with Rev1 is superpositioned on the 10S (+)-*trans*-BP adduct structure (shown in orange sticks); the structures are aligned by Rev1 protein chains. The 10S (+)-*trans*-BP and 10R (+)-*cis*-BP- moieties are oriented in the opposite directions. However, they occupy approximately the same physical space in the Rev1 active site. **c** A simulated annealing Fo-Fc omit map (contoured at 3.0σ -level at 2.24 Å resolution and colored in blue) showing clear electron density of the (-)-*cis*-BP-modified dG base (magenta sticks). Slight change in the position of the modified dG base (as compared to the unmodified dG as shown in Fig. 3a) places the N3 atom 3.6 Å away from the water molecule that is firmly coordinated via the side chains of Asp399, Trp417, and Lys681 -too far for a hydrogen bond formation. Despite the reasonably high 2.24 Å resolution, the electron density map for the BP moiety is not well defined and the conformation of the adduct cannot be established unambiguously. However, there is a residual electron density of the BP moiety below the dG base

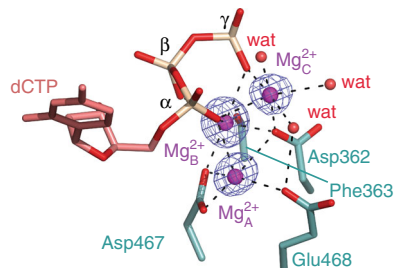


Fig. 6 The third Mg^{2+} ion (Mg^{2+}_C) at the Rev1 active site. The first Mg^{2+} , cation A, is coordinated by the invariant Asp362, Asp467, and Glu468 residues. The second Mg^{2+}_B , is chelated by the phosphate groups of the incoming dCTP and the main-chain carbonyl of Phe363. The third, Mg^{2+}_C , has six ligands and is coordinated by the two non-bridging oxygen atoms of the α - and γ -phosphates of dCTP, a carboxylate oxygen of Asp362 and three waters. The simulated annealing Fo-Fc map with Mg^{2+}_C omitted contoured at 7.0σ -level at 1.85 Å resolution is shown in blue. Hydrogen bonds and metal-ion coordination are shown as dashed lines

these complexes are strikingly distinct from the one in the Rev1 complex. In BF Pol I and Pol κ , the BP moiety lies in the minor groove and is directed toward the 5' of the template strand, adopting the conformations similar to the ones observed in the free DNA in solution^{44, 57}. In these structures, the BP moiety is protected from the solvent with one face of the pyrenyl ring system packed against the protein and the other against the DNA backbone. Whereas, in Dpo4, the BP-dG adduct is flipped out the DNA helix toward the minor groove and inserted into a solvent-protecting cleft between the fingers and the PAD domains, resulting in a non-instructional gap in the template strand. Interestingly, the structure of a base excision repair polymerase β in complex with the N^2 -dG adduct of the benzo[*c*]phenanthrene (BPh) diol epoxide also shows the adduct in a similar conformation as in free DNA⁵⁸.

The 10R (+)-*cis*-BP-dG and 10S (-)-*cis*-BP-dG ternary complex structures presented here show the *cis*-adducts in the confines of a DNA polymerase active site. The intercalative conformation of these adducts in free DNA has been ascribed to the C9-OH and C8-OH hydroxyl groups pointing inward toward

the DNA backbone, which leads to an already crowded situation in the minor groove of DNA⁵¹. We show here that the *cis*-adducts can occupy the major groove when bound to Rev1 (Fig. 5a, c) so that the C9-OH and C8-OH hydroxyl groups are far-removed from the neighboring bases. The torsion angles α' and β' for the 10R (+)-*cis*-BP-dG carcinogen-base linkage in the Rev1 active site are 175° and 112°, respectively, as compared to 160° and 136° in the free DNA⁴⁴. Thus, despite very different conformations, the torsion angles are in a similar range, as well as within the most stable potential energy surfaces of the 10R (+)-*cis*-BP-modified nucleoside⁵¹. This supports the idea that the intercalative conformation in free DNA is dictated less by the BP-dG linkage site α' and β' torsion angles and more by changes in DNA backbone torsion angles in the nucleotides adjacent to the adduct residues. Curiously, the BP moiety appears to adopt multiple conformations in our 10S (-)-*cis*-BP-dG ternary complex structure. These multiple conformations may stem from the fact that they appear to be centered on less favorable portion of the potential energy surface, namely $\alpha' = -10 \pm 40^\circ$ ⁵¹. Indeed, this allows for the BP pyrenyl aromatic ring system of the 10S (-)-*cis* adduct to point away from the finger domain of Rev1 and toward the 3'-end of the template as observed in the case of the 10R (+)-*cis*-BP stereoisomeric adduct. By contrast, when the 10S (-)-*cis*-BP-dG adduct structure is modeled with the more favorable torsion angles of $185 \pm 35^\circ$ and $-100 \pm 30^\circ$, the pyrenyl rings point into opposite directions resulting in severe steric clashes with the protein backbone of Ser398 and Asp399 and side chains of Asp399 and Trp417.

Despite strong biochemical and structural data demonstrating the ability of Rev1 to accommodate and bypass BP-dG adducts in vitro, the role of Rev1 in BP-dG bypass in vivo remains to be fully realized. In addition to its catalytic function, Rev1 also has a scaffolding function whereby its C-terminus can coordinate TLS via interactions with other TLS polymerases^{59–62}. In budding yeast, the Rev1 catalytic activity has been shown to be required in vivo for TLS across *N*²-dG-derived lesions such as 4-nitroquinoline-1-oxide (4-NQO)³⁷, as well as 1,*N*⁶-ethenoadenine adducts³⁸, but in human cells the scaffolding function of Rev1 appears to play an important role^{63, 64}. The structures of the catalytic cores of human and yeast Rev1 are very similar³⁶. The main difference is the presence of a hydrophobic “flap” on the pocket that holds the template G base, which if anything, appears to facilitate BP-*N*²-dG lesion bypass in vitro²⁴ (Supplementary Fig. 3). Taken together, the structures we present here provide a basis for further studies to dissect the relative contributions of the Rev1 catalytic and scaffolding functions in BP-*N*²-dG lesion bypass.

Methods

Preparation of the BP-*N*²-dG-modified 17-mer DNA templates. Site- and stereo-specifically modified 17-mer DNA oligonucleotides with single (+)-*trans*-anti-BP-*N*²-dG, (-)-*trans*-anti-BP-*N*²-dG, (+)-*cis*-anti-BP-*N*²-dG, or (-)-*cis*-anti-BP-*N*²-dG lesions were generated by a direct synthesis method⁶⁵ using racemic (\pm)-anti-BPDE obtained from the National Cancer Institute Carcinogen Reference Standard Repository (currently available from MRIGlobal Chemical Carcinogen Repository). The procedures used for the synthesis and adduct purification and the methods of verification of adduct stereochemistry were similar to those described previously for an 11-mer sequence⁵⁰. Briefly, ~83.8 mg (~2500 OD₂₆₀ units) of the 17-mer oligonucleotide 5'-CATCGCTACCACACCCC-3' (Integrated DNA Technologies) were dissolved in 10 ml of triethylamine acetate (TEAA)-sodium acetate pH 11 buffer. Racemic BPDE dissolved in 2 ml of tetrahydrofuran was added to a DNA solution in the molar ratio of BPDE to DNA 2:1. The mixture was allowed to react in darkness at +4 °C temperature with gentle mixing for one week to ensure complete reaction. The products in the reaction mixture were separated by HPLC using a PRP-1 polymeric HPLC preparative column (Hamilton). The HPLC conditions employed were a linear 10–30% acetonitrile/50 mM TEAA pH 7.0 buffer gradient in 60 min. A UV detector (wavelength set at 254 nm) and a fluorescence detector (emission wavelength = 400 nm, excitation wavelength = 343 nm) were used to monitor the elution profiles. A typical elution profile is

shown in Supplementary Fig. 1a. The unmodified oligonucleotide elutes first (at ~16–24 min), followed by the crude mixture of BP-DNA adducts dominated by BP-*N*²-dG adducts (collected at ~26–35 min) and by a mixture of minor BP-*N*⁶-dA adducts (at ~35–45 min), which are characterized by higher fluorescence intensities than the BP-*N*²-dG adducts. Fully hydrolyzed BP tetrol (BPT) was washed from the column with 80% acetonitrile (at ~50 min). After collection, the BP-*N*²-dG adduct mixture was vacuum dried, re-dissolved in water and further purified by HPLC in several injection steps with 10–20% acetonitrile/50 mM TEAA pH 7.0 buffer gradient in 60 min on C18 (ACE) column Supplementary Fig. 1b. Each elution peak was collected separately, combined with the corresponding peak from the other injection steps and vacuum dried. The purity check of Peak1 is shown in Supplementary Fig. 1c. Furthermore, the BP-*N*²-dG adducts were desalted with SlideSlide-A-Lyzer Dialysis Cassettes (ThermoFisher). To characterize the stereochemistry of the BPDE linkage, ~34 μ g of each purified adducted 17-mer was subjected to enzymatic hydrolysis to the BP-*N*²-dG-nucleoside levels. The stereochemistry was assigned based on the circular dichroism (CD) spectra of the hydrolyzed adducts (Supplementary Fig. 1d, e) and the previously described standards⁵⁰. The BP-*N*²-dG 17-mer yields were ~1.0 mg in the case of the (+)-*trans*-BP-modified, ~0.23 mg of the (-)-*trans*-BP-modified, ~0.84 mg of the (+)-*cis*-BP-modified, and ~0.77 mg of the (-)-*cis*-BP-modified oligonucleotides.

Expression and purification of yeast Rev1 protein. The *Saccharomyces cerevisiae* catalytic core Rev1 protein (residues 297–746) was expressed in *Escherichia coli* BL21 (DE3) codon Plus RIL (Stratagene) cells as an N-terminally tagged glutathione S-transferase (GST) fusion protein. The cells were grown in Luria-Bertani (LB) medium at 37 °C and expression of the fusion protein induced by the addition of 0.1 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG) followed by overnight incubation at 18 °C. The GST-Rev1 fusion protein was purified from bacterial lysate by affinity chromatography using a glutathione-Sepharose column. The GST tag was next cleaved on the resin with PreScission protease and Rev1 (297–746) was eluted from the column as described previously³⁵. The protein was further purified by chromatography on HiTrap Heparin column, following by passage through a Superdex 75 column (GE Healthcare). The protein was concentrated to ~12 mg ml⁻¹ in 25 mM Tris (pH 8.0), 250 mM NaCl, and 2 mM Tris(2-carboxyethyl) phosphate (TCEP) and stored in aliquots at -80 °C.

Crystallization. The crystals of the yeast Rev1 ternary complexes containing 10S (+)-*trans*-BP-*N*²-dG, 10R (+)-*cis*-BP-*N*²-dG, or 10R (-)-*trans*-BP-*N*²-dG 17-mer templates and 12-mer complementary primer terminated with 2',3'-dideoxyguanine (5'-GGGGTGTGGTAG-3') in the presence of dCTP were obtained by a hanging drop method against a reservoir solution containing 0.25 M sodium citrate pH 6.0 buffer and 15–20% PEG3350. Briefly, the template-primer DNAs were annealed and mixed with Rev1 protein in a 1.2:1 molar ratio to a final complex concentration of 0.11 mM in 22 mM Tris (pH 8.0), 160 mM NaCl, 1.5 mM TCEP, 10 mM MgCl₂, and 10 mM dCTP. The complexes were incubated at room temperature for 10 min and then centrifuged at 10,000 r.p.m. for 7 min at 4 °C. Crystallization drop was formed by mixing 1 μ L of the complex with 1 μ L of the reservoir solution and the crystals were grown at 20 °C. The complex containing 10R (-)-*trans*-BP-*N*²-dG adducted oligonucleotide failed to produce crystals. The crystals were cryoprotected in the reservoir solution supplemented with 24% PEG3350 and 20% glycerol and flash frozen in liquid nitrogen for X-ray data collection. Several rounds of microseeding were necessary to produce the large diffraction-quality crystals.

Structure determination and refinement. The X-ray diffraction data were collected at the NSLSX25 beam line at the Brookhaven National Laboratory. The data were processed and scaled using the HKL2000 suite⁶⁶. The structure of the 10R (+)-*cis*-BP-dG complex was solved by the molecular replacement method (Phaser)⁶⁷ in the CCP4 program package⁶⁸, using a previously published structure with an unmodified dG and incoming dCTP, PDB ID 2A44³⁵ (with a different DNA sequence) as search model. The model building, including substitution of the DNA sequence, was finished manually in Coot⁶⁹ based on the electron density maps calculated in REFMAC⁷⁰ in the CCP4 suite. The resulting model was refined in REFMAC in space group P 2₁2₁2 with $a = 63 \text{ \AA}$, $b = 180 \text{ \AA}$, $c = 55 \text{ \AA}$ unit cell to 1.92 \AA and $R_{\text{work/free}} = 18.3\%/21.8\%$, correspondingly. The placement of the BP and modified dG moieties was verified using the simulated annealing omit maps calculated in Phenix⁷¹ with the (+)-*cis*-BP-*N*²-dG omitted from the models before heating them to 2000 K and then slowly cooling them. The refined 10R (+)-*cis*-BP-dG structure includes Rev1 residues 307–738, nucleotides 2–17 for the template strand, nucleotides 1–12 for the primer strand, incoming dCTP, 4 Mg²⁺ ions, and 311 water molecules. The 3'-terminal and the adjacent residue of the primer strand have double conformations of their phosphate backbone reflecting the mobility of the 3' terminus in Rev1 complex. The structure of 10S (+)-*trans*-BP-dG ternary complex was obtained by molecular replacement using the 10R (+)-*cis*-BP complex as a search model. The placement of the 10S (+)-*trans*-BP-*N*²-dG adduct was verified using the simulated annealing omit maps as described above. The model was refined to 1.85 \AA in space group P 2₁2₁2; $a = 64 \text{ \AA}$, $b = 65 \text{ \AA}$, $c = 131 \text{ \AA}$ and $R_{\text{work/free}} = 17.7\%/21.7\%$, respectively.

Data availability. Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5WWM1, 5WWM8 and 5WMB for the 10S (+)-*trans*-, 10R (+)-*cis*- and 10S (-)-*cis*-BP-dG-containing Rev1 ternary complexes, respectively. Other data are available from the corresponding author upon reasonable request.

Received: 30 May 2017 Accepted: 11 August 2017

Published online: 17 October 2017

References

- Luch, A. Nature and nurture-lessons from chemical carcinogenesis. *Nat. Rev. Cancer* **5**, 113–125 (2005).
- Govindan, R. et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012).
- Imielinski, M. et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).
- Alexandrov, L. B. et al. Mutational signatures associated with tobacco smoking in human cancer. *Science* **354**, 618–622 (2016).
- Pfeifer, G. P. et al. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21**, 7435–7451 (2002).
- Kim, K. H., Jahan, S. A., Kabir, E. & Brown, R. J. A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. *Environ. Int.* **60**, 71–80 (2013).
- Poirier, M. C. Chemical-induced DNA damage and human cancer risk. *Nat. Rev. Cancer* **4**, 630–637 (2004).
- Nebert, D. W., Shi, Z., Galvez-Peralta, M., Uno, S. & Dragin, N. Oral benzo[a]pyrene: understanding pharmacokinetics, detoxication, and consequences—Cyp1 knockout mouse lines as a paradigm. *Mol. Pharmacol.* **84**, 304–313 (2013).
- Ruggeri, B. et al. Benzo[a]pyrene-induced murine skin tumors exhibit frequent and characteristic G to T mutations in the p53 gene. *Proc. Natl Acad. Sci. USA* **90**, 1013–1017 (1993).
- Monroe, J. J., Kort, K. L., Miller, J. E., Marino, D. R. & Skopek, T. R. A comparative study of in vivo mutation assays: analysis of hprt, lacI, cII/cI and as mutational targets for N-nitroso-N-methylurea and benzo[a]pyrene in Big Blue mice. *Mutat. Res.* **421**, 121–136 (1998).
- Nik-Zainal, S. et al. The genome as a record of environmental exposure. *Mutagenesis* **30**, 763–770 (2015).
- Jiang, H., Shen, Y. M., Quinn, A. M. & Penning, T. M. Competing roles of cytochrome P450 1A1/1B1 and aldo-keto reductase 1A1 in the metabolic activation of (+/-)-7,8-dihydroxy-7,8-dihydro-benzo[a]pyrene in human bronchoalveolar cell extracts. *Chem. Res. Toxicol.* **18**, 365–374 (2005).
- Ruan, Q. et al. Quantification of benzo[a]pyrene diol epoxide DNA-adducts by stable isotope dilution liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **20**, 1369–1380 (2006).
- Mocquet, V. et al. The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *EMBO J.* **26**, 2923–2932 (2007).
- Lee, Y. C. et al. The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: is there a correlation? *DNA Repair* **19**, 55–63 (2014).
- Chary, P. & Lloyd, R. S. In vitro replication by prokaryotic and eukaryotic polymerases on DNA templates containing site-specific and stereospecific benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide adducts. *Nucleic Acids Res.* **23**, 1398–1405 (1995).
- Zhang, Y., Wu, X., Guo, D., Rechkoblit, O. & Wang, Z. Activities of human DNA polymerase kappa in response to the major benzo[a]pyrene DNA adduct: error-free lesion bypass and extension synthesis from opposite the lesion. *DNA Repair* **1**, 559–569 (2002).
- Moriya, M. et al. Fidelity of translesional synthesis past benzo[a]pyrene diol epoxide-2'-deoxyguanosine DNA adducts: marked effects of host cell, sequence context, and chirality. *Biochemistry* **35**, 16646–16651 (1996).
- Avkin, S. et al. Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa. *J. Biol. Chem.* **279**, 53298–53305 (2004).
- Sale, J. E., Lehmann, A. R. & Woodgate, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat. Rev. Mol. Cell. Biol.* **13**, 141–152 (2012).
- Prakash, S., Johnson, R. E. & Prakash, L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem.* **74**, 317–353 (2005).
- Zhang, Y. et al. Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro. *Nucleic Acids Res.* **28**, 4138–4146 (2000).
- Rechkoblit, O. et al. trans-Lesion synthesis past bulky benzo[a]pyrene diol epoxide N2-dG and N6-dA lesions catalyzed by DNA bypass polymerases. *J. Biol. Chem.* **277**, 30488–30494 (2002).
- Zhang, Y. et al. Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. *Nucleic Acids Res.* **30**, 1630–1638 (2002).
- Suzuki, N. et al. Translesion synthesis by human DNA polymerase kappa on a DNA template containing a single stereoisomer of dG-(+)- or dG-(-)-anti-N-2-BPDE (7,8-dihydroxy-anti-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene). *Biochemistry* **41**, 6100–6106 (2002).
- Huang, X. et al. Effects of base sequence context on translesion synthesis past a bulky (+)-*trans*-anti-B[a]P-N2-dG lesion catalyzed by the Y-family polymerase pol kappa. *Biochemistry* **42**, 2456–2466 (2003).
- Zhang, Y. et al. Error-prone lesion bypass by human DNA polymerase eta. *Nucleic Acids Res.* **28**, 4717–4724 (2000).
- Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S. & Prakash, L. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* **406**, 1015–1019 (2000).
- Shachar, S. et al. Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J.* **28**, 383–393 (2009).
- Acharya, N., Johnson, R. E., Prakash, S. & Prakash, L. Complex formation with Rev1 enhances the proficiency of Saccharomyces cerevisiae DNA polymerase zeta for mismatch extension and for extension opposite from DNA lesions. *Mol. Cell. Biol.* **26**, 9555–9563 (2006).
- Ogi, T., Shinkai, Y., Tanaka, K. & Ohmori, H. Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. *Proc. Natl Acad. Sci. USA* **99**, 15548–15553 (2002).
- Bi, X., Slater, D. M., Ohmori, H. & Vaziri, C. DNA polymerase kappa is specifically required for recovery from the benzo[a]pyrene-dihydrodiol epoxide (BPDE)-induced S-phase checkpoint. *J. Biol. Chem.* **280**, 22343–22355 (2005).
- Mukhopadhyay, S., Clark, D. R., Watson, N. B., Zacharias, W. & McGregor, W. G. REV1 accumulates in DNA damage-induced nuclear foci in human cells and is implicated in mutagenesis by benzo[a]pyrenediol epoxide. *Nucleic Acids Res.* **32**, 5820–5826 (2004).
- Suzuki, T., Gruz, P., Honma, M., Adachi, N. & Nohmi, T. Sensitivity of human cells expressing low-fidelity or weak-catalytic-activity variants of DNA polymerase zeta to genotoxic stresses. *DNA Repair* **45**, 34–43 (2016).
- Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science* **309**, 2219–2222 (2005).
- Swan, M. K., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Structure of the human Rev1-DNA-dNTP ternary complex. *J. Mol. Biol.* **390**, 699–709 (2009).
- Wiltout, M. E. & Walker, G. C. The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. *Genetics* **187**, 21–35 (2011).
- Zhou, Y., Wang, J., Zhang, Y. & Wang, Z. The catalytic function of the Rev1 dCMP transferase is required in a lesion-specific manner for translesion synthesis and base damage-induced mutagenesis. *Nucleic Acids Res.* **38**, 5036–5046 (2010).
- Gibbs, P. E. & Lawrence, C. W. Novel mutagenic properties of abasic sites in Saccharomyces cerevisiae. *J. Mol. Biol.* **251**, 229–236 (1995).
- Otsuka, C., Loakes, D. & Negishi, K. The role of deoxycytidyl transferase activity of yeast Rev1 protein in the bypass of abasic sites. *Nucleic Acids Res. Suppl.* **2**, 87–88 (2002).
- Hoopes, J. I. et al. Avoidance of APOBEC3B-induced mutation by error-free lesion bypass. *Nucleic Acids Res.* **45**, 5243–5254 (2017).
- Cosman, M. et al. Solution conformation of the major adduct between the carcinogen (+)-anti-benzo[a]pyrene diol epoxide and DNA. *Proc. Natl Acad. Sci. USA* **89**, 1914–1918 (1992).
- de los Santos, C. et al. Influence of benzo[a]pyrene diol epoxide chirality on solution conformations of DNA covalent adducts: the (-)-*trans*-anti-[BP]G.C adduct structure and comparison with the (+)-*trans*-anti-[BP]G.C enantiomer. *Biochemistry* **31**, 5245–5252 (1992).
- Cosman, M. et al. Solution conformation of the (+)-*cis*-anti-[BP]dG adduct in a DNA duplex: intercalation of the covalently attached benzo[a]pyrenyl ring into the helix and displacement of the modified deoxyguanosine. *Biochemistry* **32**, 4145–4155 (1993).
- Cosman, M. et al. Solution conformation of the (-)-*cis*-anti-benzo[a]pyrenyl-dG adduct opposite dC in a DNA duplex: intercalation of the covalently attached BP ring into the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry* **35**, 9850–9863 (1996).
- Hsu, G. W., Huang, X., Luneva, N. P., Geacintov, N. E. & Beese, L. S. Structure of a high fidelity DNA polymerase bound to a benzo[a]pyrene adduct that blocks replication. *J. Biol. Chem.* **280**, 3764–3770 (2005).

47. Jha, V., Bian, C., Xing, G. & Ling, H. Structure and mechanism of error-free replication past the major benzo[a]pyrene adduct by human DNA polymerase kappa. *Nucleic Acids Res.* **44**, 4957–4967 (2016).
48. Jha, V. & Ling, H. Structural basis of accurate replication beyond a bulky major benzo[a]pyrene adduct by human DNA polymerase kappa. *DNA Repair (Amst)*. **49**, 43–50 (2017).
49. Bauer, J. et al. A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc. Natl Acad. Sci. USA* **104**, 14905–14910 (2007).
50. Geacintov, N. E. et al. Spectroscopic characteristics and site I/site II classification of cis and trans benzo[a]pyrene diolepoxide enantiomer-guanosine adducts in oligonucleotides and polynucleotides. *Carcinogenesis* **12**, 2099–2108 (1991).
51. Xie, X. M., Geacintov, N. E. & Brody, S. Origins of conformational differences between cis and trans DNA adducts derived from enantiomeric anti-benzo[a]pyrene diol epoxides. *Chem. Res. Toxicol.* **12**, 597–609 (1999).
52. Steitz, T. A. DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* **274**, 17395–17398 (1999).
53. Swan, M. K., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat. Struct. Mol. Biol.* **16**, 979–986 (2009).
54. Gao, Y. & Yang, W. Capture of a third Mg(2)(+) is essential for catalyzing DNA synthesis. *Science* **352**, 1334–1337 (2016).
55. Freudenthal, B. D., Beard, W. A., Shock, D. D. & Wilson, S. H. Observing a DNA polymerase choose right from wrong. *Cell* **154**, 157–168 (2013).
56. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Protein-template-directed synthesis across an acrolein-derived DNA adduct by yeast Rev1 DNA polymerase. *Structure* **16**, 239–245 (2008).
57. Geacintov, N. E. et al. NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity. *Chem. Res. Toxicol.* **10**, 111–146 (1997).
58. Batra, V. K. et al. Structure of DNA polymerase beta with a benzo[c]phenanthrene diol epoxide-adducted template exhibits mutagenic features. *Proc. Natl Acad. Sci. USA* **103**, 17231–17236 (2006).
59. Guo, C. et al. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.* **22**, 6621–6630 (2003).
60. Ohashi, E. et al. Interaction of hREV1 with three human Y-family DNA polymerases. *Genes Cells* **9**, 523–531 (2004).
61. Pustovalova, Y., Bezsonova, I. & Korzhnev, D. M. The C-terminal domain of human Rev1 contains independent binding sites for DNA polymerase eta and Rev7 subunit of polymerase zeta. *FEBS Lett.* **586**, 3051–3056 (2012).
62. Wojtaszek, J. et al. Structural basis of Rev1-mediated assembly of a quaternary vertebrate translesion polymerase complex consisting of Rev1, heterodimeric polymerase (Pol) zeta, and Pol kappa. *J. Biol. Chem.* **287**, 33836–33846 (2012).
63. Hashimoto, K. et al. The vital role of polymerase zeta and REV1 in mutagenic, but not correct, DNA synthesis across benzo[a]pyrene-dG and recruitment of polymerase zeta by REV1 to replication-stalled site. *J. Biol. Chem.* **287**, 9613–9622 (2012).
64. Yoon, J. H. et al. Rev1 promotes replication through UV lesions in conjunction with DNA polymerases eta, iota, and kappa but not DNA polymerase zeta. *Genes Dev.* **29**, 2588–2602 (2015).
65. Cosman, M., Ibanez, V., Geacintov, N. E. & Harvey, R. G. Preparation and isolation of adducts in high yield derived from the binding of two benzo[a]pyrene-7,8-dihydroxy-9,10-oxide stereoisomers to the oligonucleotide d(ATATGTATA). *Carcinogenesis* **11**, 1667–1672 (1990).
66. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
67. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
68. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
69. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
70. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
71. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

Acknowledgements

We thank the staff at beam line X25 at the Brookhaven National Laboratory for facilitating x-ray data collection. This work was partially supported by NIH grants R01 CA200575 (A.K.A.) and R01 ES024050 (N.E.G.).

Author contributions

O.R. and A.K.A. designed the experiments. O.R. and A.K. prepared the site- and stereospecifically BP- N^2 -dG-modified DNA templates under the supervision of N.E.G.; O.R. and H.L. performed protein expression, purification and crystallization. O.R. performed x-ray data collection and refinement. O.R. and A.K.A. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01013-5.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

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) 2017