

RESEARCH ARTICLE

# Variable termination sites of DNA polymerases encountering a DNA–protein cross-link

Anna V. Yudkina<sup>1,2</sup>, Antonina P. Dvornikova<sup>1</sup>, Dmitry O. Zharkov<sup>1,2\*</sup>

**1** Laboratory of Genome and Protein Engineering, Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia, **2** Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia

\* [dzharkov@niboch.nsc.ru](mailto:dzharkov@niboch.nsc.ru)



**OPEN ACCESS**

**Citation:** Yudkina AV, Dvornikova AP, Zharkov DO (2018) Variable termination sites of DNA polymerases encountering a DNA–protein cross-link. PLoS ONE 13(6): e0198480. <https://doi.org/10.1371/journal.pone.0198480>

**Editor:** Giovanni Maga, Istituto di Genetica Molecolare, ITALY

**Received:** March 9, 2018

**Accepted:** May 18, 2018

**Published:** June 1, 2018

**Copyright:** © 2018 Yudkina et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was supported by Russian Science Foundation (17-14-01190). Partial salary support from the Russian State funded 2013–2020 budget project (VI.62.1.5, 0309-2016-0003) and Russian Ministry of Science and Education (6.5773.2017/VU) is acknowledged. The funders had no involvement in any part of the study. The funders had no role in study design, data collection

## Abstract

DNA-protein cross-links (DPCs) are important DNA lesions induced by endogenous cross-linking agents such as formaldehyde or acetaldehyde, as well as ionizing radiation, cancer chemotherapeutic drugs, and abortive action of some enzymes. Due to their very bulky nature, they are expected to interfere with DNA and RNA synthesis and DNA repair. DPCs are highly genotoxic and the ability of cells to deal with them is relevant for many chemotherapeutic interventions. However, interactions of DNA polymerases with DPCs have been poorly studied due to the lack of a convenient experimental model. We have used NaBH<sub>4</sub>-induced trapping of *E. coli* formamidopyrimidine–DNA glycosylase with DNA to construct model DNA polymerase substrates containing a DPC in single-stranded template, or in the template strand of double-stranded DNA, or in the non-template (displaced) strand of double-stranded DNA. Nine DNA polymerases belonging to families A, B, X, and Y were studied with respect to their behavior upon encountering a DPC: Klenow fragment of *E. coli* DNA polymerase I, *Thermus aquaticus* DNA polymerase I, *Pyrococcus furiosus* DNA polymerase, *Sulfolobus solfataricus* DNA polymerase IV, human DNA polymerases β, κ and λ, and DNA polymerases from bacteriophages T4 and RB69. Although none were able to fully bypass DPCs in any context, Family B DNA polymerases (T4, RB69) and Family Y DNA polymerase IV were able to elongate the primer up to the site of the cross-link if a DPC was located in single-stranded template or in the displaced strand. In other cases, DNA synthesis stopped 4–5 nucleotides before the site of the cross-link in single-stranded template or in double-stranded DNA if the polymerases could displace the downstream strand. We suggest that termination of DNA polymerases on a DPC is mostly due to the unrelieved conformational strain experienced by the enzyme when pressing against the cross-linked protein molecule.

and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

DNA of all living organisms is perpetually exposed to various exogenous and endogenous genotoxic agents, and is subject to damage [1]. DNA lesions can induce mutations, which lead to cancer and contribute to aging. Since cellular DNA is tightly bound to a variety of structural, regulatory and catalytic proteins, DNA-protein cross-links (DPCs) are among the common types of DNA damage. They can be generated by genotoxic agents such as aldehydes, ionizing and UV radiation, oxidative stress, and some chemotherapy drugs, or through covalent capture of some DNA-dependent enzymes (methyltransferases, topoisomerases) in abortive reaction events [2–4]. Different analytical methods produce estimates of 0.5–70 DPCs per  $10^7$  DNA bases as a background level in human cells [5–7].

DPCs as very bulky adducts are expected to interfere with DNA replication and transcription. However, the impact of this type of damage on the genetic machinery remains poorly studied due to the scarcity of suitable experimental models and differences in the emergence mechanisms. The most basic question regarding the genotoxic effects of DPCs is how DNA polymerases behave when encountering these adducts. In a few available model *in vitro* systems DPCs are completely blocking [8–10]. However, at least some DNA polymerases can bypass cross-links with peptides up to 12-amino acids long in an error-free manner [10–13] or with skipping of one or two nucleotides [13].

Known DNA polymerases all share a similar structural fold, consisting of three domains: palm, thumb, and fingers. Based on their sequence and structure similarity and the presence of additional domains, they are currently divided into seven families (A, B, C, D, X, Y and RT) [14–16]. Family A polymerases, best represented by *E. coli* DNA polymerase I, are bacterial enzymes responsible for DNA repair synthesis and replacement of Okazaki fragment-initiating primers during replication. Some replicative phage and eukaryotic mitochondrial polymerases also belong here (T7 DNA polymerase, DNA polymerase  $\gamma$ ). Family B mostly consists of replicative archaeal, viral and eukaryotic polymerases (DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$ , *Pfu* DNA polymerase, DNA polymerases from bacteriophages T4 and RB69). Bacterial replicative polymerases, e. g., *E. coli* DNA polymerase III, make up Family C, while a subset of archaeal replicative polymerases is known as Family D. Eukaryotic DNA repair polymerases such as DNA polymerases  $\beta$  and  $\lambda$  form Family X. Family Y covers specialized polymerases from all kingdoms of life, responsible for translesion synthesis and DNA damage tolerance; due to their wide active sites, these polymerases can accommodate many non-canonical nucleotides, including bulky ones, and catalyze error-prone dNTP incorporation. Finally, reverse transcriptases, which can use RNA templates, are grouped in their own Family RT.

Biologically, DNA polymerases would directly encounter DPCs during replication only under special circumstances, since DNA helicases that separate the DNA strands in the replication fork would presumably run into the cross-link first and arrest the fork movement. However, if a DNA helicase travels on the undamaged leading strand and can unwind the DPC, the lagging strand synthesis will run into the cross-link [17]. In other biologically relevant situations such as DNA repair of complex lesions (for example, those induced by ionizing radiation) DNA polymerases could also interact with DPCs. In the analytical setting, DPCs may impede amplification of ancient DNA [18] or DNA from formalin-fixed tissues [19] by polymerase chain reaction. Therefore, it is essential to know how various DNA polymerases would behave when colliding with a DPC. Importantly, all previous studies employed a simple primer/template system with a cross-link in the single-stranded template; no data is available on the situation when DNA synthesis is coupled with the downstream strand displacement, with a DPC in either template or displaced strand of the duplex.

In this paper, we use a fully chemically defined model DPC to investigate the events after its encounter with a variety of DNA polymerases. We report that most DNA polymerases stop when coming into contact with the protein part of a DPC, but some, especially those belonging to Family B, can partially displace the cross-linked protein and continue synthesis even to the point of the cross-link.

## Materials and methods

### Oligonucleotides and enzymes

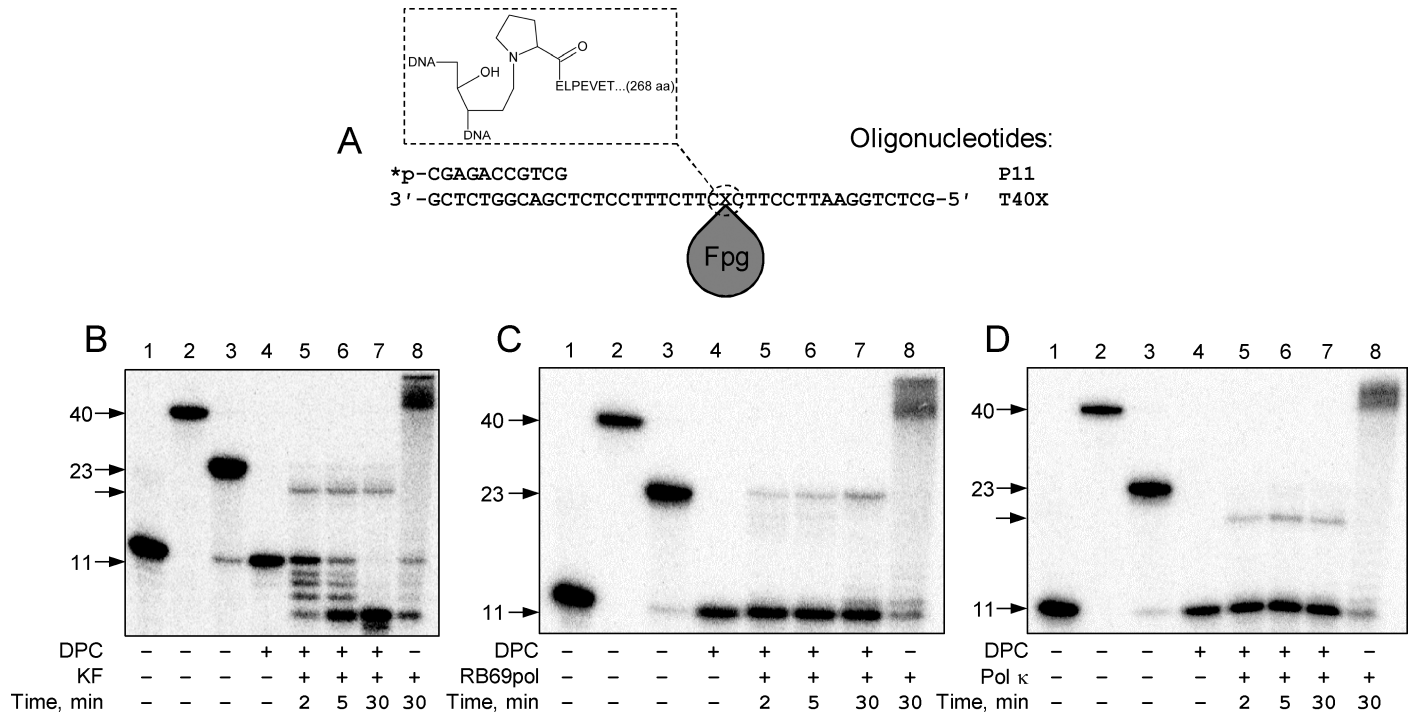
Wild-type Klenow fragment of *E. coli* DNA polymerase I (KF) and phage T4 DNA polymerase were purchased from SibEnzyme (Novosibirsk, Russia). *Taq* and *Pfu* DNA polymerases were from Thermo Scientific (Waltham, MA). *Sulfolobus solfataricus* Dpo4 DNA polymerase was from New England Biolabs (Ipswich, MA). *E. coli* Fpg, exonuclease-deficient Klenow fragment of *E. coli* DNA polymerase I (KF  $exo^-$ ), and bacteriophage RB69 DNA polymerase (RB69pol) were overexpressed and purified essentially as described [20–22]. Human DNA polymerases  $\beta$  and  $\lambda$  and human FEN1 protein were kindly provided by Dr. Olga Lavrik (SB RAS ICBFM). Oligonucleotides (S1 Table) were synthesized from commercially available phosphoramidites (Glen Research, Sterling, VA) and purified by reverse-phase HPLC on a PRP-1 C<sub>18</sub> column (Hamilton, Reno, NV). The oligonucleotides were 5'-labeled using  $\gamma$ [<sup>32</sup>P]-ATP (PerkinElmer, Waltham, MA) and phage T4 polynucleotide kinase (Biosan, Novosibirsk, Russia) according to the manufacturer's protocol.

### Model cross-link synthesis

Oligonucleotide duplexes were assembled as shown in Figs 1–3 but without the <sup>32</sup>P-labeled primer. Two model non-labeled substrates were obtained, one containing oxoG in the template strand, another, in the displaced strand. The reaction mixture contained 50 mM sodium phosphate (pH 6.8), 1 mM ethylenediaminetetraacetate (EDTA), 1 mM dithiothreitol, and 30 pmol of oligonucleotide duplex. A 20-fold molar excess of Fpg and 100 mM NaBH<sub>4</sub> were added to the reaction mixture simultaneously. The reaction was allowed to proceed at 37°C for 30 min and stopped by adding glucose to 400 mM and incubating for 30 min on ice. The efficiency of cross-linking was 65–95%, as determined by SDS-PAGE. To make a cross-link to the single-stranded template, the reaction was carried out for 60 min under the same conditions using the template strand only. The cross-linking efficiency in this case was typically ~25%. The DNA–Fpg conjugate was purified by 8% non-denaturing PAGE, the band of interest was excised and eluted with 0.1×TE buffer overnight, and the eluate was filtered using the Ultra-free-CL device with 0.22  $\mu$ m pore diameter (Merck Millipore, Billerica, MA). Purified DPC-containing DNA was concentrated on a Centricon-10 ultrafiltration unit (Merck Millipore). The <sup>32</sup>P-labeled primer (P11, Table 1) was added and incubated for 1 h at 37°C to complete the assembly of the complex. As controls, we have used primer–template or primer–downstream strand–template constructs without the damaged base, not subjected to the cross-linking procedure.

### Primer extension reactions

The reaction mixture (20  $\mu$ l) contained 500  $\mu$ M dNTP mixture (with equimolar ratio of all four dNTP), 400 fmol of a DPC substrate (see above) and the required DNA polymerase in the respective optimal buffer: 50 mM Tris–HCl (pH 7.5), 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 30 mM KCl; 0.1 mM DTT for KF and KF  $exo^-$ ; 20 mM Tris–HCl (pH 8.8), 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100 for *Pfu*, 10 mM Tris–HCl (pH 8.8), 50 mM KCl, 1.5 mM



**Fig 1. Modes of termination of DNA polymerases at DNA-protein cross-links in the template strand of single-stranded DNA.** (A) Scheme of the DPC substrate. The chemical structure of the reduced Schiff base at the site of the cross-link through Pro1 of Fpg is shown in the inset. \*p, <sup>32</sup>P-labeled 5'-end of the primer. (B) Synthesis stops before the cross-link site with the primer degradation: *E. coli* DNA polymerase I Klenow fragment (Family A). (C) Synthesis extends to the cross-link site: phage RB69 DNA polymerase (Family B). (D) Synthesis stops before the cross-link site: human DNA polymerase κ (Family Y). Lanes 1–3, size markers corresponding to the primer (lane 1, 11 nt long), full-size product (lane 2, 40 nt), and primer extended to the cross-link site (lane 3, 23 nt). Lane 4, the substrate in the absence of DNA polymerase, lanes 5–8, extension by the DNA polymerase for the indicated time. In lane 8, the reaction was carried out with a primer-undamaged template substrate.

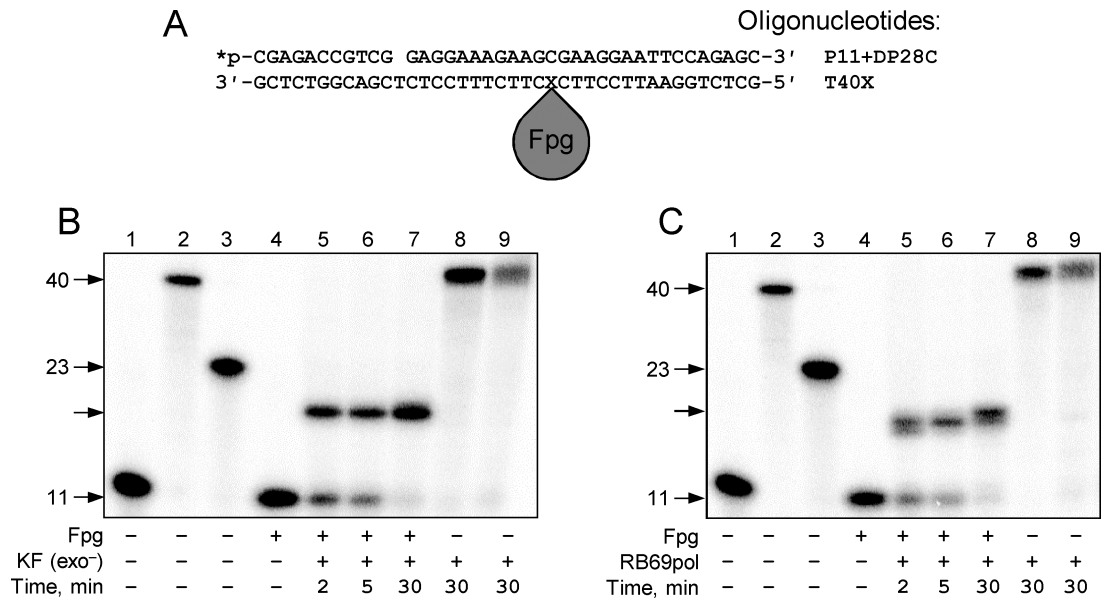
<https://doi.org/10.1371/journal.pone.0198480.g001>

MgCl<sub>2</sub>, 0.08% Nonidet-P40 for *Taq*; 10 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 1 mM DTT for Pol β; 50 mM Tris-HCl (pH 8.0), 0.5 mM MgCl<sub>2</sub>, 0.5 mM DTT for Pol λ; 50 mM Tris-HCl (pH 7.5), 5 mM MgCl<sub>2</sub>, 0.1 mM DTT for RB69pol; 67 mM Tris-HCl (pH 8.8), 6.7 mM MgCl<sub>2</sub>, 16.7 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM DTT for T4pol; 50 mM HEPES (pH 7.5), 5 mM MgCl<sub>2</sub>, 60 mM KCl, 5 mM DTT for Pol κ; 20 mM Tris-HCl (pH 8.8), 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100 for Dpo4. The buffers were selected based on the manufacturers' instructions and literature data. FEN1 was added to 400 nM, if necessary. The quantity of the DNA polymerase was optimized to give ~70% insertion of the first nucleotide in 2 min. The reaction was allowed to proceed at 37°C for 2, 5, or 30 min. At these times, aliquots were withdrawn, an equal volume of formamide dye solution (80% formamide, 20 mM Na-EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue) was added followed by 2-min heating at 95°C, and the reaction products were resolved by 20% denaturing PAGE and visualized by phosphorimaging (Typhoon FLA 9500, GE Healthcare, Chicago, IL). The lengths of extension products were determined from comparisons with the mobility markers and partial extension ladders.

## Results

### Model cross-links

To investigate bypass of DPCs by DNA polymerases belonging to different families, we have covalently cross-linked *E. coli* formamidopyrimidine-DNA glycosylase (Fpg) to oligonucleotide substrates containing 8-oxoguanine (oxoG) in the presence of NaBH<sub>4</sub> (Figs 1–3, Panel A).



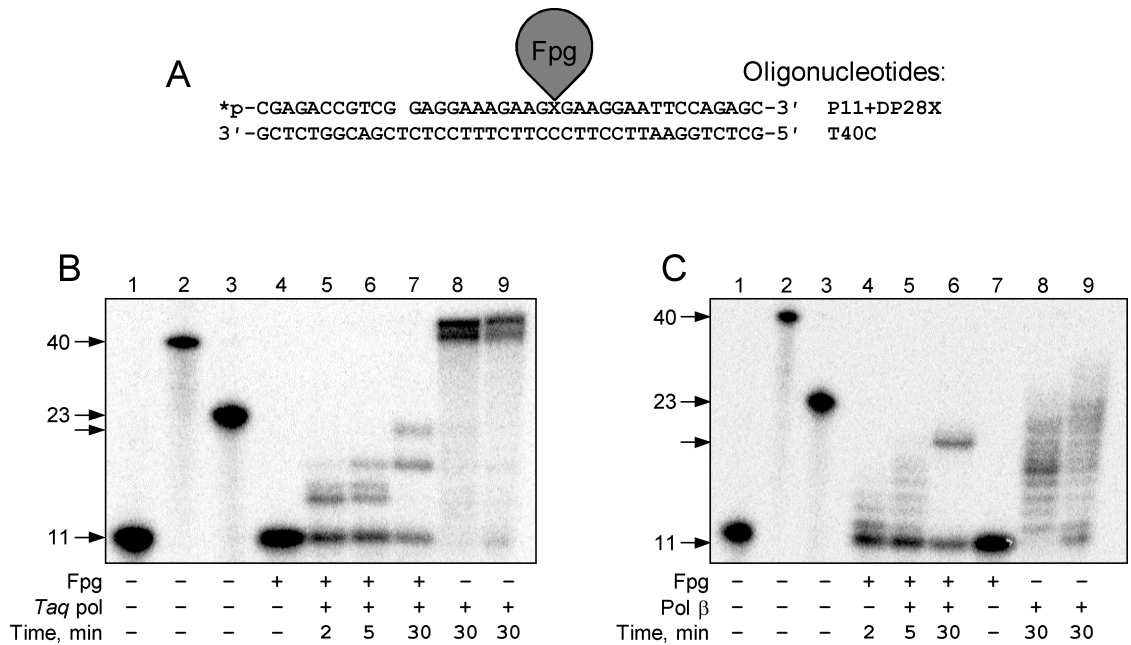
**Fig 2. Modes of termination of DNA polymerases at DNA-protein cross-links in the template strand of double-stranded DNA.** (A) Scheme of the DPC substrate. (B) Synthesis stops before the cross-link site: KF exo<sup>-</sup> (Family A). (C) Synthesis pauses before and stops closer to the cross-link site: phage RB69 DNA polymerase (Family B). Lanes 1–3, size markers corresponding to the primer (lane 1, 11 nt long), full-size product (lane 2, 40 nt), and primer extended to the cross-link site (lane 3, 23 nt). Lane 4, the substrate in the absence of DNA polymerase, lanes 5–9, extension by the DNA polymerase for the indicated time. In lane 8, the reaction was carried out with a primer-undamaged template substrate, and in lane 9, with a primer-downstream strand-undamaged template substrate.

<https://doi.org/10.1371/journal.pone.0198480.g002>

Since DPCs may be formed with a variety of proteins, we have reasoned that there is little specificity in the interaction of DNA polymerases with them, and chose the model based on the easiness of its synthesis. Fpg forms a Schiff base between the enzyme's active site amino group and C1' of the damaged nucleoside as a reaction intermediate, which can be reduced to a stable covalent amine [23] (Fig 1A, inset), a reaction that was used to prepare Fpg-DNA conjugate in large quantities for crystallization [20]. Here, Fpg was efficiently cross-linked to double-stranded DNA containing oxoG in either the template strand or the displaced strand (65–95% yield under the conditions described in Materials and Methods). Single-stranded DNA is a much worse substrate for Fpg due to both lower affinity and lower catalytic constant; however, at a large enzyme excess, the reaction can be pushed to a reasonable depth [24]. In our experiments, we obtained a ~25% yield of the cross-link with oxoG in the single-stranded template. The cross-links were purified by non-denaturing gel electrophoresis and used in the bypass studies. The control substrates were assembled with the non-damaged downstream primer DP28C and template T40G (S1 Table).

### Primer extension with DPCs in the single-stranded template by DNA polymerases of various families

Family A DNA polymerases are involved in replication of some bacteriophages (e. g., phage T7 DNA polymerase), replication of mitochondrial DNA in eukaryotes (DNA polymerase  $\gamma$ ), and in DNA repair and processing of Okazaki fragments in bacteria (DNA polymerase I) [25]. Interestingly, the Klenow fragment of DNA polymerase I (KF) have been reported to bypass dodecapeptide adducts to N<sup>6</sup> of adenine in an error-free manner [26]. Generally, DNA synthesis on DPC-containing primer-template substrates by family A DNA polymerases can be



**Fig 3. Modes of termination of DNA polymerases at DNA-protein cross-links in the displaced strand of double-stranded DNA.** (A) Scheme of the DPC substrate. (B) Synthesis with processive strand displacement with a stop before the cross-link site: *Taq* DNA polymerase (Family A). (C) Synthesis with hit-and-run strand displacement with a stop before the cross-link site: human DNA polymerase β (Family X). Lanes 1–3, size markers corresponding to the primer (lane 1, 11 nt long), full-size product (lane 2, 40 nt), and primer extended to the cross-link site (lane 3, 23 nt). Lane 4, the substrate in the absence of DNA polymerase, lanes 5–9, extension by the DNA polymerase for the indicated time. In lane 8, the reaction was carried out with a primer–template substrate, and in lane 9, with a primer–undamaged downstream strand–template substrate.

<https://doi.org/10.1371/journal.pone.0198480.g003>

illustrated by the experiment with KF, which efficiently extended the primer but stopped ~4 nucleotides before the cross-link site (Fig 1B, Table 1). Bands corresponding to products of hydrolysis of the 11-mer primer were observed together with the elongation product due to the 3'→5' exonuclease activity of KF (Fig 1B). Exonuclease-deficient KF (KF<sup>exo-</sup>) behaved in the same manner but expectedly showed no primer degradation (S2 Fig). Another member of Family A, *Taq* DNA polymerase, stopped before the cross-link at approximately the same position as KF did (S3 Fig, Table 1) and showed several pause points at earlier reaction times.

**Table 1. Termination sites of DNA polymerases on DPCs.**

Family	Poly-merase	DPC in template, single-stranded	DPC in template, double-stranded	DPC in displaced strand
A	KF	4 nt before DPC	5 nt before DPC	3–4 nt before DPC
A	<i>Taq</i>	3–4 nt before DPC	5 nt before DPC	3–5 nt before DPC
B	RB69	immediately before DPC	5–6 nt before DPC	immediately before DPC
B	T4	immediately before DPC	no strand displacement, 3'→5' degradation	no strand displacement, 3'→5' degradation
B	<i>Pfu</i>	4 nt before DPC	no strand displacement <sup>a</sup>	no strand displacement
X	Pol β	4–5 nt before DPC	3–4 nt before DPC	4–5 nt before DPC
X	Pol λ	3 nt before DPC	8 nt before DPC	8 nt before DPC
Y	Dpo4	immediately before DPC	poor strand displacement <sup>b</sup>	poor strand displacement
Y	Pol κ	4 nt before DPC	poor strand displacement	poor strand displacement <sup>c</sup>

<sup>a</sup>No strand displacement: insertion of a single nucleotide into the gap without further DNA synthesis in the undamaged control substrate.

<sup>b</sup>Poor strand displacement: insertion of 3–4 nucleotides without further DNA synthesis in the undamaged control substrate.

<sup>c</sup>For Pol κ, very limited progress to 2–3 nt before DPC was observed.

<https://doi.org/10.1371/journal.pone.0198480.t001>

Family B comprises mainly replicative DNA polymerases having high fidelity and involved in the synthesis of both leading and lagging strands during replication (eukaryotic DNA polymerases  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$ , some archaeal and phage DNA polymerases). Many of them have high 3'→5' exonuclease activity (proofreading activity) to remove erroneously incorporated nucleotides not complementary to the template. We have investigated interaction of Family B enzymes with DPCs using DNA polymerases from bacteriophages RB69 and T4 and archaeal *Pfu* DNA polymerase [27–29]. In contrast to Family A polymerases on substrates containing a DPC in the single-stranded template, RB69pol synthesized a 23-mer product, which means that it was able to reach the cross-link site despite the presence of a bulky protein adduct (Fig 1C). T4 DNA polymerase was also able to reach the cross-link site, although due to its strong exonuclease activity significant primer degradation was observed (S4 Fig). The activity of the most stable *Pfu* DNA polymerase, which has a temperature optimum at 72–74°C, was rather low at 37°C. However, it still can be observed that on a substrate with a DPC-containing single-stranded template, *Pfu* DNA polymerase stopped 4–5 nucleotides before the cross-link site, as do Family A polymerases (S6 Fig). It remains to be seen whether *Pfu* DNA polymerase could read through closer to the cross-link site at its optimal temperature.

Family X includes eukaryotic DNA polymerase  $\beta$  (Pol  $\beta$ ) involved in the process of short-patch base excision repair [30,31], as well as DNA polymerases  $\lambda$  (Pol  $\lambda$ ) and  $\mu$  likely participating in non-homologous end joining during double-stranded breaks repair [32,33]. We have investigated the activity of two representative members of Family X, Pol  $\beta$  and Pol  $\lambda$ . Unlike the previously described polymerases, these enzymes perform distributive synthesis on primer–template substrates: they catalyze single nucleotide addition to a primer, release DNA substrate, and then this cycle is repeated; however, both polymerases are processive on substrates containing short gaps, interacting with the downstream strand [34–37]. Both Pol  $\beta$  and Pol  $\lambda$  did not pass further than 3–5 nucleotides before the cross-link site, which makes them similar to Family A polymerases (S7 and S8 Figs). However, the efficiency of the polymerase reaction was clearly lower than in the cases of Families A and B, likely reflecting the preference of Family X enzymes for gapped rather than primer–template substrates. Interestingly, both Pol  $\beta$  and Pol  $\lambda$  synthesized longer products in the presence of a DPC than on an undamaged template (Panel A in S7 Fig, Panel A in S8 Fig, compare length of the products in lanes 7 and 8); we speculate that the cross-link in a single-stranded template might provide some non-specific protein–protein interactions downstream that help the polymerase to hold on the template and approach closer to the damage site.

Family Y DNA polymerases stand out in their ability to perform translesion synthesis, that is, to incorporate dNTP opposite damaged DNA bases [38]. This capacity is explained by a short fingers domain and a wide active site, which together allow the polymerase to bind DNA even if it is distorted by damage. The reverse of this flexibility is high error frequency, which is greater by 1–2 orders of magnitude in Family Y DNA polymerases than in other families. Family Y includes DNA polymerases such as human DNA polymerases  $\eta$ ,  $\iota$ ,  $\kappa$  (Pol  $\kappa$ ), Rev1, and bacterial and archaeal DNA polymerases IV and V. In our research we have investigated the activity of Family Y enzymes on substrates containing a DPC using DNA polymerase IV from *Sulfolobus solfataricus* (Dpo4) and human Pol  $\kappa$ . In the experiments with Dpo4 when a DPC was in single-stranded template, we observed several pause points after adding 2, 4 and 8 nucleotides to the primer, which did not occur with the control primer–template substrate (S9 Fig). Interestingly, they all correspond to dGMP incorporation in the growing strand, and the pause may be related to the fact that Dpo4 uses dGTP with the worst kinetics of all possible dNTPs [39]. It is possible that similar kinetic considerations explain the presence of pause points in the control reactions for some other reactions studied polymerases (for instance, *Taq* DNA polymerase). A band of low intensity corresponding to the length of 23 nucleotides was

observed after prolonged incubation of Dpo4 with the substrate, which means that this polymerase is able to continue synthesis to the site of the cross-link with some efficiency. Pol  $\kappa$  during synthesis on the same substrate has shown no pause points and stopped 4 nucleotides before the site of the cross-link (Fig 1D).

### Primer extension with DPCs in the template strand of double-stranded DNA

DNA polymerases vary widely in their ability to synthesize the new strand of DNA when the downstream strand is present. Family A DNA polymerases degrade the downstream strand through the activity located in their 5'→3'-exonuclease domain (which is not a true exonuclease but a structure-specific endonuclease, cleaving off short flaps of displaced DNA) [40]. If this domain is removed, their remaining catalytic domain maintains the ability to displace the downstream strand. Family B polymerases range from non-displacing (*Pfu* and T4 polymerases) to strongly displacing (phage RB69 and  $\phi$ 29 polymerases) [41–43]. Family X enzymes weakly to moderately displace the downstream strand through a “hit-and-run” mechanism, adding and displacing nucleotides one by one [44]. Family Y polymerases also show variable strand displacement activity [45,46].

The presence of the second strand in the DPC-containing part of the substrate profoundly affected the ability of all studied DNA polymerases to progress to the cross-link site. T4 and *Pfu* DNA polymerases, in line with their lack of displacement activity on non-damaged templates, added a single nucleotide to fill the gap (Fig 2A) and did not proceed further (S4 and S6 Figs). T4 polymerase also strongly degraded the primer. Dpo4 and Pol  $\kappa$  poorly displaced the downstream strand even from unmodified control duplexes and showed equally low or worse displacement if DPCs were present (S9 and S10 Figs).

Other DNA polymerases showed some extension but stopped earlier than in the absence of the downstream strand. This difference was especially pronounced in RB69pol (5–6 nt before DPC compared with immediately before DPC in the single-stranded template) and Pol  $\lambda$  (8 nt vs 3 nt before DPC in the single-stranded template) (Fig 2C, S8 Fig). The only exception was Pol  $\beta$  that seemed to stop even closer to the DPC (3–4 nt vs 4–5 nt before DPC in the single-stranded template; S7 Fig), which may be related to the general preference of this polymerase for substrates with a downstream strand. Otherwise, Family A DNA polymerases were the least affected by the presence of the downstream strand, inserting only 1–2 nucleotides less than in the situation with DPC in the single-stranded template (Fig 2B, Table 1).

### Primer extension with DPCs in the displaced strand of double-stranded DNA

The situation with DPCs in the displaced strand is especially interesting, since no translesion synthesis is expected and, provided the polymerase could separate the strands held together by the protein cross-linked to one of them, bypass is possible. However, as our results indicate, displacement of the cross-link-bearing strand presents a challenge even for the polymerases with the strongest displacing activity.

Family A DNA polymerases, which are able to displace or degrade the downstream strand, were about as proficient in synthesizing DNA on substrates with a DPC in the downstream strand as when the DPC was located in the single template strand. KF  $\text{exo}^+$  was able to elongate the primer by 8–9 nucleotides, then stopped 3–4 nucleotides before the cross-link site, whereas KF  $\text{exo}^-$  pushed even further (S1 and S2 Figs). Another member of Family A, *Taq* DNA polymerase, stopped before the cross-link at about the same position as KF and couldn't continue synthesis (Fig 3B). Unlike KF, *Taq* polymerase paused in the beginning of elongation after



single-nucleotide gap filling but this pause was later overcome by the enzyme. The same pause point was present in the control reaction, indicating that it could be due to suboptimal temperature conditions for strand degradation by this DNA polymerase. Thus, despite Family A DNA polymerases have an ability to tolerate some bulky peptide adducts [26], they were not able to bypass DPCs in any context.

Family B RB69pol was able to progress until very near to the cross-link site, as in the case with the adduct in the single-stranded template (S5 Fig). Asymmetric orientation of Fpg on DNA, as seen in the crystal structure [20], shifts the edge of Fpg footprint by 2 nt in the 3' direction when the cross-link is in the displaced strand relative to the DPC in the template strand (S11 Fig). However, this difference alone cannot explain the ~7 nt difference in the polymerase termination point, indicating that RB69pol can partially displace the downstream strand even having such a bulky adduct as the DPC. We have inquired whether this displacement can be further assisted by degradation of the downstream DNA by structure-specific endonuclease FEN1. However, no progress of RB69pol beyond the cross-link position was seen even in the presence of FEN1. T4 and *Pfu* DNA polymerases had no strand displacement activity on this type of substrate and behaved identically to their reaction with DPCs in the template in double-stranded DNA (S4 and S6 Figs).

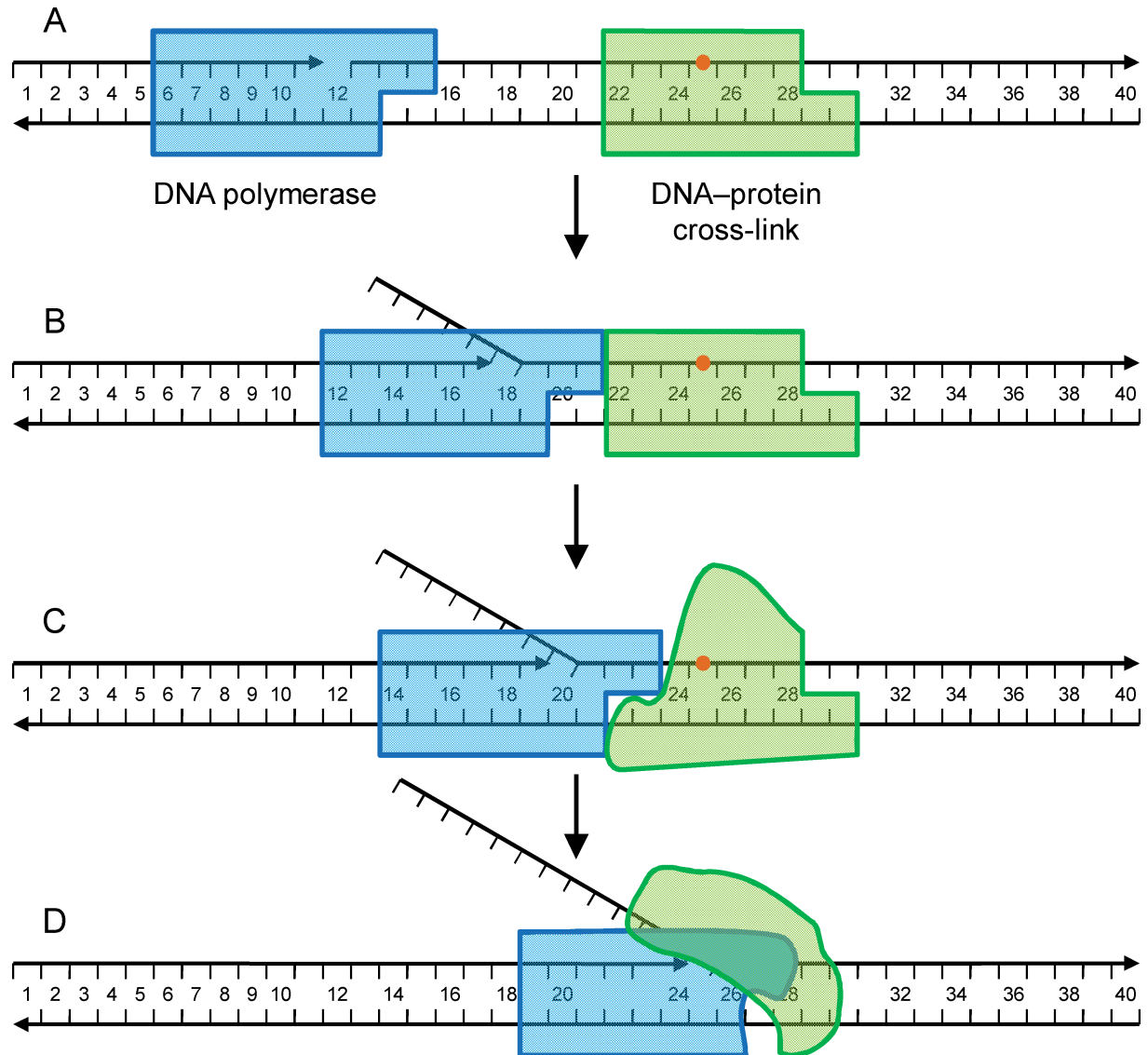
Family X polymerases performed similarly to their activity on substrates with DPCs in the template. Pol  $\beta$  stopped 4–5 nucleotides before the cross-link site whereas Pol  $\lambda$ , with its poorer strand displacement ability, terminated ~8 nucleotides before the cross-link (Fig 3B, S6 Fig). Here again the difference between the termination positions of Pol  $\beta$  when the DPC was in the template and in the displaced strand could not be explained by 2-nt asymmetry in the protein footprint (S11 Fig).

Finally, Family Y DNA polymerases (Dpo4 and Pol  $\kappa$ ) showed poor strand displacement with the control substrates. Only Dpo4 at long reaction times showed a weak signal corresponding to extension to the position 2–3 nt before DPC (S9 Fig).

## Discussion

Relative to a typical size of DNA lesions, DNA–protein cross-links are very bulky adducts. As a DPC model, we have used Fpg protein trapped on DNA by NaBH<sub>4</sub> treatment; according to its crystal structure, Fpg fully or partially covers ~7 nt on the damaged DNA strand, and ~9 nt on the complementary strand [20] (S1 Fig). In our substrates, 8-oxoguanine, the lesion that Fpg excises and gets trapped to this position, was 13 nucleotides from the 3'-terminus of the primer. In turn, DNA polymerases, even though their size varies, are also large molecules (38–114 kDa for polymerases used in our study; cf. 30 kDa for Fpg). Thus, from the steric considerations, DNA polymerase should run into the protein part of the DPC soon after the start of synthesis. The available X-ray structures [47,48] allow us to estimate the distance at the start between the polymerase and Fpg for Pol  $\beta$  and RB69pol. When Fpg is cross-linked to the displaced strand, the surfaces of Pol  $\beta$  and RB69pol are, respectively, 6 nt and 5 nt away from the surface of Fpg. Thus, the termination of Pol  $\beta$  4–5 nt before the site of the cross-link, correspond to an overlap of 2–3 nt between its footprint and the footprint of Fpg. When Fpg is cross-linked to the template strand, the surfaces of Pol  $\beta$  and Fpg are separated by 4 nt, due to the asymmetry of the Fpg footprint, and the overlap at the termination site is 0–1 nt. Even more strikingly, stalled RB69pol overlaps over at least 4 nt with Fpg in the template strand and nearly with the full footprint of Fpg cross-linked to the displaced strand.

The results with all DNA polymerases studied in this work are summarized in Table 1. Although there was considerable heterogeneity in the position of termination sites, our observations suggest that DNA polymerase interaction with DPCs may be described by a scheme



**Fig 4. Scheme of the “kiss-and-push” mechanism, shown for substrates containing a DPC in the displaced strand of double-stranded DNA.** A, substrate; B, after insertion of several nucleotides, the DNA polymerase comes to contact with the cross-linked protein (the kiss stage); C, the ability of some DNA polymerases to distort the cross-linked protein aids the incorporation of a few extra nucleotides (the push stage); D, the synthesis stops when the polymerase itself gets distorted.

<https://doi.org/10.1371/journal.pone.0198480.g004>

that we term “kiss-and-push model” (Fig 4). The incorporation of nucleotides is mostly guided by the ability of the DNA polymerase to perform synthesis on a given substrate (e. g., by its strand displacement activity) until the surfaces of DNA polymerase and protein in the DPC are in contact (the “kiss” stage). Most polymerases stop at this point but some, especially Family B polymerases and Dpo4 (in single-stranded DNA), can “push” the cross-linked protein to read through up to the site of cross-linking, until the polymerase is distorted to a degree where it cannot catalyze nucleotide incorporation anymore. Even at the kiss stage, it is likely that the protein globules of the polymerase, or the cross-linked protein, or both are partially deformed, since the termination point is closer to the cross-link site than allowed by the size of both proteins. In this respect, the kiss stage resembles the stalling of reverse transcriptases by RNA-

binding proteins in the toeprinting assay used to analyze translation complexes; in this case, cDNA synthesis stops at or very close to the 3'-edge of the ribosome or other protein obstacle [49]. Family B polymerases would have the highest ability to distort the protein obstacle.

None of the studied DNA polymerases was able to bypass DPCs. This is not unexpected when a DPC is in the template strand, since accommodation of adducts of this size in the active site should be beyond the capability even for Family Y polymerases, and the attached protein is likely too large to permit template slippage to skip the cross-link. Also, lack of the base at the site of the cross-link would contribute to stalling the polymerase, although simple abasic sites can be bypassed after a pause by polymerases from Families A, B, X, and Y [50–52]. Moreover, a comparison of the results with DPCs in single- and double-stranded template clearly indicates that the displaced strand impedes the progress of DNA polymerases near the bulky protein adduct.

The inability of strand-displacing DNA polymerases to fully extend the primer when the DPC is in the displaced strand is more intriguing. There are no available X-ray structures of DNA polymerases in the strand displacement complex. In a model of the strand displacement complex of phage  $\phi$ 29 DNA polymerase that belongs to Family B [42,53], the displaced strand forms few contacts with the protein. However DPCs are likely bulky enough to jam the polymerase at some point along the path of the displaced strand, or to distort the DNA duplex to a degree that would not allow the incoming polymerase to draw the non-damaged template strand into the active site.

Overall, it is clear that DPCs indeed present an almost insurmountable problem for bypass by DNA polymerases. This necessitates the repair that involves specific proteases, such as Wss1 in yeast and SPRTN in metazoans, which degrade the protein to a cross-linked peptide of a manageable size and are required for replication in the presence of DPC [54–57]. Biallelic mutations in the *SPRTN* gene cause Ruijs–Aalfs syndrome, a hereditary disease characterized by progeroid features, early cancer development, and hypersensitivity to DNA damage, likely a result of accumulation of unrepaired dead-end topoisomerase 1 complexes [58–60]. It would be interesting to see whether the distortion in the cross-linked protein induced by the collision with a polymerase might expose normally buried peptides and activate the nuclear misfolded protein response [61]. In addition, if replication is concerned, interaction of DNA helicases with DPCs could trigger response that stops the replicative fork and prevents the polymerase from collision with DPCs. Family X polymerases, which are involved in DNA repair, stop at least 4–5 nucleotides before the cross-link site. Therefore, in clustered damage where a DPC is combined with a modified DNA base, abasic site, or single-strand break, the types of lesions repaired with participation of Pol  $\beta$  and Pol  $\lambda$  [62], the complete repair could be difficult. These processes present an interesting topic worthy of further research, and the Fpg cross-link system used in this study provides a convenient model, now characterized at the basic level, that could be used *in vitro* or easily introduced into plasmids for experiments in living cells.

## Supporting information

**S1 Fig. Termination sites of *E. coli* DNA polymerase I Klenow fragment (Family A) at the DNA–protein cross-link in the template strand of double-stranded DNA (A) and in the displaced strand of double-stranded DNA (B).** Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 8) or containing (lane 9) the displaced strand.

(PDF)

**S2 Fig. Termination sites of *E. coli* DNA polymerase I Klenow fragment  $\text{exo}^-$  (Family A)** at the DNA-protein cross-link in the template strand of single-stranded DNA (A) and in the displaced strand of double-stranded DNA (B). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 8) or containing (lane 9) the displaced strand. (PDF)

**S3 Fig. Termination sites of *Taq* polymerase (Family A)** at the DNA-protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 8) or containing (lane 9) the displaced strand. (PDF)

**S4 Fig. Termination sites of phage T4 DNA polymerase (Family B)** at the DNA-protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B) and in the displaced strand of double-stranded DNA (C). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 3 in panel B; lane 8 in panels A, C) or containing (lane 4 in panel B; lane 9 in panel C) the displaced strand. (PDF)

**S5 Fig. Termination sites of phage RB69 DNA polymerase (Family B) at the DNA-protein cross-link in the displaced strand of double-stranded DNA.** Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 8) or containing (lane 9) the displaced strand. (PDF)

**S6 Fig. Termination sites of *Pfu* DNA polymerase (Family B)** at the DNA-protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B) and in the displaced strand of double-stranded DNA (C). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 7 in panel A; lane 8 in panels B, C) or containing (lane 9) the displaced strand. (PDF)

**S7 Fig. Termination sites of human DNA polymerase  $\beta$  (Family X)** at the DNA-protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 7 in panel A, lane 8 in panel B) or containing (lane 9 in panel B) the displaced strand. (PDF)

**S8 Fig. Termination sites of human DNA polymerase  $\lambda$  (Family X)** at the DNA–protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B) and in the displaced strand of double-stranded DNA (C). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions (lane 8) were carried out with an undamaged substrate lacking the displaced strand.  
(PDF)

**S9 Fig. Termination sites of *S. solfataricus* DNA polymerase IV (Family Y)** at the DNA–protein cross-link in the template strand of single-stranded (A) or double-stranded DNA (B) and in the displaced strand of double-stranded DNA (C). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 7 in panel A, lane 8 in panels B, C) or containing (lane 9 in panels B, C) the displaced strand.  
(PDF)

**S10 Fig. Termination sites of human DNA polymerase  $\kappa$  (Family Y)** at the DNA–protein cross-link in the template strand (A) or in the displaced strand of double-stranded DNA (B). Lanes 1–3, size markers (primer, 11 nt long; primer extended to the cross-link site, 23 nt; full-size product, 40 nt); the arrows indicate their positions. The presence of DNA polymerase, cross-linked Fpg, and the reaction time are shown under the gel images. Control reactions were carried out with an undamaged substrate lacking (lane 8) or containing (lane 9) the displaced strand.  
(PDF)

**S11 Fig. A**, arrangement of DNA polymerases and the Fpg cross-link on DNA with sizes of the proteins inferred from the structural data. PDB IDs for structures of Fpg, Pol  $\beta$  and RB69 polymerase are indicated. Arrowheads mark the 3' termini of the oligonucleotides. The orange dot shows the site of cross-linking. The sphere pattern marks the approximate position of the RB69 domain absent from the crystal structure. **B**, termination positions of Pol  $\beta$  at the Fpg cross-link covalently bound to the displaced strand or the template strand. Blue arrows mark the position of the last incorporated dNMP; red arrows, the corresponding position of the front side of Pol  $\beta$ .  
(PDF)

**S1 Table. Oligonucleotides used in this work.**  
(PDF)

## Acknowledgments

We thank Dr. Olga Lavrik (SB RAS ICBFM) for providing human Pol  $\beta$ , Pol  $\lambda$  and FEN1. We are grateful to Drs Olga Lavrik, Pavel Pestryakov (SB RAS ICBFM) and Olga Sinitsyna (Institute of Cytology and Genetics, Novosibirsk, Russia) for valuable discussions regarding the research and the structure of the paper.

## Author Contributions

**Conceptualization:** Dmitry O. Zharkov.

**Formal analysis:** Anna V. Yudkina, Dmitry O. Zharkov.

**Funding acquisition:** Dmitry O. Zharkov.

**Investigation:** Anna V. Yudkina, Antonina P. Dvornikova.

**Methodology:** Anna V. Yudkina, Antonina P. Dvornikova, Dmitry O. Zharkov.

**Project administration:** Dmitry O. Zharkov.

**Resources:** Dmitry O. Zharkov.

**Supervision:** Dmitry O. Zharkov.

**Visualization:** Dmitry O. Zharkov.

**Writing – original draft:** Anna V. Yudkina.

**Writing – review & editing:** Anna V. Yudkina, Antonina P. Dvornikova, Dmitry O. Zharkov.

## References

1. Lindahl T. Instability and decay of the primary structure of DNA. *Nature*. 1993; 362(6422):709–715. <https://doi.org/10.1038/362709a0> PMID: 8469282
2. Barker S, Weinfeld M, Murray D. DNA–protein crosslinks: Their induction, repair, and biological consequences. *Mutat Res*. 2005; 589(2):111–135. <https://doi.org/10.1016/j.mrrev.2004.11.003> PMID: 15795165
3. Reardon JT, Cheng Y, Sancar A. Repair of DNA-protein cross-links in mammalian cells. *Cell Cycle*. 2006; 5(13):1366–1370. <https://doi.org/10.4161/cc.5.13.2892> PMID: 16775425
4. Ide H, Shoukamy MI, Nakano T, Miyamoto-Matsubara M, Salem AMH. Repair and biochemical effects of DNA–protein crosslinks. *Mutat Res*. 2011; 711(1–2):113–122. <https://doi.org/10.1016/j.mrfmmm.2010.12.007> PMID: 21185846
5. Olinski R, Nackerdien Z, Dizdaroglu M. DNA-protein cross-linking between thymine and tyrosine in chromatin of  $\gamma$ -irradiated or H<sub>2</sub>O<sub>2</sub>-treated cultured human cells. *Arch Biochem Biophys*. 1992; 297(1):139–143. [https://doi.org/10.1016/0003-9861\(92\)90651-C](https://doi.org/10.1016/0003-9861(92)90651-C) PMID: 1322110
6. Barker S, Weinfeld M, Zheng J, Li L, Murray D. Identification of mammalian proteins cross-linked to DNA by ionizing radiation. *J Biol Chem*. 2005; 280(40):33826–33838. <https://doi.org/10.1074/jbc.M502477200> PMID: 16093242
7. Xu X, Muller JG, Ye Y, Burrows CJ. DNA–protein cross-links between guanine and lysine depend on the mechanism of oxidation for formation of C5 vs C8 guanosine adducts. *J Am Chem Soc*. 2008; 130(2):703–709. <https://doi.org/10.1021/ja077102a> PMID: 18081286
8. Chválová K, Brabec V, Kašpárková J. Mechanism of the formation of DNA–protein cross-links by antitumor cisplatin. *Nucleic Acids Res*. 2007; 35(6):1812–1821. <https://doi.org/10.1093/nar/gkm032> PMID: 17329374
9. Yeo JE, Wickramaratne S, Khatwani S, Wang Y-C, Vervacke J, Distefano MD, et al. Synthesis of site-specific DNA–protein conjugates and their effects on DNA replication. *ACS Chem Biol*. 2014; 9(8):1860–1868. <https://doi.org/10.1021/cb5001795> PMID: 24918113
10. Wickramaratne S, Ji S, Mukherjee S, Su Y, Pence MG, Lior-Hoffmann L, et al. Bypass of DNA-protein cross-links conjugated to the 7-deazaguanine position of DNA by translesion synthesis polymerases. *J Biol Chem*. 2016; 291(45):23589–23603. <https://doi.org/10.1074/jbc.M116.745257> PMID: 27621316
11. Minko IG, Yamanaka K, Kozekov ID, Kozekova A, Indiani C, O'Donnell ME, et al. Replication bypass of the acrolein-mediated deoxyguanine DNA-peptide cross-links by DNA polymerases of the DinB family. *Chem Res Toxicol*. 2008; 21(10):1983–1990. <https://doi.org/10.1021/tx800174a> PMID: 18788757
12. Yamanaka K, Minko IG, Takata K-i, Kolbanovskiy A, Kozekov ID, Wood RD, et al. Novel enzymatic function of DNA polymerase  $\nu$  in translesion DNA synthesis past major groove DNA–peptide and DNA–DNA cross-links. *Chem Res Toxicol*. 2010; 23(3):689–695. <https://doi.org/10.1021/tx900449u> PMID: 20102227
13. Wickramaratne S, Boldry EJ, Buehler C, Wang Y-C, Distefano MD, Tretyakova NY. Error-prone translesion synthesis past DNA-peptide cross-links conjugated to the major groove of DNA via C5 of thymidine. *J Biol Chem*. 2015; 290(2):775–787. <https://doi.org/10.1074/jbc.M114.613638> PMID: 25391658
14. Braithwaite DK, Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res*. 1993; 21(4):787–802. <https://doi.org/10.1093/nar/21.4.787> PMID: 8451181

15. Yamtich J, Sweasy JB. DNA polymerase Family X: Function, structure, and cellular roles. *Biochim Biophys Acta*. 2010; 1804(5):1136–1150. <https://doi.org/10.1016/j.bbapap.2009.07.008> PMID: 19631767
16. Pata JD. Structural diversity of the Y-family DNA polymerases. *Biochim Biophys Acta*. 2010; 1804(5):1124–1135. <https://doi.org/10.1016/j.bbapap.2010.01.020> PMID: 20123134
17. Stinglee J, Jentsch S. DNA–protein crosslink repair. *Nat Rev Mol Cell Biol*. 2015; 16(8):455–460. <https://doi.org/10.1038/nrm4015> PMID: 26130008
18. Dabney J, Meyer M, Pääbo S. Ancient DNA damage. *Cold Spring Harb Perspect Biol*. 2013; 5(7):a012567. <https://doi.org/10.1101/cshperspect.a012567> PMID: 23729639
19. Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues: Causes and strategies for minimization. *Clin Chem*. 2015; 61(1):64–71. <https://doi.org/10.1373/clinchem.2014.223040> PMID: 25421801
20. Gilboa R, Zharkov DO, Golan G, Fernandes AS, Gerchman SE, Matz E, et al. Structure of formamido-pyrimidine-DNA glycosylase covalently complexed to DNA. *J Biol Chem*. 2002; 277(22):19811–19816. <https://doi.org/10.1074/jbc.M202058200> PMID: 11912217
21. Derbyshire V, F. Grindley ND, Joyce CM. The 3'–5' exonuclease of DNA polymerase I of *Escherichia coli*: Contribution of each amino acid at the active site to the reaction. *EMBO J*. 1991; 10(1):17–24. none PMID: 1989882
22. Freisinger E, Grollman AP, Miller H, Kisker C. Lesion (in)tolerance reveals insights into DNA replication fidelity. *EMBO J*. 2004; 23(7):1494–1505. <https://doi.org/10.1038/sj.emboj.7600158> PMID: 15057282
23. Sun B, Latham KA, Dodson ML, Lloyd RS. Studies of the catalytic mechanism of five DNA glycosylases: Probing for enzyme-DNA imino intermediates. *J Biol Chem*. 1995; 270(33):19501–19508. <https://doi.org/10.1074/jbc.270.33.19501> PMID: 7642635
24. Ishchenko AA, Vasilenko NL, Sinitsina OI, Yamkovoy VI, Fedorova OS, Douglas KT, et al. Thermodynamic, kinetic, and structural basis for recognition and repair of 8-oxoguanine in DNA by Fpg protein from *Escherichia coli*. *Biochemistry*. 2002; 41(24):7540–7548. <https://doi.org/10.1021/bi0121297> PMID: 12056884
25. Patel PH, Suzuki M, Adman E, Shinkai A, Loeb LA. Prokaryotic DNA polymerase I: Evolution, structure, and “base flipping” mechanism for nucleotide selection. *J Mol Biol*. 2001; 308(5):823–837. <https://doi.org/10.1006/jmbi.2001.4619> PMID: 11352575
26. Yamanaka K, Minko IG, Finkel SE, Goodman MF, Lloyd RS. Role of high-fidelity *Escherichia coli* DNA polymerase I in replication bypass of a deoxyadenosine DNA-peptide cross-link. *J Bacteriol*. 2011; 193(15):3815–3821. <https://doi.org/10.1128/JB.01550-10> PMID: 21622737
27. Young MC, Reddy MK, von Hippel PH. Structure and function of the bacteriophage T4 DNA polymerase holoenzyme. *Biochemistry*. 1992; 31(37):8675–8690. <https://doi.org/10.1021/bi00152a001> PMID: 1390652
28. Komori K, Ishino Y. Functional interdependence of DNA polymerizing and 3'–5' exonucleolytic activities in *Pyrococcus furiosus* DNA polymerase I. *Protein Eng*. 2000; 13(1):41–47. <https://doi.org/10.1093/protein/13.1.41> PMID: 10679529
29. Xia S, Konigsberg WH. RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry*. 2014; 53(17):2752–2767. <https://doi.org/10.1021/bi4014215> PMID: 24720884
30. Sobol RW, Horton JK, Kühn R, Gu H, Singhal RK, Prasad R, et al. Requirement of mammalian DNA polymerase-β in base-excision repair. *Nature*. 1996; 379(6561):183–186. <https://doi.org/10.1038/379183a0> PMID: 8538772
31. Podlutzky AJ, Dianova II, Wilson SH, Bohr VA, Dianov GL. DNA synthesis and dRPase activities of polymerase β are both essential for single-nucleotide patch base excision repair in mammalian cell extracts. *Biochemistry*. 2001; 40(3):809–813. <https://doi.org/10.1021/bi002064s> PMID: 11170398
32. Lee JW, Blanco L, Zhou T, Garcia-Diaz M, Bebenek K, Kunkel TA, et al. Implication of DNA polymerase λ in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J Biol Chem*. 2004; 279(1):805–811. <https://doi.org/10.1074/jbc.M307913200> PMID: 14561766
33. Daley JM, Vander Laan RL, Suresh A, Wilson TE. DNA joint dependence of pol X family polymerase action in nonhomologous end joining. *J Biol Chem*. 2005; 280(32):29030–29037. <https://doi.org/10.1074/jbc.M505277200> PMID: 15964833
34. Prasad R, Beard WA, Wilson SH. Studies of gapped DNA substrate binding by mammalian DNA polymerase β: Dependence on 5'-phosphate group. *J Biol Chem*. 1994; 269(27):18096–18101. none PMID: 8027071
35. Wilson SH. Mammalian base excision repair and DNA polymerase beta. *Mutat Res*. 1998; 407(3):203–215. [https://doi.org/10.1016/S0921-8777\(98\)00002-0](https://doi.org/10.1016/S0921-8777(98)00002-0) PMID: 9653447

36. García-Díaz M, Bebenek K, Sabariego R, Domínguez O, Rodríguez J, Kirchoff T, et al. DNA polymerase  $\lambda$ , a novel DNA repair enzyme in human cells. *J Biol Chem*. 2002; 277(15):13184–13191. <https://doi.org/10.1074/jbc.M111601200> PMID: 11821417
37. Garcia-Diaz M, Bebenek K, Gao G, Pedersen LC, London RE, Kunkel TA. Structure–function studies of DNA polymerase lambda. *DNA Repair*. 2005; 4(12):1358–1367. <https://doi.org/10.1016/j.dnarep.2005.09.001> PMID: 16213194
38. Sale JE, Lehmann AR, Woodgate R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol*. 2012; 13(3):141–152. <https://doi.org/10.1038/nrm3289> PMID: 22358330
39. Boudsocq F, Iwai S, Hanaoka F, Woodgate R. *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): An archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol $\eta$ . *Nucleic Acids Res*. 2001; 29(22):4607–4616. <https://doi.org/10.1093/nar/29.22.4607> PMID: 11713310
40. Ceska TA, Sayers JR. Structure-specific DNA cleavage by 5' nucleases. *Trends Biochem Sci*. 1998; 23(9):331–336. [https://doi.org/10.1016/S0968-0004\(98\)01259-6](https://doi.org/10.1016/S0968-0004(98)01259-6) PMID: 9787638
41. Hacker KJ, Alberts BM. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix: A model for polymerase release following the termination of each Okazaki fragment. *J Biol Chem*. 1994; 269(39):24221–24228. none PMID: 7929078
42. Kamtekar S, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, et al. Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage  $\phi$ 29. *Mol Cell*. 2004; 16(4):609–618. <https://doi.org/10.1016/j.molcel.2004.10.019> PMID: 15546620
43. Steitz TA, Yin YW. Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos Trans R Soc Lond B Biol Sci*. 2004; 359(1441):17–23. <https://doi.org/10.1098/rstb.2003.1374> PMID: 15065652
44. Liu Y, Beard WA, Shock DD, Prasad R, Hou EW, Wilson SH. DNA polymerase  $\beta$  and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. *J Biol Chem*. 2005; 280(5):3665–3674. <https://doi.org/10.1074/jbc.M412922200> PMID: 15561706
45. Dorjsuren D, Wilson DM III, Beard WA, McDonald JP, Austin CP, Woodgate R, et al. A real-time fluorescence method for enzymatic characterization of specialized human DNA polymerases. *Nucleic Acids Res*. 2009; 37(19):e128. <https://doi.org/10.1093/nar/gkp641> PMID: 19684079
46. Pomerantz RT, Kurth I, Goodman MF, O'Donnell ME. Preferential D-loop extension by a translesion DNA polymerase underlies error-prone recombination. *Nat Struct Mol Biol*. 2013; 20(6):748–755. <https://doi.org/10.1038/nsmb.2573> PMID: 23686288
47. Krahn JM, Beard WA, Miller H, Grollman AP, Wilson SH. Structure of DNA polymerase  $\beta$  with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure*. 2003; 11(1):121–127. [https://doi.org/10.1016/S0969-2126\(02\)00930-9](https://doi.org/10.1016/S0969-2126(02)00930-9) PMID: 12517346
48. Zahn KE, Belrhali H, Wallace SS, Doublé S. Caught bending the A-rule: Crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry*. 2007; 46(37):10551–10561. <https://doi.org/10.1021/bi7008807> PMID: 17718515
49. Ringquist S, Gold L. Toeprinting assays: Mapping by blocks to reverse transcriptase primer extension. *Methods Mol Biol*. 1998; 77:283–295. <https://doi.org/10.1385/0-89603-397-X:283> PMID: 9770677
50. Paz-Elizur T, Takeshita M, Livneh Z. Mechanism of bypass synthesis through an abasic site analog by DNA polymerase I. *Biochemistry*. 1997; 36(7):1766–1773. <https://doi.org/10.1021/bi9621324> PMID: 9048560
51. Kokoska RJ, McCulloch SD, Kunkel TA. The efficiency and specificity of apurinic/apyrimidinic site bypass by human DNA polymerase  $\eta$  and *Sulfolobus solfataricus* Dpo4. *J Biol Chem*. 2003; 278(50):50537–50545. <https://doi.org/10.1074/jbc.M308515200> PMID: 14523013
52. Villani G, Hubscher U, Gironis N, Parkkinen S, Pospiech H, Shevelev I, et al. *In vitro* gap-directed translesion DNA synthesis of an abasic site involving human DNA polymerases  $\epsilon$ ,  $\lambda$ , and  $\beta$ . *J Biol Chem*. 2011; 286(37):32094–32104. <https://doi.org/10.1074/jbc.M111.246611> PMID: 21757740
53. Rodríguez I, Lázaro JM, Blanco L, Kamtekar S, Berman AJ, Wang J, et al. A specific subdomain in  $\phi$ 29 DNA polymerase confers both processivity and strand-displacement capacity. *Proc Natl Acad Sci USA*. 2005; 102(18):6407–6412. <https://doi.org/10.1073/pnas.0500597102> PMID: 15845765
54. Stinglele J, Schwarz MS, Bloemeke N, Wolf PG, Jentsch S. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell*. 2014; 158(2):327–338. <https://doi.org/10.1016/j.cell.2014.04.053> PMID: 24998930
55. Duxin JP, Dewar JM, Yardimci H, Walter JC. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell*. 2014; 159(2):346–357. <https://doi.org/10.1016/j.cell.2014.09.024> PMID: 25303529



56. Stinglee J, Bellelli R, Alte F, Hewitt G, Sarek G, Maslen SL, et al. Mechanism and regulation of DNA-protein crosslink repair by the DNA-dependent metalloprotease SPRTN. *Mol Cell*. 2016; 64(4):688–703. <https://doi.org/10.1016/j.molcel.2016.09.031> PMID: 27871365
57. Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S, et al. Metalloprotease SPRTN/DVC1 orchestrates replication-coupled DNA-protein crosslink repair. *Mol Cell*. 2016; 64(4):704–719. <https://doi.org/10.1016/j.molcel.2016.09.032> PMID: 27871366
58. Lessel D, Vaz B, Halder S, Lockhart PJ, Marinovic-Terzic I, Lopez-Mosqueda J, et al. Mutations in *SPRTN* cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet*. 2014; 46(11):1239–1244. <https://doi.org/10.1038/ng.3103> PMID: 25261934
59. Maskey RS, Kim MS, Baker DJ, Childs B, Malureanu LA, Jeganathan KB, et al. Spartan deficiency causes genomic instability and progeroid phenotypes. *Nat Commun*. 2014; 5:5744. <https://doi.org/10.1038/ncomms6744> PMID: 25501849
60. Maskey RS, Flatten KS, Sieben CJ, Peterson KL, Baker DJ, Nam H-J, et al. Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Res*. 2017; 45(8):4564–4576. <https://doi.org/10.1093/nar/gkx107> PMID: 28199696
61. Shibata Y, Morimoto RI. How the nucleus copes with proteotoxic stress. *Curr Biol*. 2014; 24(10):R463–R474. <https://doi.org/10.1016/j.cub.2014.03.033> PMID: 24845679
62. Belousova EA, Lavrik OI. DNA polymerases  $\beta$  and  $\lambda$  and their roles in cell. *DNA Repair*. 2015; 29:112–126. <https://doi.org/10.1016/j.dnarep.2015.02.001> PMID: 25715666