

ARTICLE

Received 1 Sep 2016 | Accepted 16 Feb 2017 | Published 3 May 2017

DOI: 10.1038/ncomms14959

OPEN

A domain in human EXOG converts apoptotic endonuclease to DNA-repair exonuclease

Michal R. Szymanski^{1,2}, Wangsheng Yu^{1,2}, Aleksandra M. Gmyrek³, Mark A. White^{2,3}, Ian J. Molineux⁴, J. Ching Lee^{2,3} & Y. Whitney Yin^{1,2}

Human EXOG (hEXOG) is a 5'-exonuclease that is crucial for mitochondrial DNA repair; the enzyme belongs to a nonspecific nuclease family that includes the apoptotic endonuclease EndoG. Here we report biochemical and structural studies of hEXOG, including structures in its apo form and in a complex with DNA at 1.81 and 1.85 Å resolution, respectively. A Wing domain, absent in other $\beta\beta\alpha$ -Me members, suppresses endonuclease activity, but confers on hEXOG a strong 5'-dsDNA exonuclease activity that precisely excises a dinucleotide using an intrinsic 'tape-measure'. The symmetrical apo hEXOG homodimer becomes asymmetrical upon binding to DNA, providing a structural basis for how substrate DNA bound to one active site allosterically regulates the activity of the other. These properties of hEXOG suggest a pathway for mitochondrial BER that provides an optimal substrate for subsequent gap-filling synthesis by DNA polymerase γ .

¹Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, Texas 77555, USA. ²Sealy Center for Structural Biology, University of Texas Medical Branch, Galveston, Texas 77555, USA. ³Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, Texas 77555, USA. ⁴Department of Molecular Biosciences, University of Texas at Austin, Austin, Texas 78712, USA. Correspondence and requests for materials should be addressed to Y.W.Y. (email: ywyin@utmb.edu).

High concentrations of reactive oxygen species in the mitochondria result in a high incidence of mitochondrial DNA (mtDNA) oxidative damage, estimated to be about ten times greater than in nuclear DNA^{1,2}. Human mtDNA encodes a subset of components for the oxidative phosphorylation electron chain that utilizes a proton potential to synthesize ATP. If left unrepaired, lesions on mtDNA can stall DNA replication and/or cause mutations, reducing mitochondrial function and cellular energy supplies³. Oxidative mtDNA damage has been attributed to increased cancer incidence and premature aging^{4,5}.

Base excision repair (BER) is the major mechanism for correcting mtDNA oxidative damage⁶. In this multi-enzyme reaction pathway, the DNA product from the previous reaction is the substrate for the next⁷. DNA intermediate 'hand-off' between the repair enzymes is therefore essential, as a prematurely released repair intermediate could be more detrimental than the original damage⁸. Increasing evidence indicates that the repair enzymes directly interact, forming a 'repairoosome'⁹.

Human mitochondrial BER is distinct from nuclear BER¹⁰. There is no designated repair polymerase in the mitochondria, and the replicase, Pol γ , is essential for both functions^{11,12}. However, Pol γ is least efficient in synthesizing DNA at a single-nucleotide gap and it lacks robust strand-displacement ability, activities necessary, respectively, for canonical nuclear single-nucleotide and long-patch (LP) BER¹³. However, the efficiency of Pol γ gap-filling increases markedly on even a two-nucleotide gap¹³. If a nuclease could generate a larger gap at a lesion site, mitochondrial BER could yield products similar to that found in nuclear LP-BER, albeit through a different mechanism.

Human mitochondrial nuclease EXOG (hEXOG) is important for mtDNA integrity. hEXOG depletion selectively increases lesions in mtDNA, but has no effect on chromosomal DNA. Conversely, depletion of nucleases that are critical for nuclear BER, flap endonuclease 1 (FEN1) and DNA2 has no or limited effect on mtDNA¹⁴. Ectopic expression of hEXOG increases resistance to oxidative stress of proliferating myoblasts¹⁵, and depletion of EXOG increases oxidative consumption rate in primary neonatal rat ventricular cardiomyocytes¹⁶. These studies provide strong support to the idea that hEXOG is involved in mtDNA repair. Importantly, hEXOG is found in a complex with other mtDNA repair enzymes, APE1, Pol γ and ligase III; their interaction is enhanced by oxidative stress, and is regulated by poly(ADP-ribose) polymerase-1¹⁷.

Human EXOG shares sequence homology with members of $\beta\beta\alpha$ -Me nuclease family¹⁸, which contains nonspecific endonucleases. One member, EndoG, is the paralog of hEXOG¹⁹. The nonspecific activity of EndoG is perfectly suited for activities during apoptosis when it digests single-stranded DNA and double-stranded DNA (dsDNA), and RNA. However, nonspecific activity is incompatible with a precise exonuclease activity necessary to process DNA 5'-end in BER^{19–21}. In addition to endonuclease activity, hEXOG also possesses 5'-exonuclease activity; relative to the endonucleases, hEXOG harbours a C-terminal extension that may contribute to the new function.

Here we present enzymatic and crystallographic studies of hEXOG. Our results provide a detailed molecular and structural mechanism for hEXOG's substrate specificity and enzymatic properties in BER. The catalytic Core domain of hEXOG is identical to EndoG, but its C-terminal extension forms a unique Wing domain that converts the enzyme from a nonspecific endonuclease into a DNA-geometry-specific exonuclease. Our studies suggest that mtDNA repair is performed by long-patch BER that provides an optimal substrate for Pol γ when fulfilling its role in DNA repair.

Results

Refolding and purification of hEXOG. hEXOG is a mitochondrial membrane protein¹⁹, and the hEXOG construct used here lacks the N-terminal trans-membrane domain (2–41 aa) and a predicted unstructured region (42–58 aa) of the native protein in an attempt to improve solubility; the resulting hEXOG- Δ N58 enzyme is defined here as wild type. However, our hEXOG construct did not produce soluble protein; hEXOG was therefore partially purified from inclusion bodies as completely denatured protein and then refolded. Success of folding was assessed by the emergence of secondary structure features in circular dichroism (CD) spectra (Supplementary Fig. 1a). Our optimized procedure consistently refolded >85% of the total protein found in the inclusion body, increasing the refolding yield by ~1,000-fold over that previously attained¹⁹. Refolded hEXOG was then further purified by gel filtration chromatography to >95% homogeneity. The elution profile indicates that the 36.3 kDa hEXOG is a dimer (Supplementary Fig. 1b). All the independent enzyme preparations used in this work were equally active in assays and crystallized under the same conditions.

Functional characterization of hEXOG. We examined hEXOG exonuclease activity on substrates relevant to BER, including a 20 bp dsDNA and a 1-nt gapped DNA (Fig. 1; Supplementary Table 1). One strand was labelled with either 5'-³²P (Fig. 1a) or 3'-fluorescein (3'-F) (Fig. 1c) to allow quantification of reaction products. For simplicity, the DNA is numbered from its 5'-end (Fig. 1c). On the 5'-³²P-labelled duplex (Fig. 1a), the reaction product is a dinucleotide, suggesting that hEXOG has 5'-exonuclease activity. When the same 20 bp duplex is labelled with 3'-F, a 3'-F 18-mer forms, confirming that the enzyme excises a dinucleotide from the 5'-end (Fig. 1c). This property distinguishes hEXOG from typical exonucleases, which cleave the phosphodiester bond between the 1st and 2nd nucleotide, yielding mononucleotides. On the BER intermediate containing a 1-nt gap, hEXOG also produces 5'-dinucleotide products (Fig. 1a). hEXOG incises duplex and gapped DNA with equal efficiency (Fig. 1b), suggesting that the enzyme has higher affinity towards a 5'-P. To test this hypothesis, we directly measured the affinity of EXOG for DNA with 5'-P and 5'-OH. The K_d values of hEXOG to 5'-P and 5'-OH dsDNAs measured by fluorescence anisotropy are 6.2 ± 2 and 22.2 ± 5.1 nM, respectively, indicating that hEXOG binds to 5'-P-DNA greater than threefold stronger than the 5'-OH-DNA (Supplementary Fig. 1c).

Under conditions of excess substrate, hEXOG generates product in two distinct phases where the first cycle of excision is much faster than subsequent reactions. The biphasic reaction persists at enzyme concentrations (Fig. 1d) that are higher than the apparent K_d for DNA binding (Supplementary Fig. 1c). The two reaction rates allow us to calculate the steady-state turnover rate. Under the conditions employed, k_{off} is 0.00028 s^{-1} (Fig. 1e). Thus, while the first cycle of reaction is completed in <10 s, the turnover rate for the second reaction is 1 h, suggesting very slow product release. The nucleolytic reaction catalysed by hEXOG therefore approximates to single-turnover, which effectively eliminates unnecessary nuclease activity that is potentially detrimental during DNA repair. The single-turnover nature of the hEXOG reaction also enables us to measure the amount of product formed in one reaction cycle. When various concentrations of hEXOG were mixed with excess substrate, the product yield was ~50% of the active sites (Supplementary Table 2).

To determine whether both monomers of EXOG dimer are capable of binding to DNA, we measured hEXOG-DNA binding using electrophoretic mobility shift assay, and found that hEXOG monomers bind to 90% of DNA at 1:1 stoichiometry

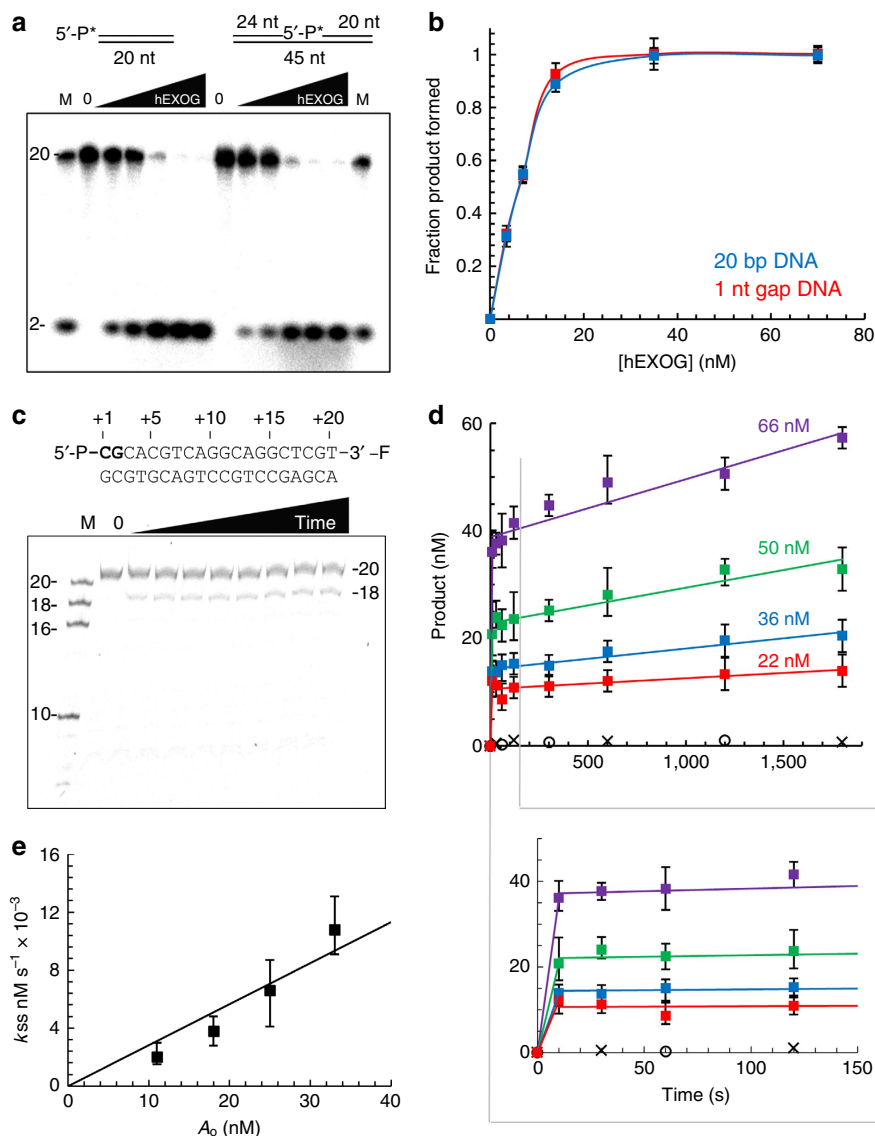


Figure 1 | hEXOg exonuclease activity. (a) Reactions used a 5'-³²P (*) 20 bp duplex or a 1-nt gapped DNA (10 nM) incubated with hEXOg_(monomer) at 7, 14, 28, 70 and 140 nM for 5 min. Products were visualized after denaturing polyacrylamide gel electrophoresis. (b) Quantification of product from reactions in a. (c) Time-dependent exonuclease reaction using 200 nM 3'-F-labelled 20 bp duplex with 66 nM hEXOg_(monomer). (d) Time-dependent product formation at various hEXOg_(monomer) concentrations. Control reactions: H140A hEXOg (o) and wild-type enzyme + EDTA (x) are indicated in black. The initial rate of production formation is underestimated. (e) Plot of k_{ss} versus A_0 . Data represent the mean \pm s.e.m. of three independent experiments.

(Supplementary Fig. 1d). The DNA-binding property of hEXOg is determined by isothermal titration calorimetry (ITC) using a nuclease-deficient (H140A) mutant. The resulting binding curve is bimodal; the simplest model that best describes the experimental data is with two distinct binding constants (K_a): 2.4×10^5 and $1.2 \times 10^7 \text{ M}^{-1}$ (Supplementary Fig. 2), implying a different mode of binding for the second monomer of a dimer. The two affinities differ by 50-fold, suggesting that within certain DNA and enzyme concentrations, hEXOg binds only one DNA molecule, and the half-site reactivity is caused by the two different DNA-binding affinities in hEXOg dimer.

Structure of homodimeric hEXOg in apo form. To provide structural insights into the enzymatic properties of hEXOg, we first determined a crystal structure of hEXOg in its apo form at 1.81 Å resolution by molecular replacement using *Drosophila* EndoG (PDB code: 3ISM) as a search model²². The C-terminal extension (residues 300–369) of hEXOg that is absent in the

model was then built into a 2Fo-Fc electron density map. The final structure was refined to $R_{\text{factor}} = 16.1\%$ and $R_{\text{free}} = 18.7\%$ (Table 1).

Apo hEXOg is a symmetrical homodimer. Each monomer contains a Core domain that is extremely similar to that of EndoG (r.m.s.d. = 0.8 Å)^{22–24}, and a helical-bundle Wing domain that corresponds to the C-terminal extension (Fig. 2a,b). However, although the N-termini of the hEXOg and EndoG possess the same fold, they exhibit very different structures. The N terminus of each hEXOg monomer undergoes a domain-swap where the folding of the region is identical to that in EndoG, but the involved residues are from the opposing monomer rather than from the same monomer in EndoG: the residues ⁵⁸KAVLEQFGFPLT⁷⁰ (hydrophobic residues are bold) of one monomer collapse on a hydrophobic patch (²⁷⁷GLVFFPHL²⁸⁴) of the opposing monomer 30 Å away (Fig. 2a,d; Supplementary Fig. 3a). The swapped domains represent 45% of the total dimer interface in hEXOg, and thus contribute substantially to dimer

Table 1 | Data collection and refinement statistics*.

	apo hEXOG complex I	apo hEXOG complex II	hEXOG-DNA complex I	hEXOG-DNA complex II
PDB	5T40	5T3V	5T5C	5T4I
hEXOG	Wild type	Wild type	H140A	H140A
Substrate DNA	—	—	10 bp	10 bp
Metal ion	Mg ²⁺	Mn ²⁺	Mg ²⁺	Mn ²⁺
Wavelength	1.0	1.89	1.0	1.89
<i>Data collection</i>				
Space group	P 2 ₁	P 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	73.37, 83.73, 75.00	73.34, 83.44, 74.77	73.50, 80.27, 139.18	73.51, 79.72, 138.87
α , β , γ (°)	90.00, 113.54, 90.00	90, 113.42, 90.00	90, 90, 90	90, 90, 90
Resolution (Å)	34.7–1.81 (1.83–1.81) [†]	49.7–2.6 (2.66–2.60)	39.2–1.85 (1.87–1.85)	40.0–2.39 (2.45–2.39)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.052 (0.814)	0.058 (0.337)	0.08 (1.0)	0.114 (0.638)
<i>I</i> / σ <i>I</i>	25.8 (1.54)	21.7 (2.3)	31.5 (1.56)	24.8 (2.4)
CC [†]	1.000 (0.926)	0.999 (0.989)	1.000 (0.957)	0.994 (0.994)
Completeness (%)	98.10 (97.8)	95.7 (93.3)	99.9 (99.1)	99.5 (97.4)
Redundancy	3.6 (3.4)	3.6 (3.6)	7.2 (6.3)	5.7 (4.4)
<i>Molecular replacement</i>				
Model	3ISM	5T40	5T40	5T5C
LLG	592.97	518.40	686.22	13,776.6
TFZ	18.2	42.6	34.6	57.2
<i>Refinement</i>				
Resolution (Å)	34.7–1.81	49.7–2.6	39.23–1.85	40.03–2.39
No. reflections	74,273	24,370	70,421	32,854
<i>R</i> _{work} / <i>R</i> _{free}	16.1/18.7	17.6/22.4	18.7/21.8	19.1/22.9
No. of atoms				
Protein	9,552	9,552	9,465	9,464
Ligand/ion	22	12	1,148	1,144
Water	440	81	272	57
<i>B</i> -factors				
Protein	39.37	63.54	58.41	63.16
Ligand/ion	57.95	127.37	32.09	57.31
Water	44.36	48.30	46.26	47.82
r.m.s.d.				
Bond lengths (Å)	0.004	0.003	0.003	0.003
Bond angles (°)	0.7	0.483	0.59	0.53

hEXOG, human EXOG.

*One crystal was used for each data set.

†Values in parentheses are for highest-resolution shell.

stability. A G227V substitution in the hydrophobic patch that disturbs the dimer interface of hEXOG displays markedly reduced activity¹⁹, suggesting that dimerization is functionally important. A Mg²⁺ is found in each active site, forming octahedron coordination with five water molecules and N₈ of N171. The nature of Mg²⁺ was confirmed by Mn²⁺ substitution and single-wavelength anomalous dispersion data (Supplementary Fig. 3b,c).

In contrast to the completely solvent exposed active site of EndoG, the active site of hEXOG is an enclosed V-shaped cleft. The cleft, measuring 32–35 Å in depth and 15 Å in width, is lined with positively charged residues from both Core and Wing domains, and is both electrostatically and dimensionally capable of accommodating dsDNA at its end (Fig. 2c). As the Core domain of hEXOG is highly similar to EndoG and other members of the $\beta\beta\alpha$ -Me superfamily that have no specificity, we hypothesized that the specific exonuclease activity of hEXOG would be attributed to the Wing domain. We therefore determined crystal structures of hEXOG complexed to substrate to understand substrate specificity and the half-site reactivity of the enzyme.

Structure of hEXOG–dsDNA complex. The nuclease-deficient variant, hEXOG–H140A, was refolded and purified similarly to

wild type (Supplementary Fig. 1) and used for co-crystallization with a 10 bp DNA. To distinguish the two strands of DNA in the duplex, only one strand contained a 5'-P. The structure was determined to 1.85 Å resolution by molecular replacement using the apo hEXOG structure. Ideal B-form dsDNA was built into the composite omit electron density map. The structure was then refined without any geometrical constraints on the DNA; the final structure has *R*_{factor} and *R*_{free} values of 18.7% and 21.8%, respectively (Table 1).

Successful crystallization experiments were performed at enzyme concentrations greatly exceeding the *K*_d value for the weaker DNA-binding site of the hEXOG dimer as determined by ITC; consequently, DNA occupied both active sites (Fig. 3a). Both DNA duplexes are oriented such that the 5'-P end of the substrate is proximal and the 5'-OH end is distal to the active site. Thus, hEXOG has a clear preference for a 5'-P on the substrate strand. The 5'-OH containing strand thus defines the complementary strand.

DNA binding disrupts the symmetry of the apo hEXOG dimer.

Binding to DNA does not change the structure of the Core domain; the r.m.s.d. between the apo and DNA complex forms is only 0.24 Å. However, the two Wing domains undergo different

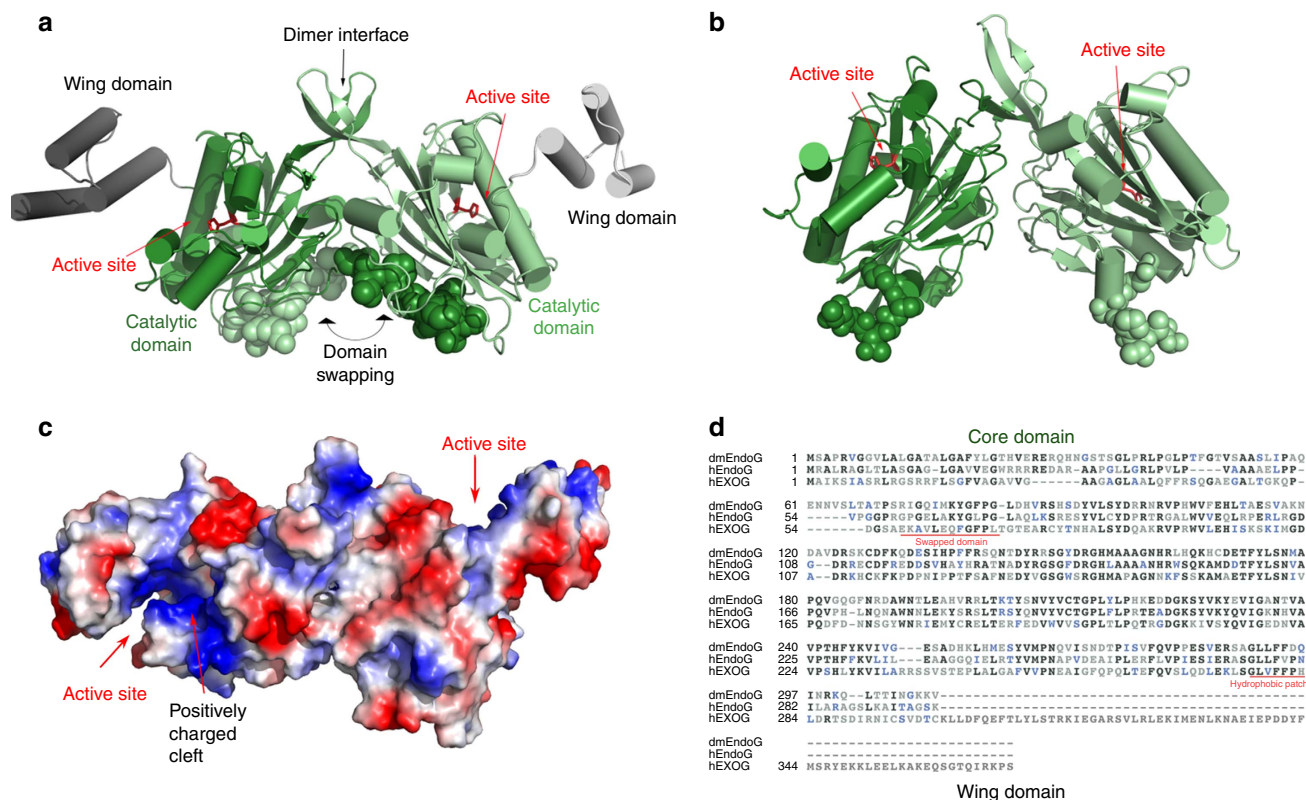


Figure 2 | Structure of apo hEXOg. (a) Structure of apo hEXOg. Each monomer has a Core (green) and a Wing domain (grey); residues involved in domain swapping are shown in CPK. (b) Structure of *Drosophila* EndoG (dmEndoG) with N-terminal residues shown in CPK. (c) Electrostatic surface of apo hEXOg. (d) Sequence alignment of hEXOg, hEndoG and dmEndoG.

conformational changes from their positions in the apo structure: one rotates only 3°, whereas the other rotates 35° away from the DNA (Fig. 3b,c,e). One active site therefore undergoes an atypical ‘closed-to-open’ conformational change. Consequently, upon DNA binding, hEXOg becomes asymmetrical where one active site remains in the ‘closed’ configuration and the other becomes ‘open’. Thermal motion of the structure is revealed by changes in the B-factor. The average B-factors for the Wing domain of the two monomers in the Apo form are ~75 Å² but in the DNA-bound form the B-factor of the open domain increases to 138 Å², whereas the closed domain remains at 76 Å² (Supplementary Fig. 4). These analyses suggest the ‘open’ Wing domain becomes less ordered than in the Apo structure, likely induced by the allosteric effect of DNA binding to the other active site.

Substrate specificity of hEXOg. The dsDNA is bound to each positively charged active site cleft; the substrate strand interacts primarily with the Core domain and the complementary strand with the Wing domain (Fig. 3a; Supplementary Fig. 5), explaining why the Wing domain of hEXOg increases specificity for dsDNA. Since neither Core domain of the dimeric hEXOg undergoes a structural change on binding DNA, the two substrate strands make identical electrostatic interactions (Fig. 3c,e). However, the conformational changes in the two Wing domains result in different DNA interactions. In the closed active site, the complementary strand interacts with Wing residues R³¹⁴, R³¹⁵, R³²⁰, R³²⁴ and K³²⁷, with R³²⁴ forming a bipartite interaction with the +3 nt DNA backbone at distances of 3.1 and 3.4 Å (Fig. 3e); R³¹⁵ and K³²⁷ are, respectively, 4.3 and 6.7 Å from the +2 nt DNA backbone. In the open active site, the complementary DNA strand lacks several interactions, because the

bipartite distances to R³²⁴ are increased threefold (to 9.4 and 9.8 Å), and R³¹⁵ and K³²⁷ are similarly displaced (10.2 and 16.4 Å, respectively) from the +2 nt backbone (Fig. 3c). As electrostatic force is proportional to distance squared, a threefold increase in distance reduces the local attraction force to <10% of the ‘closed’ active site. These reduced interactions, together with the associated reduced desolvation of DNA, explain why the open active site has a lower affinity for DNA—and thus lower reactivity. These structural conclusions are in agreement with the two DNA-binding modes of hEXOg revealed by ITC.

To verify the structure of hEXOg free of a crystal lattice, we also determined the solution structure of hEXOg–DNA complex using small-angle X-ray scattering (SAXS). The SAXS hEXOg–DNA model established with CRY SOL²⁵ fits well with the crystal structure ($\chi^2 = 1.7$; Supplementary Fig. 6), indicating good agreement between solution and crystallographic structures.

The tape measure for DNA cleavage. The structural rigidity of the Core domain enabled us to generate a composite structure of an active substrate complex by replacing A¹⁴⁰ in the hEXOg–H140A:DNA complex with H¹⁴⁰ of the wild-type enzyme (Fig. 4a,b). In this complex, the 5′-P end of the substrate strand docks at the bottom of the active site cleft against residues from the Wing domain, forming clustered π - π interactions with F³⁰⁷ and Y³¹⁰ (Fig. 4c). Two positively charged arginines, R¹³⁸ and R³¹⁴, bind to the DNA backbone at +1 and +4 nt, flanking the scissile bond (Fig. 4b). Interestingly, R¹³⁸ and R³¹⁴ bind two sulfate molecules in the apo structure in an essentially identical manner to the phosphodiester backbone of substrate DNA, reflecting their affinity for negative charges (Fig. 4a). Binding to DNA thus affords the substrate strand a register so that the oxygen bridging the +2 and +3 nt is placed in the catalytic

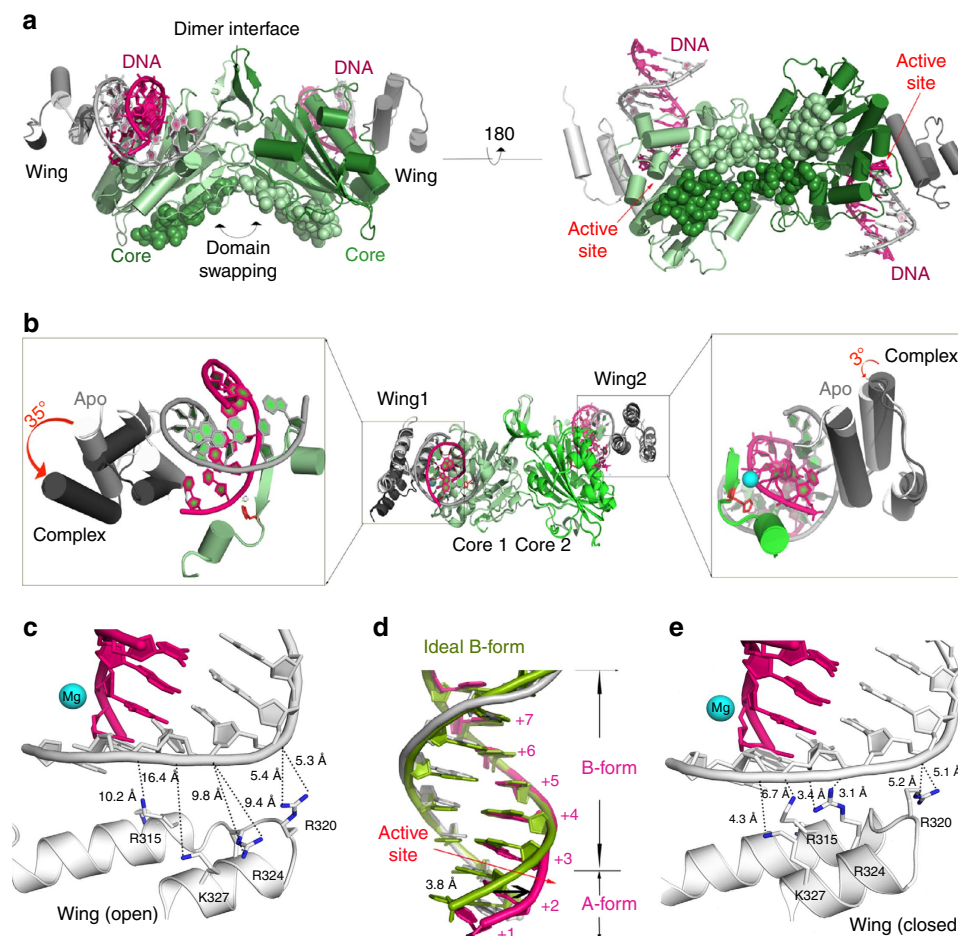


Figure 3 | Structure of hEXOG-dsDNA complex. (a) The complex has dsDNA in each active site. (b) Asymmetry of hEXOG upon DNA binding. Superposition of the apo enzyme and in a DNA complex in the open (left) and closed (right) monomer shows wing domain movements. (c) DNA interactions with the ‘open’ Wing domain. (d) The bound dsDNA (magenta and grey) partially transition to A-form from an ideal B-form (green). (e) DNA interactions with the ‘closed’ Wing domain.

active site. Because the Wing domain provides the register for DNA binding, it therefore serves as a ‘tape-measure’ for nucleolytic cleavage in a sequence-independent manner.

Structural basis for catalysis. In the substrate complex structure, DNA base pairs +1 to +3 adopt an A-form configuration with a 3'-endo sugar pucker (Fig. 3d; Supplementary Fig. 3d), while the remainder of the DNA is B-form and with 2'-endo or -exo sugar pucker (Supplementary Table 3). The B→A transition widens the minor groove, placing the substrate strand deeply into the active site cleft (Fig. 2) and bringing the O3' of the +3 nt 3.8 Å closer to the Mg²⁺ ion than if the DNA remained in B-form (Fig. 3d).

The Mg²⁺ ion that was coordinated with five water molecules and the δ-N of Asn¹⁷¹ in the apo structure is now coordinated with the bridging oxygen between the 2nd and 3rd nucleotide and the non-bridging oxygen O_p of the +2 nt in the DNA complex (Fig. 4d). The Mg²⁺ ion and H¹⁴⁰ are, respectively, 2.5 and 3.0 Å from the scissile bond, the distances that are competent for catalysis. The 3'-endo sugar pucker of the +3 nt selectively exposes the O3' atom of the 2nd to 3rd phosphodiester linkage, facilitating nucleophilic attack at the scissile bond. While H¹⁴⁰ coordinates the attacking water, Mg²⁺ serves as a Lewis acid that enhances the electrophilicity of the phosphorus so that it facilitates the nucleophilic reaction. The structure suggests that both H¹⁴⁰ and Mg²⁺ perform catalytic roles (Fig. 4d). This conclusion is consistent with biochemical assays showing that the enzyme is

inactive in the absence of either H¹⁴⁰ or Mg²⁺ (Fig. 1d). No potential attacking water molecules are observed in either the hEXOG-H140A-DNA complex or wild-type apo structures, suggesting that the attacking water requires stabilization by both the imidazole moiety of H¹⁴⁰ and the DNA backbone.

Structural basis for slow product release. In the substrate complex, the 5'-P of the substrate strand forms an electrostatic interaction with the positively charged R³¹⁴ (Fig. 5a). As hEXOG excises without fraying the DNA, the same interaction should persist in the product complex, where R³¹⁴ interacts with the 5'-P of the dinucleotide reaction product. We hypothesize that R³¹⁴ is, at least in part, responsible for both binding substrate and retention of the reaction product. Combining both properties into the same residue should reduce the rate of product release.

To test this structural prediction, we constructed the hEXOG mutant R314A that was refolded and purified similarly to wild type (Supplementary Fig. 1) and assayed its exonuclease activity. The apparent DNA-binding constant for the mutant is 35.7 ± 7.2 nM, that is, sixfold lower than wild type (Supplementary Fig. 1c), because eliminating the interaction between R³¹⁴ with 5'-P reduces substrate affinity. Notably, the reaction rate is at least twofold faster for the R314A mutant, together with its sixfold lowered binding affinity to the 5'-P-containing DNA (Fig. 5b,c), the increased catalysis is probably caused by faster product release and not increased chemistry, as R³¹⁴ is not near the active site.

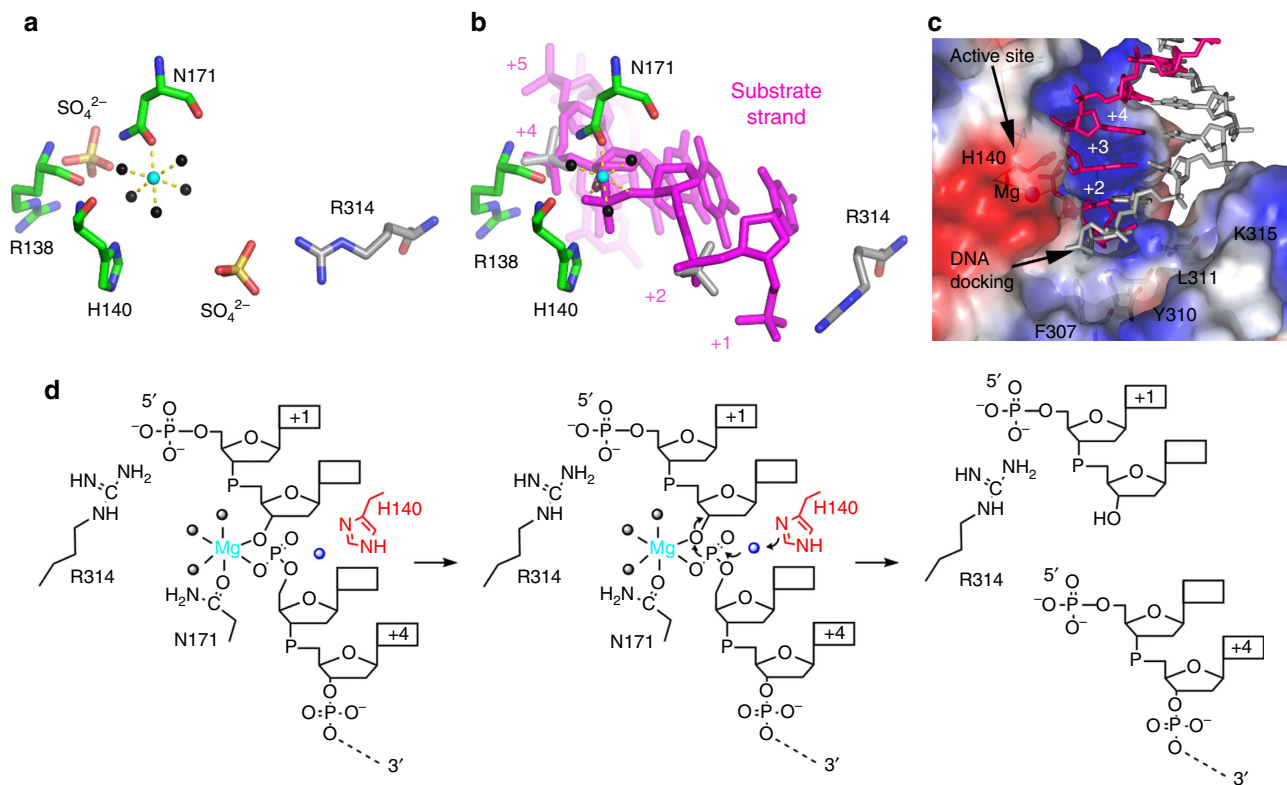


Figure 4 | Proposed reaction mechanism for hEXOg. (a) The active site of apo hEXOg illustrating R¹³⁸ and R³¹⁴ interactions with SO₄²⁻ ions, and Mg (turquoise sphere) interactions with water molecules (black spheres) and N171. (b) The active site of hEXOg-DNA complex. DNA replaces water molecules in coordinating the Mg²⁺ ion. R¹³⁸ and R³¹⁴ bind to the DNA backbone at the +1 and +4 nt, respectively. (c) The crystal structure of EXOG-DNA complex shows that the Wing domain forms a docking site for the 5'-P of the substrate strand. (d) Proposed mechanism of DNA cleavage by hEXOg.

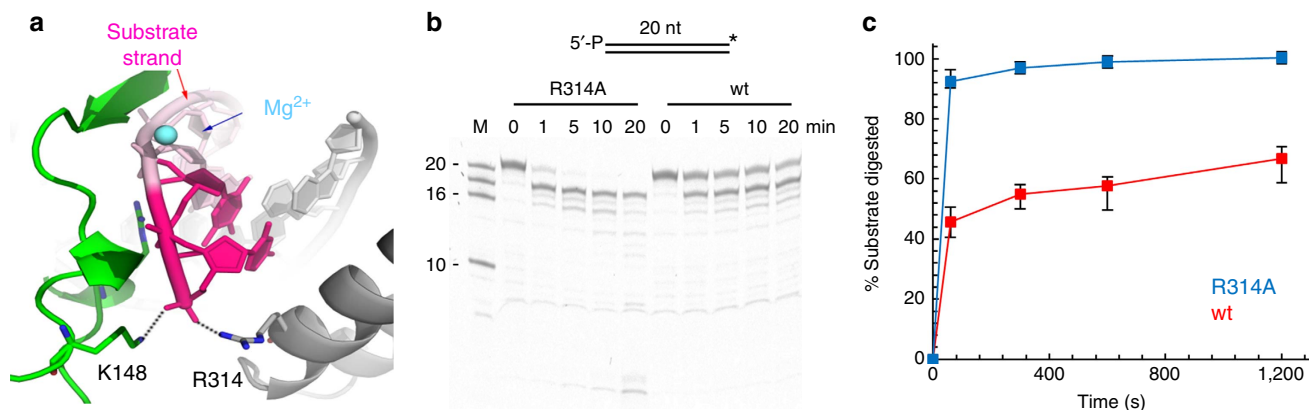


Figure 5 | Structural basis for slow product release. (a) Electrostatic interactions of R³¹⁴ and K¹⁴⁸ with the 5'-P of the substrate strand. (b) Time-dependent nucleolytic reaction of hEXOg R314A_(monomer) (150 nM) on 3'-F dsDNA (200 nM). (c) Quantification of digestion. Data represent the mean \pm s.e.m. of three independent experiments.

Deletion of the Wing domain abolishes substrate specificity.

The structure of hEXOg-DNA complex suggests that the Wing domain is responsible for the tape-measured exonuclease activity and substrate specificity. To validate this conclusion, a Wing domain deletion mutant, hEXOg- Δ C68 was constructed. The mutant refolds such as wild type and remains a dimeric enzyme (Supplementary Fig. 1). When assayed on short DNA, the deletion mutant hEXOg- Δ C68 cleaves DNA randomly, as evidenced by the appearance of multiple product bands on a gel (Fig. 6a). Similar to the R314A mutant, hEXOg- Δ C68 also displays rapid turnover, providing further evidence that the Wing domain regulates nucleolytic activity. The hEXOg- Δ C68 mutation

eliminates critical residues involved in binding the 5'-P of both the substrate and product. The active site is completely open and the enzyme thereby loses both tape-measure controlled hydrolysis and any specificity for dsDNA (Fig. 6b). However, when assayed for endonuclease activity on circular ϕ X174 DNA, hEXOg- Δ C68 exhibits catalytic activity that is eight to ten times higher than wild type (Fig. 6c,d). This increased activity of hEXOg- Δ C68 is most likely due to the loss of specificity for dsDNA, properties that should be shared by EndoG because hEXOg- Δ C68 and EndoG are similar in length and share the identical Core structure. The Wing domain of hEXOg therefore transforms a non-specific nuclease into a substrate-specific exonuclease. Conversely,

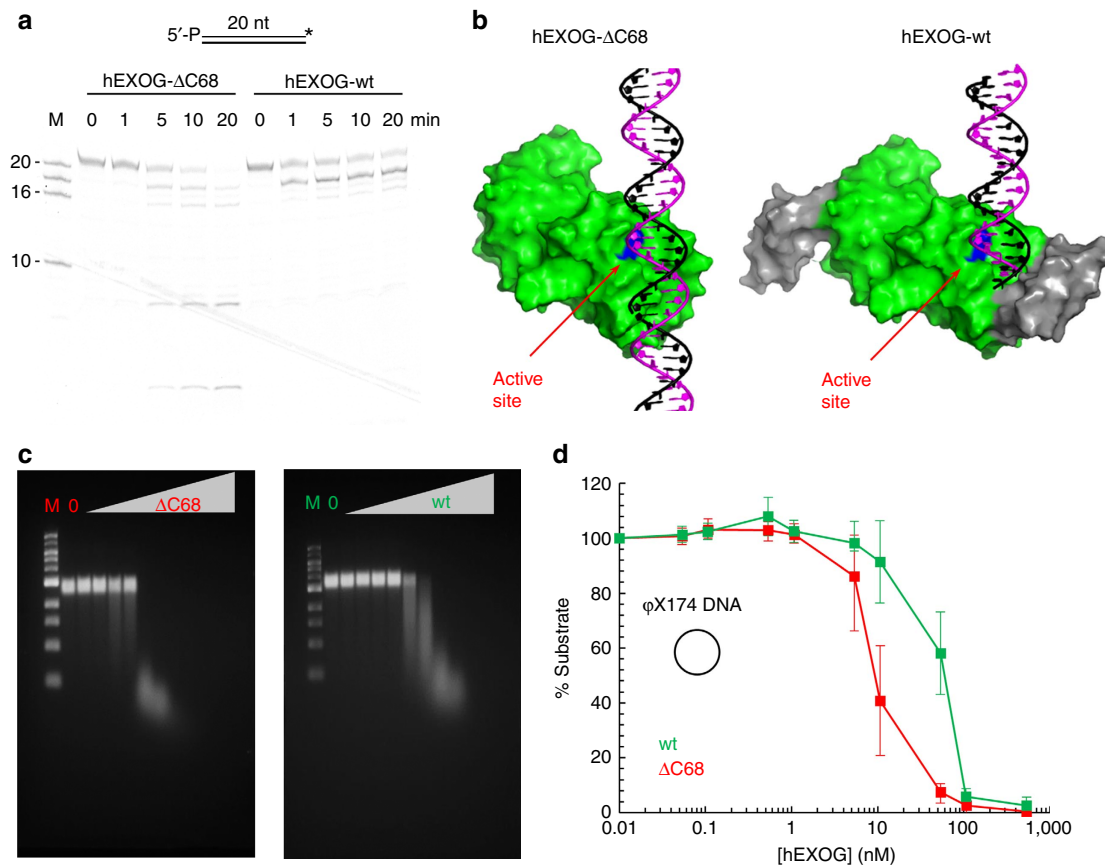


Figure 6 | Wing domain deletion and substrate specificity. (a) Digestion of 3'-F dsDNA (200 nM) by wild-type or Δ C68 hEXO_(monomert) (150 nM). (b) Modelled hEXO- Δ C68 showing loss of tape-measure and dsDNA specificity. (c) Endonuclease activity of hEXO- Δ C68 and wild type on ϕ X174 DNA (50 nM). (d) Quantification of c. Data represent the mean \pm s.e.m. of three independent experiments.

by deleting the Wing domain from hEXO, we have restored the basal nonspecific activity typical of the $\beta\beta\alpha$ -Me nuclease family.

Discussion

The Wing domain provides hEXO with features that are commonly seen in other DNA repair nucleases²⁶, including substrate specificity, slow product release and self-regulated nucleolytic reaction. Importantly, hEXO enhances Pol γ DNA synthesis by providing an optimal substrate¹³, therefore increasing the overall efficiency of the mtBER pathway. The Wing domain exerts its effect in several ways. First, it remodels the DNA-binding site of an EndoG-like enzyme to gain its preference for duplex DNA. Second, the Wing and Core domains each contribute a positively charged arginine, R³¹⁴ and R¹³⁸, in binding the substrate DNA backbone flanking the scissile bond (Fig. 7a). Third, R³¹⁴ of the Wing domain forms an electrostatic interaction with the substrate strand 5'-P, enhancing 5' exonuclease activity and reducing product release. Fourth, the Wing domain provides a tape-measure by positioning the dsDNA substrate for an incision between the 2nd and 3rd nucleotides. Canonical BER can be halted under oxidative environment because oxidized 5'-phosphosphate ribose by formation of a covalent bond with an active site lysine in the 5'-phosphosphate ribose lyase²⁷. Since the active site in hEXO is distant from the 5'-end of the DNA, crosslinking of the 5'-end of the DNA substrate with hEXO will not impact its nuclease activity. This mechanism of cleavage suggests that hEXO is likely to process a damaged 5'-end regardless of chemical structure, circumventing a potential blockage in the DNA repair pathway.

In mitochondrial BER, Pol γ performs gap-filling synthesis following end-processing of the damaged DNA. Because Pol γ is ineffective in synthesizing at a single-nucleotide gap, only a nuclease that generates DNA products with a gap size greater 2 nt is suitable for the known activities of Pol γ ¹³. Among the repair nucleases found in human mitochondria, FEN1 (ref. 28), helicase-nuclease Dna2 (refs 28,29) and hEXO¹⁴, only the latter provides this activity. Together with the ability of hEXO to interact with Pol γ and DNA ligase III, this study has established a precise mechanism of hEXO activity that makes it indispensable for mtDNA repair. A different mitochondrial nuclease, MGME1 (mitochondrial genome maintenance exonuclease 1) processes single-stranded and flap DNAs, but has low activity towards BER intermediates, gapped DNAs^{30,31}. Hence, hEXO and MGME1 divide the repair and replication functions in mtDNA maintenance.

The Wing domains become asymmetrical upon binding to DNA: one monomer interacts with the substrate tightly and the active site closes, the other active site remains open and flexible, and binds DNA loosely. This negative allosteric regulation is most likely mediated through the Core domains connecting action in the active site of one Wing domain to the other. In an attempt to map the connecting path between the two, we noticed that, although binding to DNA does not significantly change the stability of the Core domains, local thermal motion is increased in three important regions: the active site (residues 140–150), the Mg²⁺-binding site (residues 170–190) and a portion of the dimer interface (residues 203–215) (Supplementary Fig. 4). The involvement of these regions needs further investigation but their locations are suggestive of a pathway allowing inter-domain allosteric regulation.

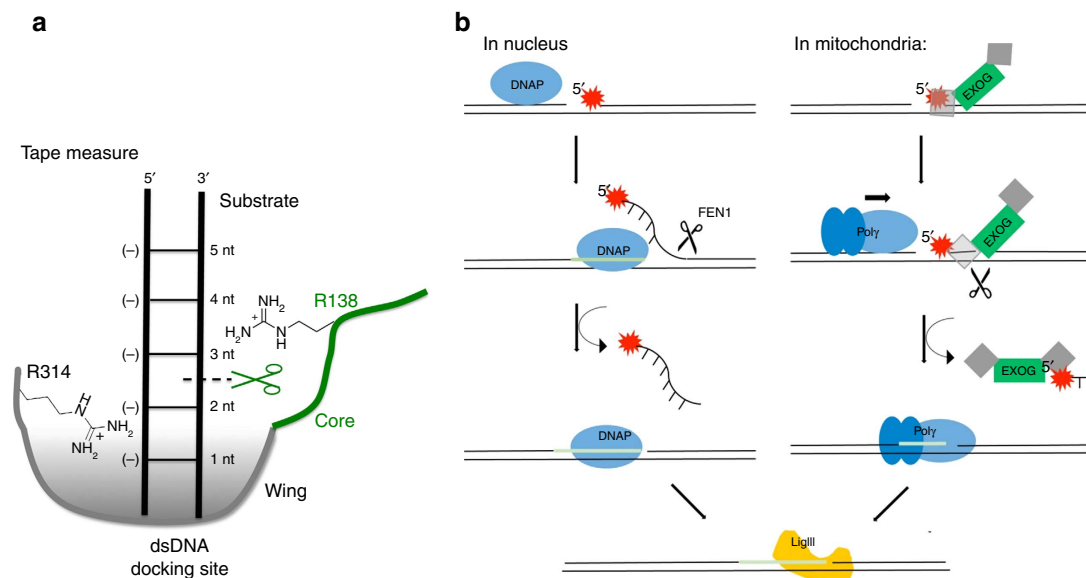


Figure 7 | Graphic summary and a mitochondrial BER model. (a) Structural basis for hEXO activity. **(b)** A proposed model for mitochondrial DNA long-patch (LP) BER and its difference from the nuclear LP-BER. In mitochondria, EXOG exerts 5'-exonuclease activity at the abasic site and generates a 2-3 nt gap on dsDNA, which is an optimal substrate for Poly to perform gap-filling synthesis. In contrast, nuclear LP-BER depends on the strand displacement activity of the DNA polymerase and FEN1 cleaves the DNA flap.

While allostery provides an explanation for the observed half-site reactivity, it leaves open the question why hEXO is a homodimer, as only one active site is necessary for BER. A possible function for the second Wing domain is to interact with other repair enzymes. Protein-protein interactions are thought to enhance repair efficiency by assuring transfer of DNA intermediates between enzymes in the repair pathway and preventing inappropriate DNA transactions^{8,9}.

The activities of hEXO that we have characterized strongly suggest that mtDNA repair is performed via a LP-BER pathway. However, the reactions in the pathway are different from nuclear LP-BER^{10,32}. The two pathways diverge after the generation of a gapped DNA by APE1 (Fig. 7b). In nuclear LP-BER, DNA synthesis occurs before the nuclease reaction; Pol δ synthesizes DNA across the lesion site and displaces the downstream strand. FEN1 then removes the DNA flap before ligation. Conversely, in mitochondrial LP-BER, the nuclease reaction occurs before DNA synthesis: hEXO processes the 5'-end of DNA downstream of the lesion, generating an optimal substrate for DNA synthesis catalysed by Pol γ . Mitochondrial BER thus successfully circumvents the lack of strand-displacement activity by Pol γ and its inefficiency in synthesizing at single-nucleotide gaps. This model also explains why FEN1 activity is dispensable in the mitochondria^{14,33}.

The structural and functional studies presented here provide a missing piece of the puzzle towards our understanding the mechanism of mitochondrial BER. The LP-BER pathway we propose allows DNA repair to proceed through a well-controlled 'cut-and-fill' mechanism that generates an optimal substrate for Pol γ without the need for strand-displacement and without requiring flap-endonuclease activity.

Methods

Cloning and purification of hEXO. hEXO was cloned in pET22b without N-terminal residues 2-58 and with a C-terminal His-tag. This clone served as the parental construct for the H140A and R314A substitutions and for the C-terminal 68 aa deletion (hEXO- Δ C68). All hEXO variants were expressed in *Escherichia coli* ROSETTA (DE3) (Novagen). Cells were inoculated with 0.1% of an overnight culture in LB media containing ampicillin ($100 \mu\text{g ml}^{-1}$) and chloramphenicol ($34 \mu\text{g ml}^{-1}$) at 37°C to $\text{OD}_{600} = 0.4$. Isopropyl- β -D-thiogalactoside (1 mM) was then added and the culture was incubated for 12 h at 25°C . All proteins were purified using a modification of a previous protocol³⁴. Briefly, the insoluble hEXO

in inclusion bodies was solubilized in denaturing buffer (10 mM Tris, pH 8.2, 200 mM NaCl, 20 mM B-ME, 10 mM imidazole and 8 M urea). The unfolded protein was purified on a Ni-NTA affinity column. The eluted hEXO was flash diluted in denaturing buffer to concentration of 1-2 μM , and then refolded, first by dialysing against denaturing buffer where 8 M urea was substituted with 1.5 M sorbitol and then against GF buffer (20 mM HEPES, pH 8.0, 300 mM NaCl, 1 mM TCEP) with 5% glycerol (v/v). The refolded hEXO was loaded onto a Hi-Load Superdex 200 gel filtration column (GE Healthcare Life Sciences), and peak fractions were concentrated using a Centricon concentrator (Millipore) to $\sim 7 \text{ mg ml}^{-1}$. The enzyme used for activity assays was flash-frozen in GF buffer with 40% glycerol in liquid nitrogen and stored at -80°C . Before the experiments, the protein was dialysed against buffer EK140 (20 mM HEPES (pH 7.5), 140 mM KCl, 1 mM TCEP and 5% (v/v) glycerol) and the concentration was determined spectrophotometrically using extinction coefficient $\epsilon_{280(\text{dimer})} = 77,030 \text{ M}^{-1} \text{ cm}^{-1}$, for wild type, R314A and H140A, and $\epsilon_{280(\text{dimer})} = 68,090 \text{ M}^{-1} \text{ cm}^{-1}$ for hEXO- Δ C68.

Circular dichroism spectroscopy. Unfolded hEXO-H140A ($\sim 2 \mu\text{M}$) was used as the baseline for monitoring folding; the protein was the Ni-NTA eluant dialysed against 20 mM Na_2HPO_4 (pH 7.5) containing 8 M urea to remove the imidazole. After testing various folding procedures, hEXO wild type, H140A or Δ C68 variants ($2 \mu\text{M}$) was dialysed against 20 mM Na_2HPO_4 (pH 7.5) with 140 mM KCl overnight at 4°C . CD spectra (195-260 nm) were collected on a Jasco J-815 CD Spectrometer at 20°C . Each reported CD curve is an average of three scans, corrected for the contribution from the buffer alone.

Small-angle X-ray scattering. SAXS data were collected using a Rigaku BioSAXS-1000 with an FRED++ Copper x-ray source. Exposure times were 8, 12 and 16 h at each apo hEXO concentration, and 12, 14 and 16 h at each DNA-bound hEXO concentrations, with matching times for their buffers. Hourly partial scans were visually examined for signs of radiation damage and the χ^2 -squared metric used to ascertain when damage began (<http://xray.utmb.edu/saxns-XCS.html>). Images were processed in SAXLab (Rigaku) and buffer subtraction performed with the SAXNS_ES web server (<http://www.xray.utmb.edu/saxns-es.html>). The apo hEXO data could be merged from the 1.0 and 0.75 mg ml^{-1} samples based on their XCS values. The hEXO-DNA sample data from the 6, 4 and 2 mg ml^{-1} data were merged, using cross- χ^2 as a guide to exclude regions with problems due to near-beamstop parasitic scatter or a poor signal to noise ratio. All data were examined using PRIMUS, $P(r)$ curves generated with GNOM, and ten molecular bead models generated with DAMMIF then aligned and filtered with DAMAVER, all from the ATSAS suite³⁵. Crystal structures were compared to the SAXS data using CRYSOLO and rigid-body models refined using CORAL³⁵.

Preparation of oligonucleotide substrates. Oligonucleotide containing 3'-fluorescein (Midland Certified Reagent Co) (Supplementary Table 2) or 5'-³²P, prepared using T4 kinase and [γ -³²P-ATP], was annealed with the complementary strand by heating at 95°C for 5 min in a solution containing 20 mM Tris, pH 8.1, 100 mM NaCl, 10% (v/v) glycerol and slowly cooled to room temperature.

Isothermal titration calorimetry. hEXOg–H140A and DNA were dialysed separately against EK140 buffer at 4 °C. Titrations were conducted using a VP-ITC calorimeter (Microcal, GE Healthcare) at 20 °C with 23 injections (1 µl for the initial injection, followed by five injections at 5 µl, eight injections at 10 and 15 µl for the subsequent injections) of 60 µM DNA titrant into 1.437 ml 2.8–3.2 µM hEXOg–H140A, with 360 s between injections. Reference titrations were performed without enzyme. The first injection was excluded from the analysis. Thermodynamic binding parameters were derived from fitting isotherm to two sequential binding sites model, using SEDPHAT³⁶.

Endonuclease activity assay. hEXOg wild type or mutants were incubated with circular ssDNA (100 ng µl⁻¹ φX174 virion DNA) in EK140 buffer for 10 min at 25 °C. Reactions were initiated by the addition of MgCl₂ to 5 mM, and incubated for 1 h at 25 °C using varying enzyme concentrations. Reactions were stopped by the addition of stop BF (50 mM EDTA, 0.1% SDS, 80% formamide). Cleavage products were resolved by electrophoresis in a 1% agarose gel, stained with ethidium bromide, and quantified using ImageJ (<http://rsb.info.nih.gov/ij/>). Substrate remaining was quantified with the initial concentration set at 100%.

Exonuclease activity assay. hEXOg wild type or a variant was pre-incubated with MgCl₂ (5 mM) for 10 min at room temperature in EK140 buffer. The reaction was initiated by the addition of DNA substrate; enzyme and DNA concentrations are given with the figures. All experiments were performed at 25 °C. Aliquots were taken and mixed with an equal volume of stop BF and incubated at 95 °C for 4 min. Reaction products were separated using denaturing 23% polyacrylamide/7 M urea polyacrylamide gel electrophoresis and visualized using a Molecular Dynamics Storm 860 Phosphorimager. Product bands were quantified using ImageQuant (GE Healthcare). The time dependence of product formation is fitted to the equation [product] = A₀ + k_{ss} t, where A₀ is amplitude, k_{ss} is steady-state turnover rate and t is time. The burst product is approximated to A₀. The off rate k_{off} = k_{ss}/A₀. Reactions were stopped with stop BF and incubated at 95 °C for 4 min. Error bars represent ± s.d. from three parallel experiments.

DNA-binding assays. Substrate 10 bp DNA (10 nM) containing a 5'-P and a 3'-fluorescein (Supplementary Table 1) was mixed with increasing concentrations of hEXOg wild type or R314A mutant in EK140 buffer plus 20 mM EDTA. To determine the affinity of hEXOg to 5'-P and 5'-OH on DNA, the same 3'-F-labelled DNA was synthesized containing a phosphate or a hydroxyl group. The measurements were taken on a ISS PC1 photon counting spectrofluorimeter by monitoring the emission of the F-labelled nucleic acid (λ_{ex} = 480 nm and λ_{em} = 520 nm) as described previously and the experimental titrations were fit to single-site binding isotherm³⁷. Electrophoretic mobility shift assay was performed with 10 nM 5'-³²P-labelled 10 bp DNA mixed with increasing concentrations of hEXOg wild type or R314A mutant in EK140 buffer plus 20 mM EDTA. The mixtures were resolved on a 10% acrylamide gel at 4 °C in × 0.5 TBE supplemented with 20 mM EDTA. Gels were scanned and quantified as above. The fraction of DNA bound was calculated after background correction. Apparent K_d values were estimated by fitting using Mathematica.

Crystallization and data collection. Apo hEXOg crystals were obtained by the hanging-drop method at 19 °C using 5.6 mg ml⁻¹ enzyme with 5 mM MgCl₂ against a well solution containing 2.0 M (NH₄)₂SO₄, 0.1 M MgCl₂ and 0.1 M Tris-HCl pH 8.5. Crystals containing Mn²⁺ were obtained by soaking hEXOg-Mg²⁺ crystals in mother liquor containing 100–200 mM MnCl₂ for 3–24 h. Complex crystals of hEXOg–H140A with Mg²⁺ and DNA were obtained using the sitting-drop method at 17 °C using hEXOg–H140A (6–7.0 mg ml⁻¹), a 2.3 M ratio of DNA and 10 mM MgCl₂ against a well solution containing 22–28% PEG 2K MME, 0.2–0.25 M (NH₄)₂SO₄ and 0.1 M Na acetate pH 4.6. Crystals with Mn²⁺ were obtained by soaking EXOG–H140A:Mg²⁺:dsDNA complex crystals in mother liquor with 75–150 mM MnCl₂ for 5–24 h. All crystals were cryo-protected by stepwise increasing concentrations of glycerol and MPD in mother liquor until both reached 15%. They were then flash-frozen in liquid nitrogen. Single-wavelength anomalous diffraction data were collected using the hEXOg apo–Mn²⁺ or DNA–Mn²⁺ complex crystals at an energy level of 6.539 keV—the K-edge of Mn²⁺ anomalous scattering. Data were processed using HKL2000 (ref. 38). The difference anomalous map calculated from the single-wavelength anomalous diffraction data shows peaks above 5-sigma at the same position as the original Mg²⁺ ions, confirming their replacement with Mn²⁺ (Supplementary Fig. 3b,c). This suggests that the metal ion in the original crystals contain Mg²⁺ at the same locations.

Structure determination. The crystal structure of hEXOg:Mg²⁺ was determined by molecular replacement using Phaser and *Drosophila* EndoG (PDB accession number 3ISM) as search model³⁹. The missing Wing domain was built using Coot⁴⁰, and refined using Phenix⁴¹. Crystal structures of hEXOg–DNA complexes were determined by molecular replacement using the apo EXOG structure. MolProbity⁴² was used to evaluate the final structures. Single-wavelength anomalous dispersion data were collected for apo and DNA complex containing Mn²⁺. The crystal structures of hEXOg:Mn²⁺ and hEXOg:DNA:Mn²⁺ were

determined by molecular replacement using hEXOg:Mg²⁺ and hEXOg–DNA structure as respective models. Metal locations were determined from an anomalous difference map. After refinement, Ramachandran favoured residues were ≥98% and Ramachandran outlier residues were ≤0.57% for all four structures. All structural figures were prepared using PyMol.

Data availability. Coordinates and structure factor of the structure reported here have been deposited into the Protein Data Bank with PDB codes: 5T40, 5T3V, 5T5C and 5T4I. All additional experimental data are available from the corresponding author on request.

References

- Brown, W. M., George, Jr M. & Wilson, A. C. Rapid evolution of animal mitochondrial DNA. *Proc. Natl Acad. Sci. USA* **76**, 1967–1971 (1979).
- Haag-Liautard, C. *et al.* Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biol.* **6**, e204 (2008).
- Bohr, V. A., Stevnsner, T. & de Souza-Pinto, N. C. Mitochondrial DNA repair of oxidative damage in mammalian cells. *Gene* **286**, 127–134 (2002).
- Greaves, L. C. *et al.* Comparison of mitochondrial mutation spectra in ageing human colonic epithelium and disease: absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. *PLoS Genet.* **8**, e1003082 (2012).
- Young, M. J. & Copeland, W. C. Human mitochondrial DNA replication machinery and disease. *Curr. Opin. Genet. Dev.* **38**, 52–62 (2016).
- Kazak, L., Reyes, A. & Holt, I. J. Minimizing the damage: repair pathways keep mitochondrial DNA intact. *Nat. Rev. Mol. Cell Biol.* **13**, 659–671 (2012).
- Mol, C. D., Izumi, T., Mitra, S. & Tainer, J. A. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. *Nature* **403**, 451–456 (2000).
- Prasad, R., Shock, D. D., Beard, W. A. & Wilson, S. H. Substrate channeling in mammalian base excision repair pathways: passing the baton. *J. Biol. Chem.* **285**, 40479–40488 (2010).
- Wilson, S. H. & Kunkel, T. A. Passing the baton in base excision repair. *Nat. Struct. Biol.* **7**, 176–178 (2000).
- Krokan, H. E. & Bjoras, M. Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583 (2013).
- Longley, M. J., Ropp, P. A., Lim, S. E. & Copeland, W. C. Characterization of the native and recombinant catalytic subunit of human DNA polymerase gamma: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. *Biochemistry* **37**, 10529–10539 (1998).
- Longley, M. J., Prasad, R., Srivastava, D. K., Wilson, S. H. & Copeland, W. C. Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro. *Proc. Natl Acad. Sci. USA* **95**, 12244–12248 (1998).
- He, Q., Shumate, C. K., White, M. A., Molineux, I. J. & Yin, Y. W. Exonuclease of human DNA polymerase gamma disengages its strand displacement function. *Mitochondrion* **13**, 592–601 (2013).
- Tann, A. W. *et al.* Apoptosis induced by persistent single-strand breaks in mitochondrial genome: critical role of EXOG (5'-EXO/endonuclease) in their repair. *J. Biol. Chem.* **286**, 31975–31983 (2011).
- Szczesny, B. *et al.* Deficiency in repair of the mitochondrial genome sensitizes proliferating myoblasts to oxidative damage. *PLoS ONE* **8**, e75201 (2013).
- Tighelehar, W. *et al.* Loss of mitochondrial exo/endonuclease EXOG affects mitochondrial respiration and induces ROS-mediated cardiomyocyte hypertrophy. *Am. J. Physiol. Cell Physiol.* **308**, C155–C163 (2015).
- Szczesny, B., Brunyanski, A., Olah, G., Mitra, S. & Szabo, C. Opposing roles of mitochondrial and nuclear PARP1 in the regulation of mitochondrial and nuclear DNA integrity: implications for the regulation of mitochondrial function. *Nucleic Acids Res.* **42**, 13161–13173 (2014).
- Yang, W. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* **44**, 1–93 (2011).
- Cymerman, I. A., Chung, I., Beckmann, B. M., Bujnicki, J. M. & Meiss, G. EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.* **36**, 1369–1379 (2008).
- Parrish, J. *et al.* Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature* **412**, 90–94 (2001).
- Zhang, J. *et al.* Endonuclease G is required for early embryogenesis and normal apoptosis in mice. *Proc. Natl Acad. Sci. USA* **100**, 15782–15787 (2003).
- Loll, B., Gebhardt, M., Wahle, E. & Meinhardt, A. Crystal structure of the EndoG/EndoG1 complex: mechanism of EndoG inhibition. *Nucleic Acids Res.* **37**, 7312–7320 (2009).
- Shlyapnikov, S. V. *et al.* Atomic structure of the Serratia marcescens endonuclease at 1.1 Å resolution and the enzyme reaction mechanism. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 567–572 (2000).
- Ghosh, M., Meiss, G., Pingoud, A., London, R. E. & Pedersen, L. C. Structural insights into the mechanism of nuclease A, a betabeta alpha metal nuclease from *Anabaena*. *J. Biol. Chem.* **280**, 27990–27997 (2005).

25. Petoukhov, M. V. *et al.* New developments in the program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
26. Tsutakawa, S. E., Lafrance-Vanasse, J. & Tainer, J. A. The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. *DNA Repair* **19**, 95–107 (2014).
27. Sung, J. S. & Dimple, B. Analysis of base excision DNA repair of the oxidative lesion 2-deoxyribonolactone and the formation of DNA-protein cross-links. *Methods Enzymol.* **408**, 48–64 (2006).
28. Balakrishnan, L. & Bambara, R. A. Flap endonuclease 1. *Ann. Rev. Biochem.* **82**, 119–138 (2013).
29. Zheng, L. *et al.* Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. *Mol Cell* **32**, 325–336 (2008).
30. Kornblum, C. *et al.* Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nat. Genet.* **45**, 214–219 (2013).
31. Uhler, J. P. *et al.* MGME1 processes flaps into ligatable nicks in concert with DNA polymerase gamma during mtDNA replication. *Nucleic Acids Res.* **44**, 5861–5871 (2016).
32. Dianov, G. L. *et al.* Base excision repair in nuclear and mitochondrial DNA. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 285–297 (2001).
33. Szczesny, B., Tann, A. W., Longley, M. J., Copeland, W. C. & Mitra, S. Long patch base excision repair in mammalian mitochondrial genomes. *J. Biol. Chem.* **283**, 26349–26356 (2008).
34. Kieper, J. *et al.* Production and characterization of recombinant protein preparations of Endonuclease G-homologs from yeast, *C. elegans* and humans. *Protein Expr. Purif.* **73**, 99–106 (2010).
35. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
36. Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT—a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137–148 (2015).
37. Szymanski, M. R., Jezewska, M. J. & Bujalowski, W. The *Escherichia coli* PriA helicase specifically recognizes gapped DNA substrates: effect of the two nucleotide-binding sites of the enzyme on the recognition process. *J. Biol. Chem.* **285**, 9683–9696 (2010).
38. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307–326 (1997).
39. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica Section D, Biological crystallography* **66**, 486–501 (2010).
41. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
42. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

Acknowledgements

We thank C. Shumate for critical reading of the manuscript, Seipo Co. for artistic rendering, B. Szczesny and S. Mitra for the original clone of hEXOG. We thank W. Bujalowski for the usage of ISS PC1 photon counting spectrofluorimeter and helpful discussions. We thank staff at the beamlines 19-ID, 19-BM and 21-ID at the Advanced Photon Source, Argonne National Laboratory, and beamlines 8.2.1 and 8.2.2 at the Advanced Light Source, Lawrence Berkeley National Laboratory. M.R.S. was a J.B. Kempner Postdoctoral Fellow. A.M.G. and J.C.L. are supported by an endowment (H-0013) from the Welch Foundation. The work was supported by grants from NIH (GM 083703 and GM110591) to Y.W.Y., and an endowment from Sealy and Smith Foundation to Sealy Center for Structural Biology at UTMB.

Author contributions

M.R.S., J.C.L. and Y.W.Y. conceived the project and designed the experiments. M.R.S., W.Y., A.M.G., M.A.W. and Y.W.Y. performed the experiments. M.R.S., I.J.M. and Y.W.Y. wrote the manuscript; all authors provided editorial suggestions and criticisms.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing interests: The authors declare no competing financial interests.

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

How to cite this article: Szymanski, M. R. *et al.* A domain in human EXOG converts apoptotic endonuclease to DNA-repair exonuclease. *Nat. Commun.* **8**, 14959 doi: 10.1038/ncomms14959 (2017).

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



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017