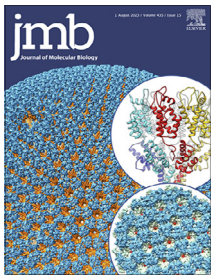
Research Article



Biochemical Reconstitution of the Mimiviral Base Excision Repair Pathway

Shailesh B. Lad [1], Monica Upadhyay [1,3], Pracheta Thorat [1], Divya Nair [1], Gregory W. Moseley [3], Sanjeeva Srivastava [1], Pradeepkumar P. I. [2] and Kiran Kondabagil [1]*

*1 - **Department of Biosciences and Bioengineering,** Indian Institute of Technology Bombay, Powai, Mumbai 400076, Maharashtra, India*
*2 - **Department of Chemistry,** Indian Institute of Technology Bombay, Powai, Mumbai 400076, Maharashtra, India*
*3 - **Biomedicine Discovery Institute,** Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia*

***Correspondence to Kiran Kondabagil:** kirankondabagil@iitb.ac.in, kirankondabagil@gmail.com (K. Kondabagil) @shailesh2603 (S.B. Lad), @upadhyaymonica (M. Upadhyay), @sanjeeva_IITB (S. Srivastava), @pradeepkumarpi (P.I. Pradeepkumar), @kirankondabagil (K. Kondabagil)*
https://doi.org/10.1016/j.jmb.2023.168188
Edited by M Gottesman

Abstract

Viruses are believed to be the obligate intracellular parasites that only carry genes essential for infecting and hijacking the host cell machinery. However, a recently discovered group of viruses belonging to the phylum nucleocytovirocota, also known as the nucleo-cytoplasmic large DNA viruses (NCLDVs), possess a number of genes that code for proteins predicted to be involved in metabolism, and DNA replication, and repair. In the present study, first, using proteomics of viral particles, we show that several proteins required for the completion of the DNA base excision repair (BER) pathway are packaged within the virions of Mimivirus as well as related viruses while they are absent from the virions of Marseillevirus and Kurlavirus that are NCLDVs with smaller genomes. We have thoroughly characterized three putative base excision repair enzymes from Mimivirus, a prototype NCLDV and successfully reconstituted the BER pathway using the purified recombinant proteins. The mimiviral uracil-DNA glycosylase (mvUDG) excises uracil from both ssDNA and dsDNA, a novel finding contrary to earlier studies. The putative AP-endonuclease (mvAPE) specifically cleaves at the abasic site created by the glycosylase while also exhibiting the 3′-5′ exonuclease activity. The Mimivirus polymerase X protein (mvPolX) can bind to gapped DNA substrates and perform single nucleotide gap-filling followed by downstream strand displacement. Furthermore, we show that when reconstituted *in vitro*, mvUDG, mvAPE, and mvPolX function cohesively to repair a uracil-containing DNA predominantly by long patch BER and together, may participate in the BER pathway during the early phase of Mimivirus life-cycle.

Introduction

DNA is constantly exposed to external as well as internal genotoxic agents. For instance, reactive oxygen species (ROS) produced during cellular metabolism can damage DNA. Lesions thus formed on the DNA cannot be read by replicative polymerases, leading to mutations that may prove lethal if not repaired.[1] For successful cell proliferation, any alterations in the DNA must be corrected

before the subsequent round of replication begins. The cellular response to DNA damage can proceed through DNA damage response (DDR) pathways or DNA damage tolerance. The latter of the two routes does not repair the DNA damage *per se*, but it avoids replication fork collapse by employing non-replicative polymerases that "bypass" the lesions on the DNA and allow continued replication.[2,3] On the contrary, DDR consists of several sub-pathways such as non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), homologous recombination (HR), nucleotide excision repair (NER), base excision repair (BER), and mismatch repair (MMR). While NHEJ, MMEJ, and HR are involved in the repair of DNA strand breaks, NER plays a key role in the removal of bulky and helix distorting modifications such as pyrimidine-pyrimidine dimers. The BER pathway is mainly involved in the repair of oxidized bases such as 8-oxo-guanine (8-oxo-dG), thymine and cytosine glycols, the deaminated base hypoxanthine, and alkylated bases such as 3-methyladenine and 7-methylguanosine.[4,5] Another important role of BER is to remove uracil bases from the DNA which may be the erroneously incorporated by DNA polymerases or formed due to spontaneous deamination of cytosine.[6]

The first step of the BER pathway consists of the action of DNA glycosylase/s that are responsible for the cleavage of the N-glycosidic bond that attaches the modified base to the DNA-sugar backbone. The removal of the base leads to the creation of an apurinic-apyrimidinic (abasic) site which is further processed by an AP-endonuclease (Apurinic/apyrimidinic (AP)-endonuclease) by producing a break in the sugar-phosphate backbone of the DNA.[7,8] Several glycosylases are bi-functional in nature and may also cleave the abasic site after base excision and therefore do not require the action of an endonuclease. The products formed after the incision of the abasic site may require further processing to remove the 5′ or 3′ blocking groups such as a 3′-phosphate, 5′-deoxyribose phosphate or a 3′ phospho-α,β-unsaturated aldehyde. Removal of these blocking moieties is often performed by polynucleotide kinase/phosphate enzyme or polymerase β.[7,8] After removal of the blocking groups, the polymerase utilizes the 3′-OH as a primer and fills the AP-site gap. In the final step, a DNA ligase is engaged which seals the gap created in the DNA to complete the repair process.[9–11] *In vivo*, several accessory proteins are also required for the completion of the BER pathway while *in vitro*, the minimal pathway can be reconstituted with as little as four proteins.[12]

The BER pathway has been studied extensively in eukaryotic and bacterial systems, however reports of BER enzymes in viruses are limited. The 1.2 Mb Mimivirus genome encodes about 1000 proteins with several predicted to be involved in DNA repair or related processes.[13–15] Very few of these proteins have been biochemically or functionally characterized such as the mimiviral primpol, and the bifunctional DNA glycosylase enzymes mvNEIL1 (ORF L315) and mvNEIL2 (ORF L720).[16–19] mvNEIL1 and mvNEIL2 proteins are similar to the human NEIL proteins with respect to their sequence and even substrate specificity. Apart from glycosylases, Mimivirus encodes endonuclease IV like protein (encoded by the R296 gene), an AP-endonuclease family 2 member (encoded by the L687 gene), and a putative phosphatase/polynucleotide kinase protein (PNK, gene L469).[15,20] Additionally, the Mimivirus genome codes for proteins involved in methylation-induced DNA repair pathways.[13] Finally, Mimivirus also contains a family X polymerase (product of gene L318), which is similar to the human pol β/pol λ and may be involved in the viral BER pathway.[20] Interestingly, family X polymerase is present in the viral particle.[21] The requirement for a reparative DNA polymerase in nucleo-cytoplasmic large DNA viruses probably stems from the need to maintain a large genome in a predominantly cytoplasmic life-cycle where the virus is devoid of the host nuclear factors.[22]

Therefore, with the holistic goal of understanding the BER pathway of Mimivirus, in the present study, we have performed a biochemical characterization of putative BER proteins, namely, a uracil DNA glycosylase (mvUDG), an AP-endonuclease (mvAPE), and a family-X polymerase (mvPolX) encoded by genes L249, R296, and L318, respectively. We also show that mvUDG, mvAPE, and mvPolX can function cohesively to repair a uracil containing DNA predominantly by the long patch BER pathway. In addition, using proteomics of mature virions, we also show that several proteins required for the BER pathway are packaged inside the virion and may participate in the BER pathway during the early phase of Mimivirus life cycle.

Results

Proteomics of Mimivirus and Marseillevirus virions

Mass spectrometry-based proteomics of the Mimivirus viral factory as well as of mature virions has been reported previously.[23,21] In the present study, we focused on identifying the proteins packaged within the Mimivirus virions that may potentially be involved in the DNA base excision repair (BER) pathway. Our analysis showed the presence of a number of proteins packaged within the capsids including several unidentified proteins as well as proteins having cellular homologs (data not shown). We identified the presence of several putative BER related proteins within the capsids (Table 1). *Acanthamoeba polyphaga* mimivirus (APMV) virions possess the full complement of proteins needed for the BER pathway (Table 1).[23,21] The presence

Table 1 The Putative BER-related proteins identified within the capsids of APMV, MVB, PLMV, MaV, and KV.

Virus (Genome size)	Packaged proteins								References
	UDG	DNA glycosylase	NEIL-1	NEIL-2	APE	XPG	PNKP	PolX	Ligase	
Mimivirus (~1.2 Mb)	-	+	+	+	+	+	+	+	+	Renesto et al., 2006; Fridmann-Sirkis et al., 2016; This study
Mimivirus Bombay (~1.2 Mb)	-	+	+	+	+	+	+	+	+	This study
Powai-lake megavirus (~1.2 Mb)	-	+	+	+	+	+	+	+	+	This study
Marseillevirus (~0.36 Mb)	+	-	-	-	-	-	-	-	+	Boyer et al., 2009; This study
Kurlavirus (~0.36 Mb)	-	-	-	-	-	-	-	-	+	This study

UDG-Uracil DNA Glycosylase, NEIL- Nei-like DNA Glycosylase, APE- Apurinic/ Apyrimidinic Endonuclease, XPG- Xeroderma pigmentosum group G (XPG) Endonuclease, PNKP- Polynucleotide Kinase/ Phosphatase, PolX- DNA Polymerase X. The "+" symbol in bold indicates proteins identified in the present study.

of three putative repair enzymes, namely, the endonuclease VIII-like protein (gp779/L720), the putative endonuclease IV (gp324/R296), and the polynucleotide kinase-phosphatase (gp502/L469) in the mimiviral capsid has not previously been reported (Table 1). In addition to APMV, we also performed identical proteomics analysis of the two related Mimivirus strains, Mimivirus bombay (MVB) and Powai-lake megavirus (PLMV) which were previously isolated by our lab.[24,25] Our proteomics study also showed that all the above proteins are also present in the virions of MVB and PLMV (Table 1) indicating that they may have essential roles in maintaining the genomic stability of these viruses having a large DNA (~1.2 Mb). In contrast, we found only DNA ligase and uracil DNA glycosylase enzymes in the virions of Marseillevirus[26] and Kurlavirus, both of which carry much smaller genomes (~360 kb) than the Mimiviruses (Table 1). The proteomics analysis shows that giant viruses with much larger genomes tend to package proteins involved in DNA repair within their virions which may aid in genome maintenance during the early phase of the viral life-cycle.

DNA polymerase activity of mvPolX

Sequence and secondary structure analysis of the mvPolX [gene product (gp) 347 encoded by gene L318] shows the presence of the finger, palm, and thumb subdomains that are typical of polymerases but lacks a proof-reading domain similar to other family X polymerases (Supplemental Figure S1). BLAST search identified additional family X polymerases from other NCLDVs such as Megavirus chiliensis, Klosneuvirus, Saudi momouvirus, Entomopoxvirus, *Phaeocystis globosa* virus, and African swine fever virus (ASFV). Although the overall sequence identity between mvPolX and other viral family X polymerases varies widely, the putative active site motifs including the key acidic residues involved in binding divalent cations ($D_{201}xD_{203}$ and xD_{269}) are conserved (Supplemental Figure S1). A comparison with eukaryotic homologs shows that mvPolX shares about 33% similarity with pol β (Supplemental Figure S1). Unlike pol λ, pol β and mvPolX do not contain the N-terminal BRCA1 carboxyl-terminal homology (BRCT) domain required for interaction with DNA and other proteins during non-homologous end joining (NHEJ).

The recombinant gp347 protein (mvPolX) was purified to near homogeneity using successive heparin, Ni-NTA, and size exclusion chromatographies as described in the materials and methods section. The ~42 kDa protein eluted as a monomer as confirmed by gel filtration chromatography (Figures 1(A and B)). The purified protein was assessed for its ability to extend a 15-mer DNA primer annealed to a 27-mer template DNA. DNA polymerases generally

utilize a two-metal ion binding mode of catalysis.[27] Using site-directed mutagenesis, we replaced D_{201} and D_{203} with alanine and purified mvPolX DID-AIA mutant (AIA) (Figure 1(B)). As expected, the AIA mutant showed a reduction in the DNA polymerase activity as compared to the wild-type protein (Figure 1(C and D)) confirming that these residues are essential for mvPolX function. We also tested the DNA-dependent RNA polymerase activity of mvPolX by replacing dNTPs with NTPs in the polymerization assays. Our results show that mvPolX can utilize NTPs for polymerization although with reduced efficiency as compared to dNTPs and therefore cannot form the full-length extension products (Figure 1(E)).

Further, to understand the effect of different divalent cations on enzyme processivity, we performed the DNA polymerase assay using varying concentrations of $MgCl_2$ and $MnCl_2$. The reactions lacking either of the two metal ions showed only residual polymerase activity as compared to those reactions containing them, thereby reaffirming the dependence of mvPolX on divalent cations (Supplemental Figures S2A and B). At concentrations of up to 5 mM, both Mg^{2+} and Mn^{2+} showed similar amount of full-length extension products. However, a significant difference was observed at higher concentrations (10 and 20 mM) of both cations with Mg^{2+} supporting better catalysis as evident by the formation of more full-length extension products (Figure 1(F and G)). Hence, 10 mM of $MgCl_2$ was used for all further assays.

Bypass of 8-oxo-guanine by mvPolX

To check the ability of mvPolX to bypass the 8-oxo-dG modification, we performed the full-length extension assays using mvPolX with 50-mer templates containing either an unmodified guanine

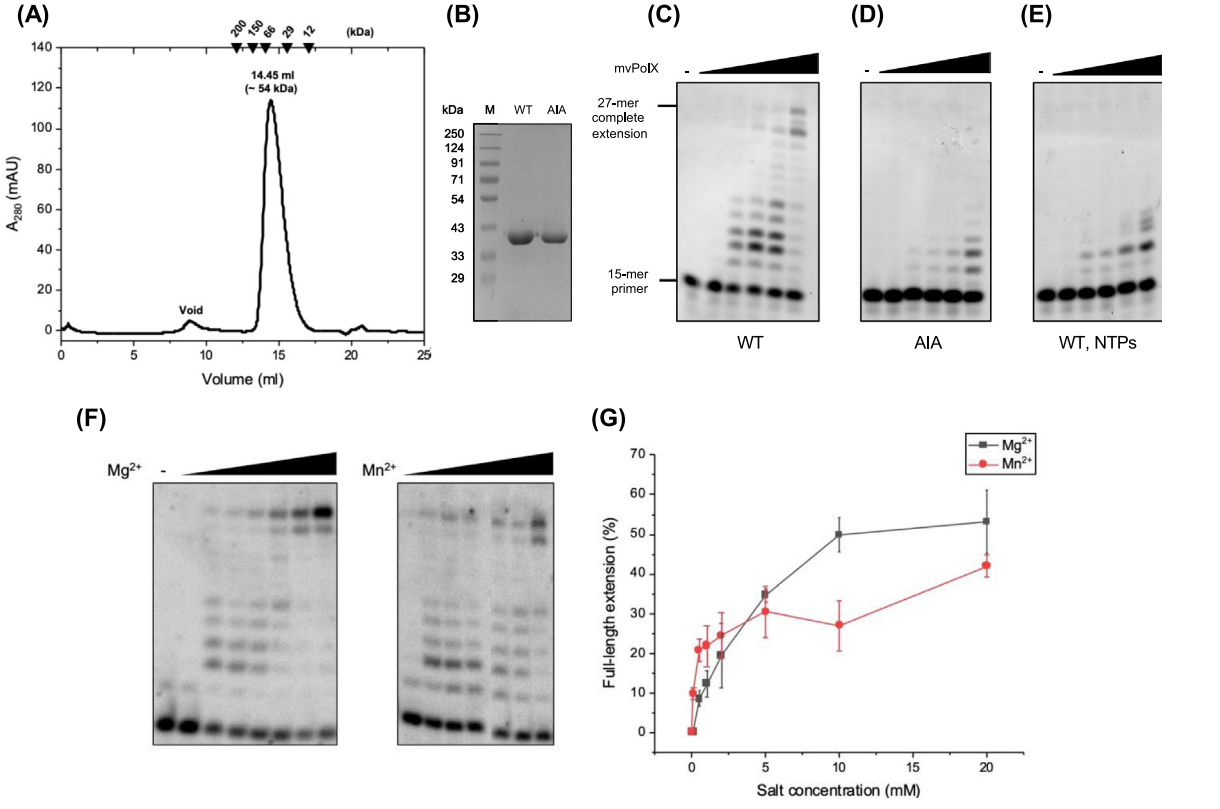


Figure 1. Purified recombinant mvPolX demonstrates DNA-dependent DNA polymerase activity. (A) Chromatogram showing the elution of mvPolX from a Superose-12 10/300 gel filtration column. The filled triangles indicate the elution volumes of standard proteins. mvPolX eluted at ~14.45 ml which corresponded to ~54 kDa and is close to its expected monomeric M_r of 43 kDa. **(B)** A 10% SDS-PAGE showing the purified wildtype (WT) and mutant (DID-AIA) mvPolX. **(C)** and **(D)** Show the DNA polymerization activity of mvPolX wildtype and AIA mutant, respectively. **(E)** RNA polymerization activity of mvPolX on a primer-template DNA. All reactions were performed using the polymerase buffer supplemented with 5 mM $MgCl_2$ as the cofactor and incubated at 30 °C for 10 min. Varying concentrations (No enzyme, 10, 50, 100, 200 and 500 nM) of the mvPolX wildtype or mutant enzyme were used. **(F)** A comparative analysis of Mg^{2+} and Mn^{2+} as cofactors for the DNA polymerase activity of mvPolX. Varying concentrations of both metal ions (No metal ions, 0.1, 0.5, 1, 2, 5, 10 and 20 mM) were used and reactions were incubated at 30 °C for 10 min. **(G)** Quantification of the full-length products formed in **(F)**. The error bars indicate standard deviation.

or the 8-oxo-dG modification at the 16th position (Figure 2(A)). Extension of the 15-mer primer beyond the unmodified or modified base was monitored by stopping the reactions at various timepoints and visualizing the products on a 20% urea-PAGE as described in the materials and methods. Our results show that mvPolX efficiently extends the primer beyond the unmodified guanine and nearly forms the full-length (50-mer) product (Figure 2(B)). On the contrary, while extending the 8-oxo-dG substrate, mvPolX incorporates several nucleotides beyond the modified base but fails to form the full-length product (Figure 2(C)). Additionally, a pause at the 16th and 17th nucleotide positions indicates low processivity of mvPolX while extending beyond the 8-oxo-dG modification (Figure 2(C)).

To assess the fidelity of DNA synthesis by mvPolX, we performed single nucleotide incorporation experiments using a primer-template substrate. The addition of individual deoxyribonucleotides across the template nucleotide (dG) was studied. mvPolX strictly incorporated only the complementary nucleotide dCTP across the template dG (Figure 2(D)). The addition of the incorrect nucleotide dCTP across the second template base dT, indicated by the "+2" product was observed only at large molar excess of dCTP (40 μM) (Figure 2(D)). Generally, Mn^{2+} ions tend to enhance polymerase activity by compromising with its fidelity leading to the incorporation of non-complementary nucleotides across the template base.[28] In the case of mvPolX however, Mn^{2+} did not induce misincorporation across the template dG (Supplemental Figure S3). Overall, the use of either Mg^{2+} or Mn^{2+} ions as metal cofactors did not significantly alter the fidelity of mvPolX while incorporating nucleotides across a template dG.

Family-X polymerases, β and λ have been shown to bypass modified DNA bases like 8-oxo-dG, especially during the MUTYH glycosylase initiated BER.[29] Therefore, we also tested the ability of mvPolX to bypass the 8-oxo-dG modified base in the template DNA (Figure 2(E)). The bypass reactions were performed using varying concentrations of the individual deoxyribonucleotides. mvPolX efficiently bypassed the 8-oxo-dG template using dATP probably due to the pseudo-complementarity of the incoming dATP with the syn-conformation of the 8-oxo-dG in the template, which mimics a thymine residue.[30] The enzyme also efficiently added dATP across the second template base dT. mvPolX did not support the addition of dTTP or dGTP across the 8-oxo-dG template (Figure 2(E)). Bypass of the lesion also occurred *via* the addition of the complementary nucleotide dCTP (Figure 2(E)). Synthesis across the 8-oxo-

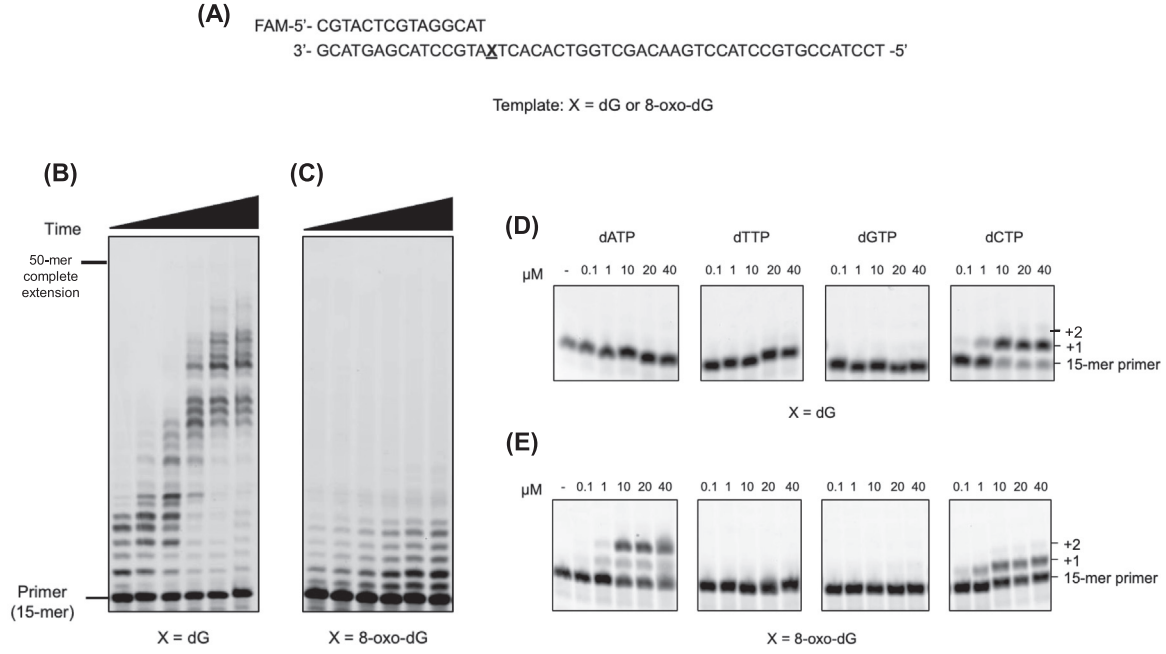


Figure 2. mvPolX**can bypass a template 8-oxo-dGresidue. (A)** A schematic representation of the substrate used for assessing the fidelity and translesion synthesis activity of mvPolX. The template nucleotide indicated in bold and underlined (X) is either dG or 8-oxo-dG. **(B)** and **(C)** Extension of the primer beyond the dG and 8-oxo-dG template nucleotides, respectively. The respective substrates were incubated with all 4 dNTPs and 100 nM of mvPolX and aliquots were taken at 2, 5, 10, 30, 60 and 90 min after initiating the reaction. The products formed were separated on a 20% urea-PAGE. **(D)** and **(E)** Single nucleotide incorporation against a template dG or template 8-oxo-dG, respectively. The concentrations of individual dNTPs is indicated at the top. These assays were performed using 100 nM mvPolX and incubated at 30 °C for 5 min.

dG modification was also supported by Mn^{2+} (data not shown). In summary, our results show that mvPolX selectively incorporates the complementary nucleotide across an unmodified base while it uses dATP and dCTP to bypass the 8-oxo-dG in a template DNA.

mvPolX possesses gap filling and strand displacement activity

Eukaryotic homologs of mvPolX like human pol β often act on gapped DNA substrates while participating in the BER pathway. To test the ability of mvPolX to utilize such substrates, we compared the polymerase activity on a primer-template (P-T) DNA with a gapped DNA substrate. The gapped template included the usual P-T DNA with a downstream blocking primer annealed to the template strand thereby producing either a 1 or a 5-nucleotide gap (see schematics in Figure 3). The blocking primer (indicated in red) either contained a 5′-hydroxyl or a 5′-phosphate group and was always used in two-fold molar excess. The P-T DNA serves as a control to measure the reduced processivity, stalling, or pause sites that may be formed due to the blocking primer. Efficient polymerase activity was observed on the P-T DNA with no stalling or pause sites (Figure 3(A)). With one nucleotide gapped template, mvPolX was able to bind and extend the primer beyond the gap. At shorter incubation times (2 and 5 min), a comparable pause site was observed in substrates with the 5′-OH and the 5′-phosphate groups (Figures 3(B and C)). Although the primer is extended beyond the gapped site, full-length extension does not occur and products with shorter length are formed. While utilizing the 5-nucleotide gap substrate, a pause site at the 20th nucleotide was formed possibly due to the stalling of the enzyme. This pause site disappeared after about 30 min of incubation and completely extended products are formed that are similar to the P-T unblocked DNA (Figure 3(D and E)). In summary, our results conclusively show that mvPolX possesses gap-filling and strand displacement activities.

Purification of mvUDG

The recombinant Mimivirus uracil DNA glycosylase (mvUDG) protein [gene product (gp) 271 encoded by gene L249] was purified using affinity and gel filtration chromatographies. The elution volume of the purified protein corresponded to ~73 kDa suggesting mvUDG protein to be a dimer (Figure 4(A)). The purified mvUDG protein (Figure 4(B)) without the tag was used for all biochemical studies. It is worth noting that a previous study found the mvUDG protein to be in the monodisperse state with the molecular mass being higher than the expected size of the monomer.[31] Further analysis using SEC-MALLS

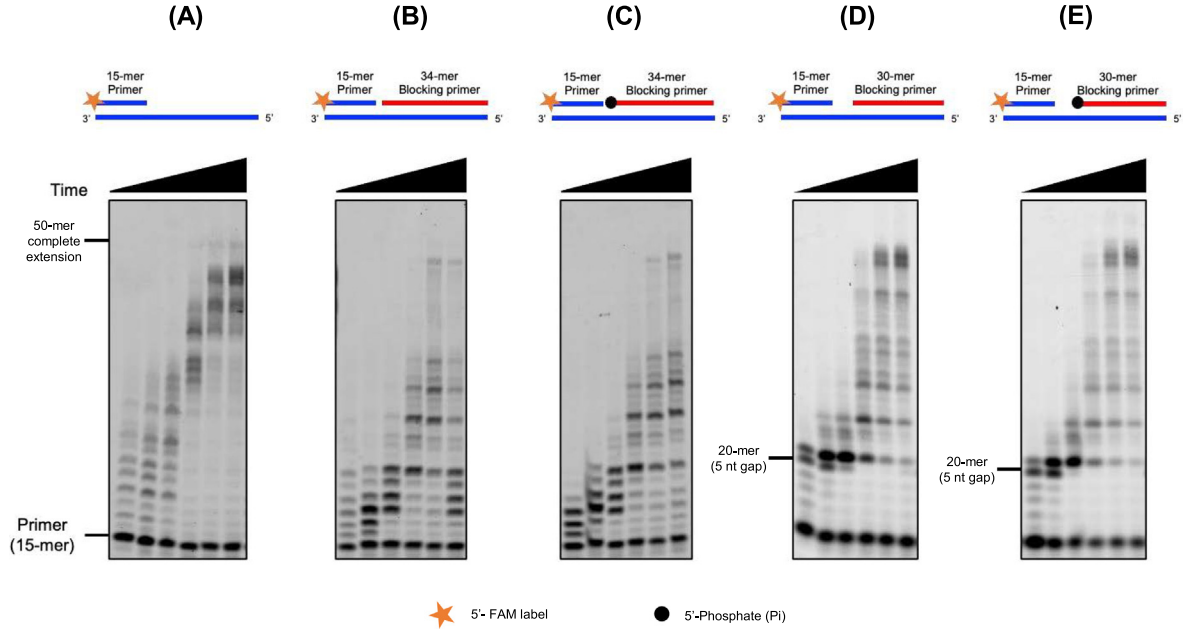


Figure 3. mvPolX **can perform gap-filling and strand displacement. (A)** DNA polymerization by mvPolX on a primer-template DNA. Gap-filling and strand displacement by mvPolX on a template containing a downstream blocking primer with 5′-OH **(B)** and 5′-phosphate **(C)**. DNA polymerization by mvPolX on a primer-template DNA with a downstream primer with 5 nucleotide gap with either 5′-OH **(D)** or 5′-phosphate **(E).** The respective substrates were incubated with a mixture of all 4 dNTPs and 100 nM of mvPolX and aliquots were taken at 2, 5, 10, 30, 60 and 90 min after initiating the reaction. All reactions were incubated at 30 °C and the products formed were separated on a 20% urea-PAGE.

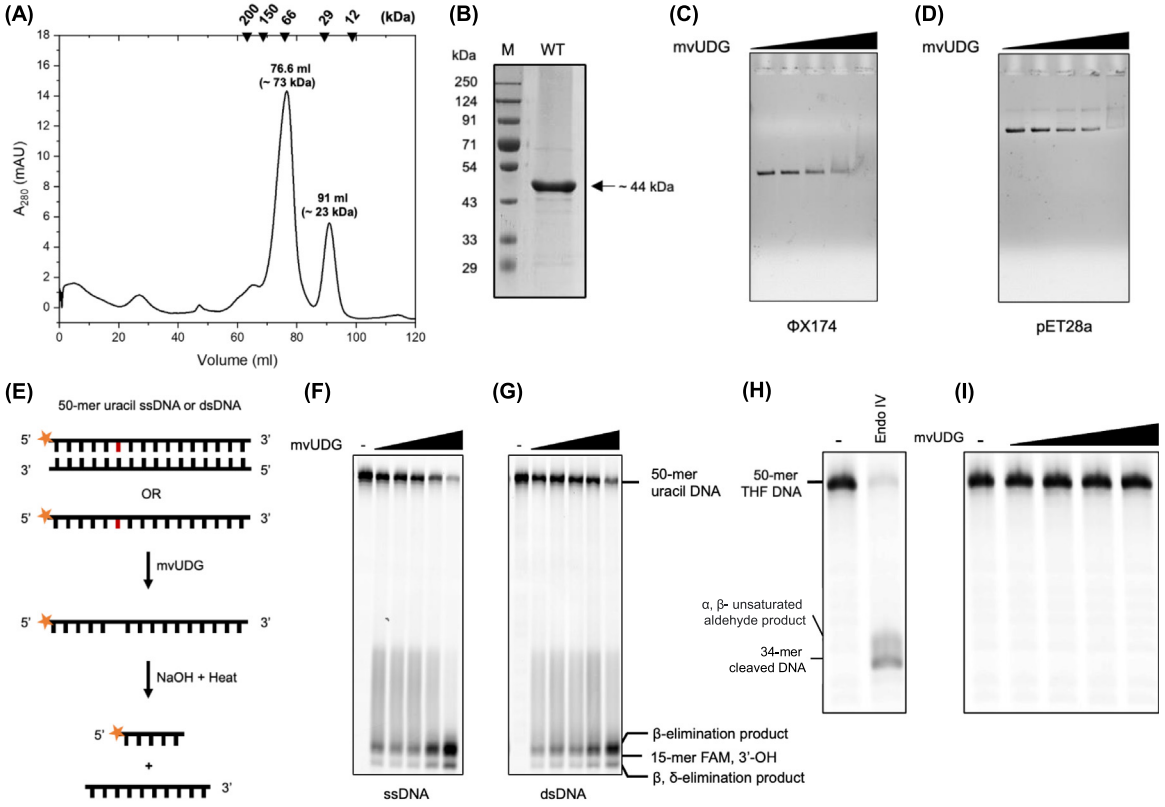


Figure 4. mvUDG**is a mono-functional uracil glycosylase. (**A) Chromatogram from 16/600 Superdex 200 column showing the elution of mvUDG as a dimer (76.6 ml) and the thioredoxin tag (91 ml). The filled triangles indicate the elution volumes of standard proteins. (B) A 10% SDS-PAGE gel showing the purified mvUDG corresponding to ∼44 kDa. (C) and (D) Binding of mvUDG to ΦX174 circular ssDNA and pET28a circular plasmid dsDNA, respectively. Approximately 30 ng of respective DNA (∼1 nM pET28a vector or 2 nM ΦX174 ssDNA) was incubated with either 200, 500, 1000 or 2000 nM of the recombinant mvUDG enzyme at 30 °C for 30 min and the products were separated on a 0.8% agarose gel. (E) A schematic showing the assay used for testing the activity of mvUDG on a linear ssDNA or dsDNA. (F) and (G) 20% urea-PAGE showing the excision of uracil from ssDNA and dsDNA by mvUDG. A 5′-FAM labelled primer with a 3′-OH end acts as a marker to distinguish between the two products formed by mvUDG. Varying concentrations of the enzyme (No enzyme, 20, 50, 100, 200, and 500 nM) were incubated with ss or dsDNA at 30 °C for 30 mins. The reactions were stopped by heating in the presence of alkali. (H) AP-lyase activity of *E. coli* endonuclease IV enzyme. (I) Testing the AP-lyase activity using 50, 100, 200, and 500 nM of mvUDG.

showed that the protein is a monomer in solution. [31]

Uracil glycosylase activity of mvUDG

A previous study by Kwon et al. (2017) had identified the key catalytic residues of mvUDG. We reconfirmed this finding using sequence analysis with known proteins (Supplemental Figure S4). Kwon et al., also showed that mvUDG could only act on ssDNA and fails to excise uracil from a dsDNA substrate.[31] We therefore began our study by understanding the ability of the purified recombinant mvUDG protein to bind to ssDNA and dsDNA. Our results show that mvUDG has the ability to bind circular ssDNA (ΦX174) as well as dsDNA (pET28a vector) leading to a shift in the mobility of DNA when separated on an agarose gel (Figures 4(C and D)). Hence, we checked the ability of mvUDG to excise uracil from linear ssDNA as well as dsDNA substrates. We utilized a linear ssDNA with a 5′-FAM label and uracil at the 16th position as a substrate for testing the uracil glycosylase activity. This DNA was annealed to a complementary 50-mer ssDNA to produce dsDNA containing uracil. The respective substrates were incubated with varying concentrations of mvUDG and the reactions were stopped by the addition of 0.1 mM NaOH and heating to 95 °C to cleave the abasic site created by mvUDG (Figure 4(E)). With increasing concentrations of the mvUDG enzyme, a decrease in the amount of substrate and a corresponding increase of the two products was observed. The major product formed is due to the β-elimination reaction, whereas the β, δ-elimination product forms the minor product (Fig-

ure 4(F)). Similar products are formed when mvNEIL1 and 2 enzymes excise bases from DNA.[16] Contrary to the previous report (Kwon et al., 2017), mvUDG can excise uracil from a dsDNA substrate and forms similar end products as those seen for ssDNA (Figure 4(G)). As mvUDG acts on dsDNA substrate, we hypothesized that Mimivirus may utilize mvUDG initiated BER pathway for the removal of uracil from viral DNA. Such a pathway would require the subsequent action of other enzymes such as apurinic-apyrimidinic endonuclease, polymerase, and ligase on the damaged DNA.

The two other DNA glycosylases encoded by Mimivirus (NEIL1 and NEIL2) are bifunctional in nature and therefore also cleave at abasic sites.[16] Therefore, we also tested the AP-lyase activity of mvUDG using an AP site-containing DNA. While the commercially available endonuclease IV cleaves the abasic site resulting in the formation of two products (Figure 4(H)), no such cleavage of the AP site was observed when using the mvUDG enzyme (Figure 4(I)). This result confirms that mvUDG is a monofunctional uracil DNA glycosylase.

Purification of mvAPE

The mvAPE protein [gene product (gp) 324 encoded by gene R296] was purified by successive Ni^{2+} affinity and gel filtration chromatographies. During gel filtration chromatography, the protein predominantly eluted as a monomer (~38 kDa, Figure 5(A)).

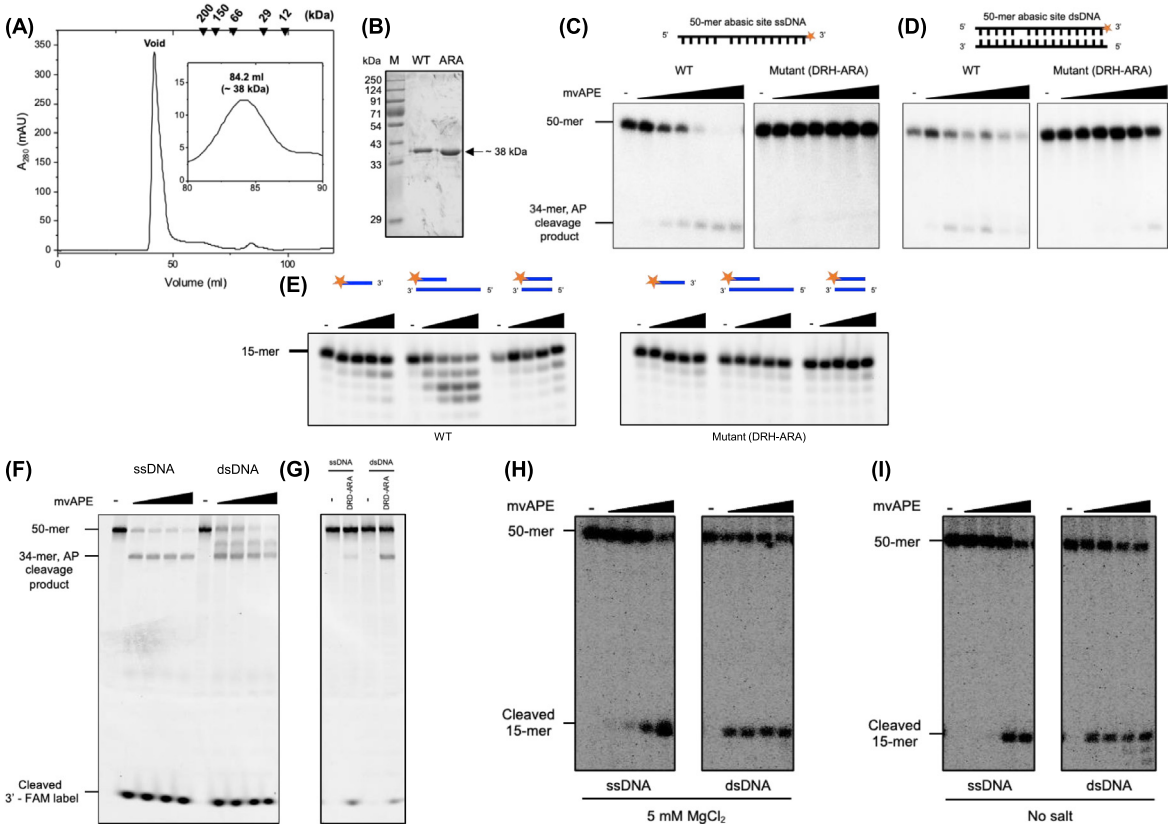


Figure 5. mvAPEpossesses AP-**endonuclease and 3′-5′ exonuclease activities.** (A) Chromatogram from 16/600 Superdex 200 column showing the elution of mvAPE protein. The inset shows the elution of monomeric mvAPE at ~84 ml. (B) A 10% SDS-PAGE showing the purified mvAPE wildtype (WT) and DRH-ARA mutant (ARA). The M_r of the monomeric mvAPE is ~38 kDa. (C) and (D) show the attempted AP endonuclease assay on 3′-FAM labelled single and double stranded DNA respectively, using either the WT or ARA mutant enzyme. The reactions were either devoid of or contained 20, 50, 100, 200, 500 or 1000 nM mvAPE enzyme. (E) Testing the 3′-5′ exonuclease activity of mvAPE using the WT and ARA mutant enzymes. The schematics show the type of DNA substrate used: ssDNA (left), recessed 3′-DNA (middle) and dsDNA (right). AP-lyase activity on ss and dsDNA by mvAPE WT (F) and ARA mutant (G). Either 100, 200, 500 and 1000 nM of the WT protein or 500 nM of the mutant were used for these assays and reactions were incubated at 30 °C for 30 min. (H) and (I) Show the AP-lyase activity of mvAPE in the presence and absence of $MgCl_2$. The reactions contained a 5′-radiolabelled 50-mer oligonucleotide with an AP site at the 16th position and 20, 50, 100 or 200 nM of mvAPE.

Additionally, some protein also eluted in the void volume, probably as aggregates or DNA bound form (Figure 5(A)). For the experiments presented in the current study, only the monomeric form of the protein was used.

mvAPE is an AP endonuclease and 3′-5′ exonuclease

We studied the AP-lyase activity of mvAPE using both ssDNA and dsDNA substrates containing an abasic site and a 3′-FAM label (Figure 5(C and D)). An interesting observation made during the AP-lyase assay was that upon treating the labelled DNA with the purified wildtype enzyme, a 34-mer cleavage product corresponding to the AP-cleavage site could be seen (Figure 5(C and D), panel WT). Interestingly, the amount of the product released and the leftover substrate were found to be lesser than the original amount of substrate added to the reaction. The observed results could be possible if the DNA substrate was undergoing degradation during the course of the reaction. However, the DNA substrates treated with the active site mutant mvAPE did not show any reduction in the amount (Figure 5(C and D), panel Mutant). This confirms that the degradation of DNA observed is probably due to the mvAPE enzyme. We hypothesized that the reduction in total DNA is probably due to the removal of the 3′-FAM label by 3′-5′ exonuclease-like activity of mvAPE.

To test the 3′-5′ exonuclease activity of the enzyme, we incubated the enzyme with three different forms of DNA labelled at the 5′-end. The DNA used included a 5′-end labelled ssDNA, a partially dsDNA with a 3′ recessed end and a blunt dsDNA. Of the three substrates, only the

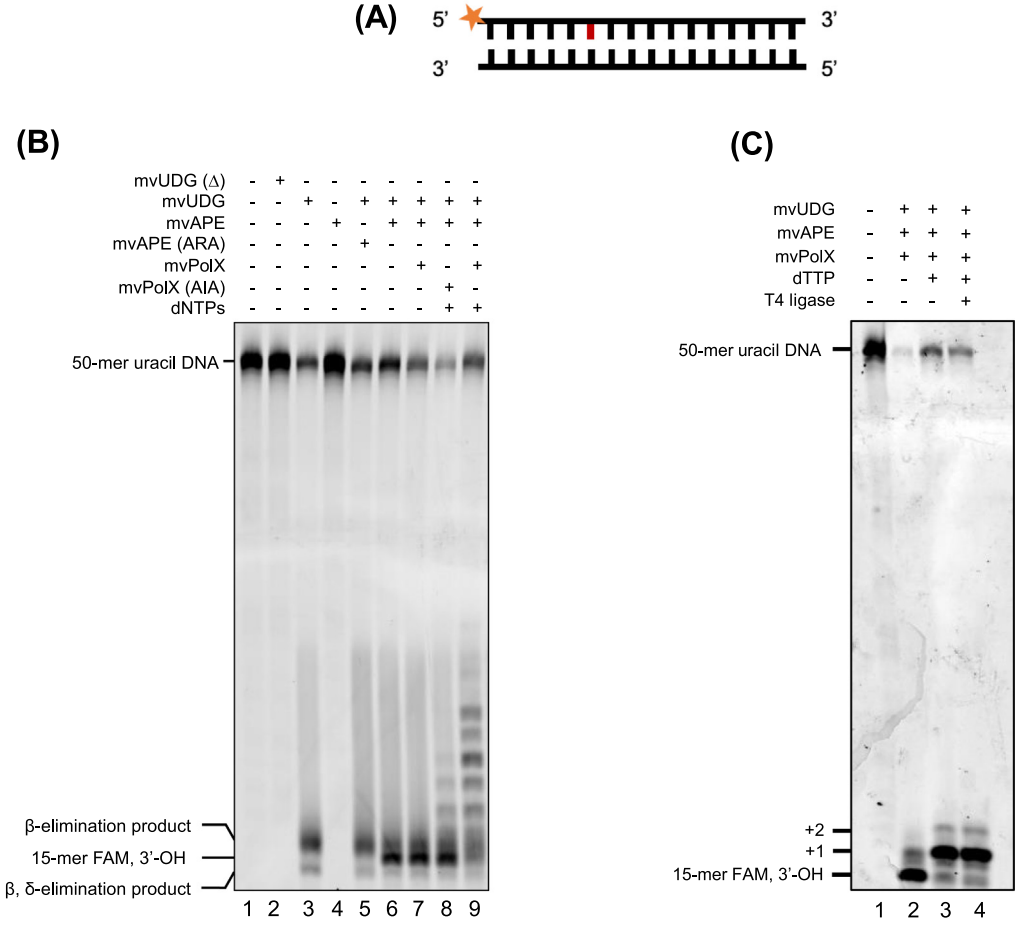


Figure 6. A "one-pot" assay demonstrates effective reconstitution of mvBER. (A) The 5'- FAM labelled dsDNA substrate containing uracil used for reconstituting the viral BER *in vitro*. **(**B) Step-wise *in vitro* reconstitution of long-patch mvBER using the purified mvUDG, mvAPE and mvPolX. mvUDG(Δ) indicates the catalytically null form of the mvUDG enzyme produced by heating the enzyme at 95 °C for 10 mins. Reactions contained 500 nM of 5'-FAM labelled uracil dsDNA with 100 nM mvUDG, 50 nM mvAPE, and 100 nM mvPolX in the indicated combinations. (C) Reconstitution of mimiviral short-patch BER. The reactions were similar to long-patch but contained 10 μM of dTTP instead of all four dNTPs in the presence or absence of T4 DNA ligase.

DNA with a recessed 3′-end was utilized efficiently and products of size lesser than the substrate were clearly seen thereby confirming the exonuclease activity of the mvAPE enzyme (Figure 5(E), panel WT). Using multiple sequence alignment with known proteins, we identified the putative metal binding residues in mvAPE to be $D_{252}RH_{254}$ and mutated them to ARA (Supplemental Figure S5). The mvAPE mutant, as expected, does not cleave the substrate (Figure 5 (E), panel Mutant). The mutation of DRH to ARA impeded both the AP-endonuclease and the 3′-5′ exonuclease activities (Figure 5(C–G)). We also trapped the cleaved product containing the 3′-FAM label by separating it on a 20% urea-PAGE for a shorter duration than usual (90 min instead of 180 min at a constant power of 30 W). We indeed observed that the cleaved 3′-FAM label is trapped near the end of the gel (Figure 5(F)). This product is observed with both ssDNA and dsDNA substrates and is significantly reduced with the ARA mutant enzyme (Figure 5(G)). The significance of the exonuclease activity of mvAPE *in vivo* remains to be studied.

Finally, we checked the dependence of the mvAPE on divalent cations while performing cleavage at an AP site. We used a 5′-radiolabeled ssDNA or dsDNA containing a tetrahydrofuran (THF) site at the 16th position as a substrate. The reactions were performed using reduced ratio of enzyme to DNA and a shorter incubation time of 15 min to keep the 3′-5′ exonuclease activity to a minimum. All reactions were performed either in the presence or absence of 5 mM $MgCl_2$. Our results clearly show that the AP-endonuclease activity of mvAPE is independent of divalent cations (Figure 5(H and I)). In summary, our study shows that mvAPE is capable of cleaving abasic sites in ssDNA or dsDNA and additionally also possesses the 3′-5′ exonuclease activity.

In vitro reconstitution of the mvBER pathway

Finally, to check if all the three purified BER enzymes can work in unison, we performed a "one-pot" assay. All three enzymes were incubated with uracil containing dsDNA (Figure 6 (A)) and the ability of mvUDG, mvAPE, and mvPolX to repair the uracil in the substrate DNA was examined. To track the reaction products formed, the uracil DNA was labeled at the 5'-end using 6-FAM (see materials and methods). Despite our several attempts we could not purify the mvUDG catalytically null mutant and hence heat activated wild type mvUDG was used. The heat inactivated enzyme showed complete loss of glycosylase activity as compared to the wild type protein (Figure 6(B), lanes 2 and 3). Further, in the absence of the mvUDG enzyme, mvAPE is unable to act on the uracil-containing DNA and the 50-mer uracil DNA substrate remains intact (Figure 6(B), lane 4). Next, upon addition of the mvAPE mutant (ARA) and the wild type enzyme to the reactions containing the wild type mvUDG enzyme, formation of two distinct products is visible. The former reaction produces a product similar to that observed in the case of mvUDG alone as mvAPE (ARA) does not have the AP-endonuclease activity (Figure 6(B), lanes 5 and 6). The latter reaction shows formation of a product equivalent to a 15-mer oligonucleotide which is the expected product of coupled glycosylase and endonuclease reactions (Figure 6(B), lane 6). This product is expected to possess a 3'–OH end that can be extended by a DNA polymerase when supplemented with a mixture of all four dNTPs. We expected that after the removal of uracil by mvUDG and processing of the abasic site by mvAPE, mvPolX would perform gap-filling synthesis and displace the downstream strand using its strand displacement activity. This activity could be traced by the formation of the polymerase extension products beyond the "+1" insertion site. If no dNTPs are provided the 3'–OH group remains unextended (Figure 6(B), lane 7). Addition of the mvPolX-AIA mutant to the above reaction leads to limited extension of the 3'–OH end as the AIA mutant has reduced polymerase activity as compared to the corresponding wild type enzyme (Figure 6(B), lanes 8 and 9). These results clearly show that the three Mimivirus enzymes selected for the present study can potentially perform the long patch BER.

In the next set of reactions, we added all three wild type enzymes to the reaction mixture along with only dTTP (instead of all 4 dNTPs) as the next template base in the dsDNA substrate is a "dA". We observed that in the presence or absence of the T4 DNA ligase enzyme, the "+1" product formed due to the addition of dTTP is not ligated to the downstream 34-mer DNA strand (Figure 6(C)). The "+1" product accumulates and is also partly extended to form the "+2" product. Thus, the results indicate that the mvUDG, mvAPE and mvPolX can together excise the uracil base in the DNA and incorporate a nucleotide across the template but they cannot form the full-length repaired DNA product to complete short patch BER.

Discussion

In continuation of our previous work regarding DNA damage tolerance in Mimivirus,[19] the present study was aimed at understanding the mechanisms of DNA damage tolerance and repair in Mimivirus. Here, we identified and biochemically characterized three Mimivirus proteins, namely, uracil DNA glycosylase (mvUDG), AP-endonuclease (mvAPE), and a family X polymerase (mvPolX).

We purified mvPolX and demonstrated its ability to extend a primer annealed to a template DNA (Figure 1(C)). The dependence of the enzyme on

divalent cations was confirmed by mutating the putative cofactor binding residues, D_{201} and D_{203} to alanine (Figure 1(D) and Supplemental Figure S2). The DID-AIA double mutant shows reduced polymerase activity (Figure 1(D)). Polymerases are known to utilize a variety of divalent cations for the nucleotidyl transferase reaction resulting in differing fidelity and processivity.[32] Similar to previous reports for polymerase β,[33–35] mvPolX also utilized ribonucleotides as substrates during its polymerase activity although with lesser efficiency (Figure 1(C and E)). This can be explained as mvPolX contains a conserved "sugar selector" tyrosine amino acid in the thumb subdomain (Y271) highlighted in green (Supplemental Figure S1). The sequence "YFTGS" present in polymerase β and λ correlates with their preference for dNTPs. Conversely, Pol μ and terminal transferase (TdT) have the sequence "GWTGS", losing the sugar selector tyrosine, and consequently incorporating ribonucleotides with high efficiency.[36] As seen in the alignment of mvPolX with polymerase β, λ, μ, and TdT, the sequence "YFTGP" is present in mvPolX which is congruent with the demonstrated preference for dNTPs (Figures 3(C and E)). As dNTPs are utilized more efficiently, NTPs may not be incorporated into the growing primer across an unmodified base; however, incorporation of NTPs across a damaged base may still occur as reported for polymerase β.[35,37] We also assessed the ability of mvPolX to bypass an oxidized base, 8-oxo-dG, in the template DNA. Akin to human polymerase β, mvPolX could bypass the lesion both in an error-free manner and *via* the incorporation of dATP as well as dCTP (Figure 2).[32,38] Additionally, mvPolX can also perform both gap-filling and strand displacement synthesis on various DNA substrates which may be essential for a viral short or long-patch base excision repair pathway (Figure 3).

The first step of the BER pathway is generally initiated by a DNA glycosylase protein that excises the modified base from the DNA. Sequence analysis and a previous report showed the presence of gene L249 which codes for a uracil DNA glycosylase (Supplemental Figure S3).[14,15,31] A previous study showed that mvUDG selectively excises uracil from a ssDNA and fails to act on dsDNA.[31] However, a mvUDG-like enzyme would be required to act on dsDNA as the Mimivirus genome is known to be a linear dsDNA. We therefore tested the binding as well as the uracil excision ability of mvUDG using ssDNA and dsDNA substrates. Our activity assays using a uracil-containing DNA revealed that mvUDG can bind to and excise uracil from both ss and dsDNA substrates (Figure 4). The two products formed by mvUDG upon treatment with hot alkali appear to be due to β-elimination and β, δ-elimination reactions.

Next, to examine whether mvUDG products can be further used by downstream viral enzymes, we purified the putative endonuclease IV protein, mvAPE, which shares limited sequence homology with other known AP-endonucleases but performs AP site cleavage on ssDNA and dsDNA (Supplemental Figure S5 and Figure 5(C-I)). Additionally, mvAPE also has a 3′-5′ exonuclease activity (Figure 5(E)). An AP-endonuclease from ASFV, pE296R protein, has been shown to possess 3′-5′ exonuclease, 3′-phosphatase, and 3′-diesterase activities in addition to AP site cleavage and has an essential role in the viral life-cycle.[39–41] We believe that mvAPE may also possess similar activities and may participate in a viral BER pathway.

Finally, we demonstrated the ability of all three viral proteins to work in unison by *in vitro* reconstitution of the mvBER to repair uracil-containing dsDNA (Figure 6). Despite all our attempts we were unable to conclusively demonstrate the dRP-lyase activity of mvPolX. The addition of ligase does not produce a full-length repaired DNA product probably owing to the lack of mvPolX to excise the dRP group on the downstream DNA strand (Figure 6(C)). mvPolX may thus aid in the repair of DNA during BER using its strand displacement activity and may form a critical part of the long patch BER (Figure 6(B)).

Interestingly, the Mimivirus polymerase X is packaged inside the viral particle along with a DNA glycosylase (L687), endonucleases NEIL1 (L315), NEIL2 (L720), and AP-endonuclease (R296), a FEN1 homolog (L386), and a DNA ligase enzyme (R303) which could together be useful for the removal of erroneously incorporated uracil or oxidized bases[16,23] (Table 1). These packaged proteins may perform base excision repair during the early phase of infection before initiation of DNA replication. Alternatively, as the expression profile of Mimivirus indicates expression of mvPolX, a bifunctional glycosylase (Gene L315), a putative AP-endonuclease (Gene L687) and a FEN1-like protein (Gene L386) during the late stages of infection, they could work in conjunction to complete a post-replicative DNA repair pathway.[16,42]

Our studies reported here add to the list of other previously characterized DNA repair proteins from the nucleo-cytoplasmic large DNA viruses. The mimiviral NEIL1 (Gene L315) and NEIL2 (Gene L720) proteins that are similar to human NEIL-1 and NEIL-2 proteins are the first examples of oxidative DNA base excision repair mechanisms in viruses.[16] Additionally, ASFV AP-endonuclease and polymerase X have also been shown to be involved in a viral BER pathway that prevents oxidative damage to viral genomic DNA.[39,41,43] More recently, another PolX fusion protein from *Phaeocystis globosa* virus has been shown to be a lyase, polymerase and ligase that can repair BER intermediates like an AP site.[44] As Mimivirus codes for a PrimPol, which can also bypass bulky modified

bases like N^2-Bn-dG,[19] as well as several components of the BER and MMR pathways, it appears to have robust DNA repair mechanisms.[14,15,21] NCLDVs like mimiviruses have been predicted to acquire DNA from the amoebal "melting pot" and the availability of repair proteins may aid in the maintenance of the newly acquired DNA. Such DNA accretion events, for instance, the accumulation of repeat domain-containing proteins (RDCPs), may provide viruses with additional mechanisms for host adaptation, thereby contributing to the viral fitness.[45] As the energy cost of viral DNA replication is minuscule, NCLDVs can probably afford to expand their genome which may ultimately aid in achieving greater physiological autonomy.[46] Thus, it is highly likely that Mimivirus and other related viruses have acquired DNA repair genes from the amoebal melting pot that may be useful for maintaining genomic integrity. Additionally, the low non-synonymous to synonymous substitution ratio of the *Megavirales* genomes and the central location of the viral replication and repair enzymes together point towards their fixation due to the essential roles.[45,47,48]

In summary, our study establishes the various biochemical activities of a family X polymerase, a uracil-DNA glycosylase and an AP-endonuclease from Mimivirus. We believe that mvPolX along with other proteins like mvUDG, mvAPE, mvPNKP, and mvLigase may play an indispensable role in the repair of damaged DNA and maintaining viral genomic integrity *via* any of the two BER sub-pathways (Figure 7). Further studies targeted towards knockdown of one or a combination of these repair proteins may aid in understanding their role and importance in maintaining the viral genome. Additionally, as demonstrated for Herpes Simplex Virus-1, it could be interesting to understand the interaction of mimiviral enzymes with the host amoebal proteins.[49,50]

Materials and Methods

Giant virus purification and Mass spectrometry (MS)

Acanthamoeba castellanii cells were cultured at 28 °C to achieve a monolayer. Giant virus (GV) particles were harvested by infecting *A. castellanii* cells with an MOI ~5. The cell lysate was

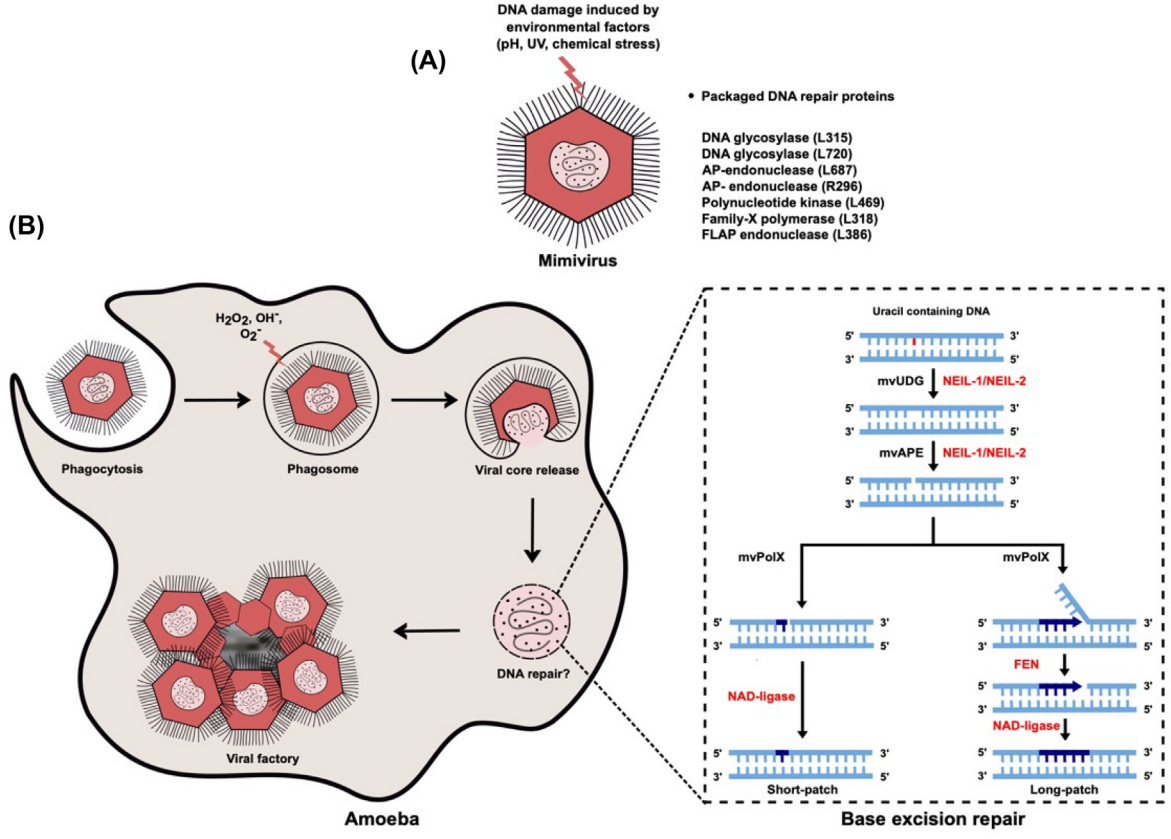


Figure 7. Hypothetical mimiviral DNA repair pathway. (A) A schematic representation of mimiviral particle and the packaged replication proteins. **(B)** A hypothetical pathway showing the repair of uracil containing DNA inside mimivirus replication factory. Mimivirus encodes proteins required for both short and long patch base excision repair. Proteins that were used in this study are shown in black, whereas those present in the virus with potential role in BER are shown in red.

subjected to centrifugation to remove the cell debris and supernatant was used for virus purification. *Acanthamoeba polyphaga* mimivirus (APMV), Mimivirus bombay (MVB), and Powai Lake megavirus (PLMV) were purified by centrifugation at 7900 *g* for 30 min and Marseillevirus (MaV) and Kurlavirus (KV) were centrifuged at 26,000 *g* for 90 min. Further purification was performed using sucrose density gradient centrifugation at 5000 *g* in a swinging bucket rotor. Purified GV particles were subjected to lysis before in-gel and in-solution trypsin digestion. Q-Exactive Plus-Orbitrap Mass spectrometer (Thermo Scientific) instrument and Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fischer Scientific) with easy nano-LC 1200 system were used for in-gel and in-solution digested samples, respectively. The data obtained from MS was analyzed using the MaxQuant and/ or the Proteome Discoverer software.

Expression and purification of Mimivirus polymerase X and its mutant

Mimiviral gene, L318 encoding a putative family X polymerase was PCR amplified from the purified mimiviral genome using appropriate primers and cloned into pET28a vector between *NcoI* and *HindIII* restriction enzyme sites using the primers (Forward: 5′ -GCCACCATGGGGAT GAATTCAAAAATTATCGAAC −3′ and Reverse: 5′ -CGCGAAGCTTGTGTATTTTTCCTGAACTGT −3′). Two conserved aspartates, D_{201} and D_{203} identified in the putative DNA polymerization motif (DID) were mutated to alanine (AIA) by site-directed mutagenesis using the primers: (Forward: 5′-GGTAAAAAAACTTCTGGTGCTATCGCTGTTC TTATGTATCATC-3′ and Reverse:

5′-GATGATACATAAGAACAGCGATAGCACCA GAAGTTTTTTTACC-3′). The recombinant plasmids (wildtype and mutant) were confirmed using sequencing and transformed into *E. coli* BL21 (DE3) Rosetta cells for expression of the protein.

Wild-type mimiviral polymerase (mvPolX) and its mutant (DID-AIA) were purified using the following protocol: One-liter culture of *E. coli* BL21 (DE3) Rosetta cells were grown to an O.D. of 0.6–0.8 and expression of mvPolX was induced by the addition of 0.3 mM isopropyl-β-D-thiogalactopyrano side (IPTG). Following the induction at 16 °C for 12 h, cells were harvested, resuspended in Buffer R [50 mM Tris-Cl pH-7.4, 10% glycerol, 500 mM NaCl] and lysed using ultrasonication. The cell-free lysate containing soluble protein was diluted using buffer D [50 mM Tris-Cl pH-8.0 and 10% glycerol] to maintain the NaCl concentration in the lysate at 400 mM. The diluted cell lysate was centrifuged at ∼22,000 *g* for 30 min at 4 °C, filtered, and loaded onto a HiTrap Heparin HP column (GE Healthcare) equilibrated with Buffer HA [50 mM Tris-Cl pH-7.4, 10% glycerol, and 400 mM NaCl]. The column was washed with 20 CV of Buffer HA and protein was eluted from 0.4-1 M NaCl elution gradient with Buffer HB [50 mM Tris-Cl pH-8.0, 10% glycerol, and 1 M NaCl]. The elution fractions were analyzed on a 10% SDS-PAGE gel. The peak fractions were pooled and loaded on a HiTrap chelating HP column (GE Healthcare) equilibrated using Buffer A [50 mM Tris-Cl pH-7.4, 10% glycerol, 500 mM NaCl, and 20 mM imidazole]. The column was washed with 10 CV Buffer A and the protein was eluted with an imidazole gradient produced using buffer A and B [50 mM Tris-Cl pH-7.4, 10% glycerol, 500 mM NaCl, and 600 mM imidazole]. The peak fractions were subjected to 12% SDS-PAGE analysis and pure fractions were pooled, concentrated to 0.5 ml, and further purified by HiLoad Superose 12 10/300 GL gel filtration column (GE Healthcare) connected to the ÄKTA purifier pre-equilibrated with Buffer G [50 mM Tris-Cl pH-8.0, 10% glycerol, and 500 mM NaCl]. The elution fractions were checked for protein purity by SDS-PAGE analysis. Pure fractions were pooled and concentrated to 1–2 mg/ml and stored at −80 °C.

Primer extension assays

A 15-mer 5′-FAM labelled DNA primer was annealed to a 27-mer DNA template (Supplemental Table S1, Sequences 4 and 5) to generate a primer-template complex used to test the primer extension activity of mvPolX. Reactions (10 μl), contained buffer P (10 mM Tris-Cl, 1 mM DTT, 0.1 mg/ml BSA, and 5% glycerol), 50 nM primer-template DNA, 5 mM $MgCl_2$, varying concentrations of the wildtype or the DID-AIA mutant enzyme, and 80 μM of dNTPs. Reactions with the wildtype and the mutant enzyme were incubated at 30 °C for 10 min unless otherwise specified. Similar reactions were performed using 80 μM of NTPs as well. Varying concentrations of the two divalent cations (Mg^{2+} and Mn^{2+}) were also used to study the effect of metal ions on the activity of the polymerase. All reactions were stopped by adding 2X loading dye (80% formamide, 20 mM EDTA and 1 mg/ml bromophenol blue). Samples were heated at 95 °C for 5 min and separated using 20% denaturing PAGE (7 M urea) as previously described.[19] The reaction products were visualized using a Amersham Typhoon imager (Cytiva). Full-length extension products formed were quantified using ImageJ. The data points shown represent mean of duplicate reactions followed by standard deviation.

Bypass of a modified base by mvPolX

The ability of mvPolX to bypass a modified nucleotide in the template base was studied using the standing-start assay. 50-mer DNA template with either 8-oxo-deoxyguanosine (8-oxo- dG) or unmodified guanine at the 16th position was used

(Supplemental Table S1, Sequence 4 or 7). The above template was annealed to a 5′-FAM labelled 15-mer primer (Supplemental Table S1, Sequence 4). Reactions were incubated with a mixture of all four dNTPs (80 μM) at 30 °C and aliquots of equal volume were drawn at the indicated time points to assess the bypass of the modified nucleotide.

Further, to understand the fidelity of bypass, a single nucleotide bypass assay was also performed by replacing the mixture of dNTPs with all 4 individual dNTPs. The concentration of individual dNTPs was varied from 0.1 to 40 μM and the reactions were incubated for 5 min, terminated, and then separated and visualized as described previously.

Gap-filling and strand displacement assay

A gap-filling assay was used to test whether the mvPolX can bind to a gapped DNA substrate and extend it by strand displacement. The primer-template combination used was identical to the one used for bypass assay except it also contained a downstream 30-mer or a 34-mer blocking primer thereby producing a 5 and 1 nucleotide gap, respectively (Supplemental Table S1, sequences 4 and 5 or 6). The three oligonucleotides i.e. The 50-mer template, the 5′-FAM labelled primer and the downstream primer were annealed in a molar ratio of 1:1:2. For annealing, the downstream primers were either used directly (5′-OH group) or phosphorylated at the 5′-end using polynucleotide kinase (Thermo Fisher Scientific). The ability of mvPolX to perform strand-displacement synthesis was assessed by comparing extension on an unblocked template. The assay mixture contained 500 nM of the respective templates, 100 nM mvPolX, 80 μM of dNTPs, and 10 mM $MgCl_2$. Reactions were incubated at 30 °C for indicated incubation times, stopped, separated on a 20% PAGE (7 M urea) gel, and visualized similar to the primer extension assays.

Expression and purification of mimiviral uracil DNA glycosylase

The L249 gene from Mimivirus which codes for a uracil DNA glycosylase protein[31] was PCR amplified from mimviral genome using the primers (Forward: 5′ -AAGGCCATGGGGATGTCTAAGAA AA ATGTGG- 3′ and reverse: 5′ -GCGCAAGCTTG AGATCCCAAAAAATTGG- 3′) and cloned into pET32b vector using *NcoI* and *HindIII* restriction enzyme sites.

One-liter culture of *E. coli* BL21 (DE3) Rosetta cells were grown in the presence of appropriate antibiotic till O.D. 0.6 and protein expression was induced by the addition of 0.3 mM IPTG. Followed by induction at 16 °C for 12 h, cells were harvested by centrifugation, resuspended in Buffer A [50 mM Tris-Cl pH-7.4, 10% glycerol and 500 mM NaCl] and lysed using ultrasonication. The lysate was centrifuged at 22,000 *g* for 30 min at 4 °C and the supernatant containing soluble protein was applied to a HiTrap chelating HP column (GE Healthcare) equilibrated using Buffer A with 20 mM imidazole. The column was washed with 25 CV Buffer A and about 20 CV with 10% of buffer B (containing 600 mM imidazole). Thereafter, the protein was eluted with an imidazole gradient. The peak fractions were subjected to 12% SDS-PAGE analysis and the pure fractions were pooled, concentrated to 1 ml and further purified by HiLoad 16/600 Superdex 200 pg gel filtration column (GE Healthcare) connected to ÄKTA purifier which was pre-equilibrated with Buffer C [50 mM Tris-Cl pH-7.4, 10% glycerol, 500 mM NaCl]. The elution fractions were checked for protein purity by 12% SDS-PAGE analysis. mvUDG protein was further subjected to cleavage using 10 units of enterokinase (Merck Millipore) at 1U enterokinase per 100 μg protein as per manufacturer's recommendations. The enterokinase reaction was incubated at 4 °C for 14 h. The untagged protein was separated from the free thioredoxin, S tag and the histidine tag by gel filtration chromatography using a Superdex 200 pg gel filtration column. The eluted fractions were checked for purity using 12% SDS-PAGE and the pure fractions were concentrated to ∼0.5 mg/ml using vivaspin turbo centrifugal concentrators (Merck-Millipore). Protein concentration was estimated by Bradford's method and the protein was stored as aliquots at −80 °C.

Electrophoretic mobility shift assay for mvUDG

Indicated amounts of the purified proteins were incubated with ∼30 ng of circular ssDNA (ΦX174 phage genome) or dsDNA (pET28a vector). The reactions were incubated at 30 °C for 30 min and the DNA-protein complexes formed were separated on a 0.8% agarose gel.

Uracil removal assays

Uracil removal by mvUDG was tested using a 5′-FAM labelled DNA containing the uracil base at the 16th position (Supplemental Table S1, Sequence 8). The uracil DNA was annealed to a complementary 50-mer DNA to produce dsDNA with uracil (Supplemental Table S1, Sequence 4). Varying concentrations of mvUDG enzyme were incubated with 500 nM ssDNA or dsDNA. All reactions were incubated at 30 °C for 30 min. Positive control used was *E. coli* UDG (Thermo Fisher Scientific). The abasic site formed due to uracil removal was further cleaved by the addition of 0.1 mM NaOH (final concentration) and heating to 95 °C for 5 min. A 2X loading dye was added

and the samples heated again at 95 °C for 5 min before loading on a 20% urea-PAGE. The products were visualized using an Amersham Typhoon scanner (Cytiva).

Testing the AP-lyase activity of mvUDG

The AP-lyase activity was tested using a 3′-FAM labelled DNA containing an abasic/THF site at the 16th position (Supplemental Table S1, Sequence 9). Varying concentrations of the enzyme were incubated with this substrate for 30 min at 30 °C. The endonuclease IV enzyme from *E. coli* (NEB, M0304S) acted as a positive control. The resulting products were separated on a 20% urea-PAGE and visualized as described for uracil removal assays.

Expression and purification of mimiviral apurinic-apyrimidinic endonuclease and its mutant

The R296 gene was PCR amplified from mimviral genome using the primers (Forward: 5′ -CGCCGA ATTCATGCAATCGAATATAAATATTTG- 3′ and Reverse: 5′ -GGTGCTCGAGTTATTTAAGCCAT GATTTGAC- 3′) and cloned into pET28a vector using *EcoRI* and *XhoI* restriction enzyme sites. A double mutation (DRH to ARA) was introduced in the wildtype protein using mutant primers (Forward: 5′- GTTGGAATCAAAAGTTGCTAGAG CTGCTGATTTATGCTAC −3′ and Reverse: 5′- GTAGCATAAATCAGCAGCTC TAGCAACTTTT GATTCCAAC −3′). The recombinant plasmids were transformed into *E. coli* BL21 (DE3) Rosetta cells for protein expression. Purification of mvAPE and its mutant was similar to that of mvUDG protein except that buffer A contained 20 mM imidazole. Pure fractions obtained after gel filtration chromatography were pooled, concentrated and stored as aliquots at −80 °C.

AP site incision assays for mvAPE

The apurinic-apyrimidinic lyase activity of mvAPE was tested using 3′-FAM labelled 50-mer ssDNA containing an abasic site (dSpacer/ Tetrahydrofuran site) at the 16th base position (Supplemental Table S1, Sequence 9). This ssDNA was annealed to a complementary 50-mer to produce a dsDNA with an abasic site. The AP endonuclease activity was tested using 1 μM ssDNA or dsDNA and varying concentrations of the mvAPE enzyme in buffer P (10 mM Tris-Cl, 1 mM DTT, 0.1 mg/ml BSA, and 5% glycerol). Reactions were incubated at 30 °C for 30 min and stopped using 2X gel loading dye. The products formed were separated on a 20% urea-PAGE and visualized as described for the UDG assay.

Further, apurinic-apyrimidinic lyase activity of mvAPE was also checked using a 50-mer abasic site DNA labelled with [γ-^{32}P] ATP at its 5′-end (Supplemental Table S1, Sequence 10). The radiolabeled DNA was annealed to a complementary 50-mer to form dsDNA with an abasic site. Reaction conditions were identical to those used above for 3′-FAM labelled DNA cleavage. The products separated on urea-PAGE were visualized by exposure to a storage phosphor screen (GE Healthcare) and scanning on an Amersham Typhoon scanner (Cytiva).

3′-5′ exonuclease assays for mvAPE

For testing the 3′-5′ exonuclease activity, the reaction conditions were identical to the AP-lyase assay except that 3 different DNAs were used: (a) 15-mer ssDNA with a 5′-FAM label, (b) partially dsDNA with a recessed 3′-end (15-mer ssDNA with a 5′-FAM label annealed to a 27-mer ssDNA) and (c) a blunt ended dsDNA (15-mer ssDNA with a 5′-FAM label annealed to a 15-mer completely complementary ssDNA) (Supplemental Table S1). Reactions were incubated, stopped, separated on 20% urea-PAGE and visualized as described above for AP endonuclease assay.

In vitro reconstitution of BER pathway

Purified mimiviral proteins hypothesized to form the viral base excision repair pathway were reconstituted *in vitro*. A 5′-FAM labelled uracil containing oligonucleotide (500 nM) was incubated with purified wildtype enzymes, their heat inactivated form (for mvUDG) or the respective active site mutants (for mvAPE and mvPolX). The reactions were supplemented with 5 mM $MgCl_2$, 80 μM dNTPs, and assay buffer (20 mM Tris-HCl, 1 mM EDTA, and 10 mM NaCl). The concentrations of individual enzymes were 100 nM for mvUDG, 50 nM for mvAPE and 100 nM for mvPolX.

To check the repaired ligated product formed at the end of BER, one reaction also included 5 units of T4 DNA ligase (Thermo Fisher Scientific). All reactions were incubated at 30 °C for 30 mins and terminated by addition of 50 mM EDTA (final concentration) and 2X gel loading dye. Products were separated on a 20% urea-PAGE and visualized as described above for polymerase assays.

CRediT authorship contribution statement

Shailesh B. Lad: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Monica Upadhyay:** Methodology, Investigation, Formal analysis, Writing – review & editing. **Pracheta Thorat:** Methodology, Investigation, Formal analysis. **Divya Nair:** Methodology, Formal analysis, Writing – review & editing. **Gregory W Moseley:** . **Sanjeeva Srivastava:** Methodology,

Resources, Supervision, Writing – review & editing. **P.I. Pradeepkumar:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing. **Kiran Kondabagil:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors thank the Sophisticated Analytical Instruments Facility (SAIF) at IIT Bombay for providing access to the Typhoon 9500 phosphorimaging system and the Orbitrap Fusion Tribrid Mass Spectrometer. We also thank the Department of Chemistry, IIT Bombay for access to the Amersham Typhoon V imager.

Funding

Research in KK lab is supported by the Department of Biotechnology, DBT [BT/PR35928/BRB/10/1841/2019], the Department of Science and Technology, SERB [EMR/2016/005155], and The Board of Research in Nuclear Sciences, BRNS [58/14/11/2020-BRNS/37188] grants. Research in PIP lab is supported by grant from Science and Engineering Research Board, SERB (CRG/2021/001992). S.B.L acknowledges financial support provided by the IIT Bombay Senior Research Fellowship. M.U. acknowledges IITB-Monash Research Academy fellowship. P.T. thanks financial assistance from IRCC-IIT Bombay. MASSFIITB (Mass Spectrometry Facility, IIT Bombay) from the Department of Biotechnology (BT/PR13114/INF/22/206/2015) to SS is gratefully acknowledged for MS-based proteomics work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2023.168188.

Received 5 February 2023;
Accepted 20 June 2023;
Available online 26 June 2023

Keywords:
DNA repair;
BER pathway;
Mimivirus;
DNA polymerase;
glycosylase

References

1. McGlynn, P., Lloyd, R.G., (2002). Recombinational repair and restart of damaged replication forks. *Nature Rev. Mol. Cell Biol.* **3**, 859–870. https://doi.org/10.1038/nrm951.
2. Sale, J.E., Lehmann, A.R., Woodgate, R., (2012). Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nature Rev. Mol. Cell Biol.* **13**, 141–152. https://doi.org/10.1038/nrm3289.
3. Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L., Green, C.M., (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst).* **6**, 891–899. https://doi.org/10.1016/j.dnarep.2007.02.003.
4. Lindahl, T., (1999). Quality control by DNA repair. *Science (80-.)* **286**, 1897–1905. https://doi.org/10.1126/science.286.5446.1897.
5. Baute, J., Depicker, A., (2008). Base excision repair and its role in maintaining genome stability. *Crit. Rev. Biochem. Mol. Biol.* **43**, 239–276. https://doi.org/10.1080/10409230802309905.
6. Yonekura, S.I., Nakamura, N., Yonei, S., Zhang-Akiyama, Q.M., (2009). Generation, biological consequences and repair mechanisms of cytosine deamination in DNA. *J. Radiat. Res.* **50**, 19–26. https://doi.org/10.1269/jrr.08080.
7. Sung, J.S., Demple, B., (2006). Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *FEBS J.* **273**, 1620–1629. https://doi.org/10.1111/j.1742-4658.2006.05192.x.
8. Zharkov, D.O., (2008). Base excision DNA repair. *Cell. Mol. Life Sci.* **65**, 1544–1565. https://doi.org/10.1007/s00018-008-7543-2.
9. Yamtich, J., Sweasy, J.B., (1804). DNA polymerase Family X: Function, structure, and cellular roles. *Biochim. Biophys. Acta - Proteins Proteomics.* **2010**, 1136–1150. https://doi.org/10.1016/j.bbapap.2009.07.008.
10. Izumi, T., Mellon, I., (2016). Base Excision Repair and Nucleotide Excision Repair. Elsevier Inc.. Doi: 10.1016/B978-0-12-803309-8.00017-3.
11. Krokan, H.E., Bjørås, M., (2013). Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, https://doi.org/10.1101/cshperspect.a012583 a012583.
12. Pascucci, B., Maga, G., Hübscher, U., Bjoras, M., Seeberg, E., Hickson, I.D., Villani, G., Giordano, C., et al., (2002). Reconstitution of the base excision repair pathway for 7,8-dihydro-8-oxoguanine with purified human proteins. *Nucleic Acids Res.* **30**, 2124–2130. https://doi.org/10.1093/nar/30.10.2124.
13. Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., et al., (2004). The 1.2-megabase genome sequence of Mimivirus. *Science (80-.)* **306**, 1344–1350. https://doi.org/10.1126/science.1101485.
14. Legendre, M., Santini, S., Rico, A., Abergel, C., Claverie, J.-M., (2011). Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. *Virol. J.* **8**, 99. https://doi.org/10.1186/1743-422X-8-99.
15. Legendre, M., Audic, S., Poirot, O., Hingamp, P., Seltzer, V., Byrne, D., Lartigue, A., Lescot, M., et al., (2010). mRNA deep sequencing reveals 75 new genes and a complex

transcriptional landscape in Mimivirus. *Genome Res.* **20**, 664–674. https://doi.org/10.1101/gr.102582.109.
16. Bandaru, V., Zhao, X., Newton, M.R., Burrows, C.J., Wallace, S.S., (2007). Human endonuclease VIII-like (NEIL) proteins in the giant DNA Mimivirus. *DNA Repair (Amst).* **6**, 1629–1641. https://doi.org/10.1016/j.dnarep.2007.05.011.
17. Imamura, K., Averill, A., Wallace, S.S., Doublié, S., (2012). Structural Characterization of Viral Ortholog of Human DNA Glycosylase NEIL1 Bound to Thymine Glycol or 5-Hydroxyuracil-containing DNA. *J. Biol. Chem.* **287**, 4288–4298. https://doi.org/10.1074/jbc.M111.315309.
18. Prakash, A., Eckenroth, B.E., Averill, A.M., Imamura, K., Wallace, S.S., Doublié, S., (2013). Structural investigation of a viral ortholog of human NEIL2 / 3 DNA glycosylases. *DNA Repair (Amst).* **12**, 1062–1071. https://doi.org/10.1016/j.dnarep.2013.09.004.
19. Gupta, A., Lad, S.B., Ghodke, P.P., Pradeepkumar, P.I., Kondabagil, K., (2019). Mimivirus encodes a multifunctional primase with DNA/RNA polymerase, terminal transferase and translesion synthesis activities. *Nucleic Acids Res.* **47**, 6932–6945. https://doi.org/10.1093/nar/gkz236.
20. Redrejo-Rodríguez, M., Salas, M.L., (2014). Repair of base damage and genome maintenance in the nucleo-cytoplasmic large DNA viruses. *Virus Res.* **179**, 12–25. https://doi.org/10.1016/j.virusres.2013.10.017.
21. Renesto, P., Abergel, C., Decloquement, P., Moinier, D., Azza, S., Ogata, H., Fourquet, P., Gorvel, J.-P., et al., (2006). Mimivirus Giant Particles Incorporate a Large Fraction of Anonymous and Unique Gene Products. *J. Virol.* **80**, 11678–11685. https://doi.org/10.1128/jvi.00940-06.
22. Mutsafi, Y., Zauberman, N., Sabanay, I., Minsky, A., (2010). Vaccinia-like cytoplasmic replication of the giant Mimivirus. *Proc. Natl. Acad. Sci.* **107**, 5978–5982. https://doi.org/10.1073/pnas.0912737107.
23. Fridmann-Sirkis, Y., Milrot, E., Mutsafi, Y., Ben-Dor, S., Levin, Y., Savidor, A., Kartvelishvily, E., Minsky, A., (2016). Efficiency in Complexity: Composition and Dynamic Nature of Mimivirus Replication Factories. *J. Virol.* **90**, 10039–10047. https://doi.org/10.1128/jvi.01319-16.
24. Chatterjee, A., Ali, F., Bange, D., Kondabagil, K., (2016). Complete Genome Sequence of a New Megavirus Family Member Isolated from an Inland Water Lake for the First Time in India. *Genome Announc.* **4** https://doi.org/10.1128/genomeA.00402-16.
25. Chatterjee, A., Ali, F., Bange, D., Kondabagil, K., (2016). Isolation and complete genome sequencing of Mimivirus bombay, a Giant Virus in sewage of Mumbai. *India, Genomics Data.* **9**, 1–3. https://doi.org/10.1016/j.gdata.2016.05.013.
26. Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., Robert, C., Azza, S., et al., (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci.* **106**, 21848–21853. https://doi.org/10.1073/pnas.0911354106.
27. Yang, W., Lee, J.Y., Nowotny, M., (2006). Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity. *Mol. Cell.* **22**, 5–13. https://doi.org/10.1016/j.molcel.2006.03.013.
28. Beckman, R.A., Mildvan, A.S., Loeb, L.A., (1985). On the Fidelity of DNA Replication: Manganese Mutagenesis in Vitro†. *Biochemistry.* **24**, 5810–5817. https://doi.org/10.1021/bi00342a019.
29. Kaminski, A.M., Kunkel, T.A., Pedersen, L.C., Bebenek, K., (2022). Structural insights into the specificity of 8-oxo-7,8-dihydro-2′-deoxyguanosine bypass by family x dna polymerases. *Genes (Basel).* **13** https://doi.org/10.3390/genes13010015.
30. Batra, V.K., Shock, D.D., Beard, W.A., McKenna, C.E., Wilson, S.H., (2012). Binary complex crystal structure of DNA polymerase β reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 113–118. https://doi.org/10.1073/pnas.1112235108.
31. Kwon, E., Pathak, D., Chang, H.W., Kim, D.Y., (2017). Crystal structure of mimivirus uracil-DNA glycosylase. *PLoS One.* **12** https://doi.org/10.1371/journal.pone.0182382.
32. Vashishtha, A.K., Wang, J., Konigsberg, W.H., (2016). Different divalent cations alter the kinetics and fidelity of DNA polymerases. *J. Biol. Chem.* **291**, 20869–20875. https://doi.org/10.1074/jbc.R116.742494.
33. Brown, J.A., Suo, Z., (2011). Unlocking the sugar "steric gate" of DNA polymerases. *Biochemistry.* **50**, 1135–1142. https://doi.org/10.1021/bi101915z.
34. Cavanaugh, N.A., Beard, W.A., Wilson, S.H., (2010). DNA Polymerase β Ribonucleotide Discrimination. *J. Biol. Chem.* **285**, 24457–24465. https://doi.org/10.1074/jbc.M110.132407.
35. Crespan, E., Furrer, A., Rösinger, M., Bertoletti, F., Mentegari, E., Chiapparini, G., Imhof, R., Ziegler, N., et al., (2016). Impact of ribonucleotide incorporation by DNA polymerases β and λ on oxidative base excision repair. *Nature Commun.* **7**, 10805. https://doi.org/10.1038/ncomms10805.
36. Brown, J.A., Fiala, K.A., Fowler, J.D., Sherrer, S.M., Newmister, S.A., Duym, W.W., Suo, Z., (2010). A Novel Mechanism of Sugar Selection Utilized by a Human X-Family DNA Polymerase. *J. Mol. Biol.* **395**, 282–290. https://doi.org/10.1016/j.jmb.2009.11.003.
37. Bergoglio, V., Ferrari, E., Hübscher, U., Cazaux, C., Hoffmann, J.S., (2003). DNA polymerase β can incorporate ribonucleotides during DNA synthesis of undamaged and CPD-damaged DNA. *J. Mol. Biol.* **331**, 1017–1023. https://doi.org/10.1016/S0022-2836(03)00837-4.
38. Brown, J.A., Duym, W.W., Fowler, J.D., Suo, Z., (2007). Single-turnover Kinetic Analysis of the Mutagenic Potential of 8-Oxo-7,8-dihydro-2′-deoxyguanosine during Gap-filling Synthesis Catalyzed by Human DNA Polymerases λ and β. *J. Mol. Biol.* **367**, 1258–1269. https://doi.org/10.1016/j.jmb.2007.01.069.
39. Lamarche, B.J., Tsai, M.D., (2006). Contributions of an endonuclease IV homologue to DNA repair in the African swine fever virus. *Biochemistry.* **45**, 2790–2803. https://doi.org/10.1021/bi051772g.
40. Redrejo-Rodríguez, M., Ishchenko, A.A., Saparbaev, M.K., Salas, M.L., Salas, J., (2009). African swine fever virus AP endonuclease is a redox-sensitive enzyme that repairs alkylating and oxidative damage to DNA. *Virology.* **390**, 102–109. https://doi.org/10.1016/j.virol.2009.04.021.

41. Redrejo-Rodríguez, M., García-Escudero, R., Yáñez-Muñoz, R.J., Salas, M.L., Salas, J., (2006). African Swine Fever Virus Protein pE296R Is a DNA Repair Apurinic/Apyrimidinic Endonuclease Required for Virus Growth in Swine Macrophages. *J. Virol.* **80**, 4847–4857. https://doi.org/10.1128/jvi.80.10.4847-4857.2006.
42. Blanc-Mathieu, R., Ogata, H., (2016). DNA repair genes in the Megavirales pangenome. *Curr. Opin. Microbiol.* **31**, 94–100. https://doi.org/10.1016/j.mib.2016.03.011.
43. Redrejo-Rodriguez, M., Rodriguez, J.M., Suarez, C., Salas, J., Salas, M.L., (2013). Involvement of the Reparative DNA Polymerase Pol X of African Swine Fever Virus in the Maintenance of Viral Genome Stability In Vivo. *J. Virol.* **87**, 9780–9787. https://doi.org/10.1128/JVI.01173-13.
44. Fernández-Garciá, J.L., De Ory, A., Brussaard, C.P.D., De Vega, M., (2017). Phaeocystis globosa Virus DNA Polymerase X: A "swiss Army knife", Multifunctional DNA polymerase-lyase-ligase for Base Excision Repair. *Sci. Rep.* **7**, 1–13. https://doi.org/10.1038/s41598-017-07378-3.
45. Shukla, A., Chatterjee, A., Kondabagil, K., (2018). The number of genes encoding repeat domain-containing proteins positively correlates with genome size in amoebal giant viruses. *Virus Evol.* **4**, vex039. https://doi.org/10.1093/ve/vex039.
46. Chaudhari, H.V., Inamdar, M.M., Kondabagil, K., (2021). Scaling relation between genome length and particle size of viruses provides insights into viral life history. *IScience.* **24**, https://doi.org/10.1016/j.isci.2021.102452 102452.
47. Patil, S., Kondabagil, K., (2021). Coevolutionary and Phylogenetic Analysis of Mimiviral Replication Machinery Suggest the Cellular Origin of Mimiviruses. *Mol. Biol. Evol.* **38**, 2014–2029. https://doi.org/10.1093/molbev/msab003.
48. Colson, P., La Scola, B., Levasseur, A., Caetano-Anollés, G., Raoult, D., (2017). Mimivirus: Leading the way in the discovery of giant viruses of amoebae. *Nature Rev. Microbiol.* **15**, 243–254. https://doi.org/10.1038/nrmicro.2016.197.
49. Bogani, F., Chua, C.N., Boehmer, P.E., (2009). Reconstitution of uracil DNA glcosylase-initiated base excision repair in herpes simplex virus-1. *J. Biol. Chem.* **284**, 16784–16790. https://doi.org/10.1074/jbc.M109.010413.
50. Bogani, F., Corredeira, I., Fernandez, V., Sattler, U., Rutvisuttinunt, W., Defais, M., Boehmer, P.E., (2010). Association between the herpes simplex virus-1 DNA polymerase and uracil DNA glycosylase. *J. Biol. Chem.* **285**, 27664–27672. https://doi.org/10.1074/jbc.M110.131235.

Supplementary information

Biochemical reconstitution of the mimiviral base excision repair pathway

Shailesh B. Lad[1], Monica Upadhyay[1,3], Pracheta Thorat[1], Divya Nair[1], Gregory W. Moseley[3], Sanjeeva Srivastava[1], Pradeepkumar P. I.[2], and Kiran Kondabagil[1*]

Supplementary Table 1: Oligonucleotides used in the present study.

Sr. No.	**Sequence (5'-3')**	**Application**
1	FAM- CGTACTCGTAGGCATUAGTGTGACCAGCTGTTCAGGT AGGCACGGTAGGA	Uracil 50-mer with 5'-FAM label
2	CGTACTCGTAGGCAT/THF/AGTGTGACCAGCTGTTCAGGTAG GCACGGTAGGA-FAM	THF site 50-mer with 3-FAM label
3	CGTACTCGTAGGCAT/THF/AGTGTGACCAGCTGTTCAGGTAG GCACGGTAGGA	THF site 50-mer
4	CGTACTCGTAGGCAT	15-mer 5'-FAM labelled primer
5	TCCTACCGTGCCTACCTGAACAGCTGG	27-mer template for polymerase/exonuclease assays
6	ATGCCTACGAGTACG	15-mer FAM complementary
7	TCCTACCGTGCCTACCTGAACAGCTGGTCACACTGATGCCTAC GAGTACG	50-mer template for strand displacement/gap-filling assays
8	TGACCAGCTGTTCAGGTAGGCACGGTAGGA	Blocking primer for gap-filling/strand displacement assays
9	AGTGTGACCAGCTGTTCAGGTAGGCACGGTAGGA	Blocking primer gap-filling/strand displacement assays
10	TCCTACCGTGCCTACCTGAACAGCTGGTCACACT**8-oxo-dG**ATGCCTACGAGTACG	TLS 8-oxo-dG modified 50-mer

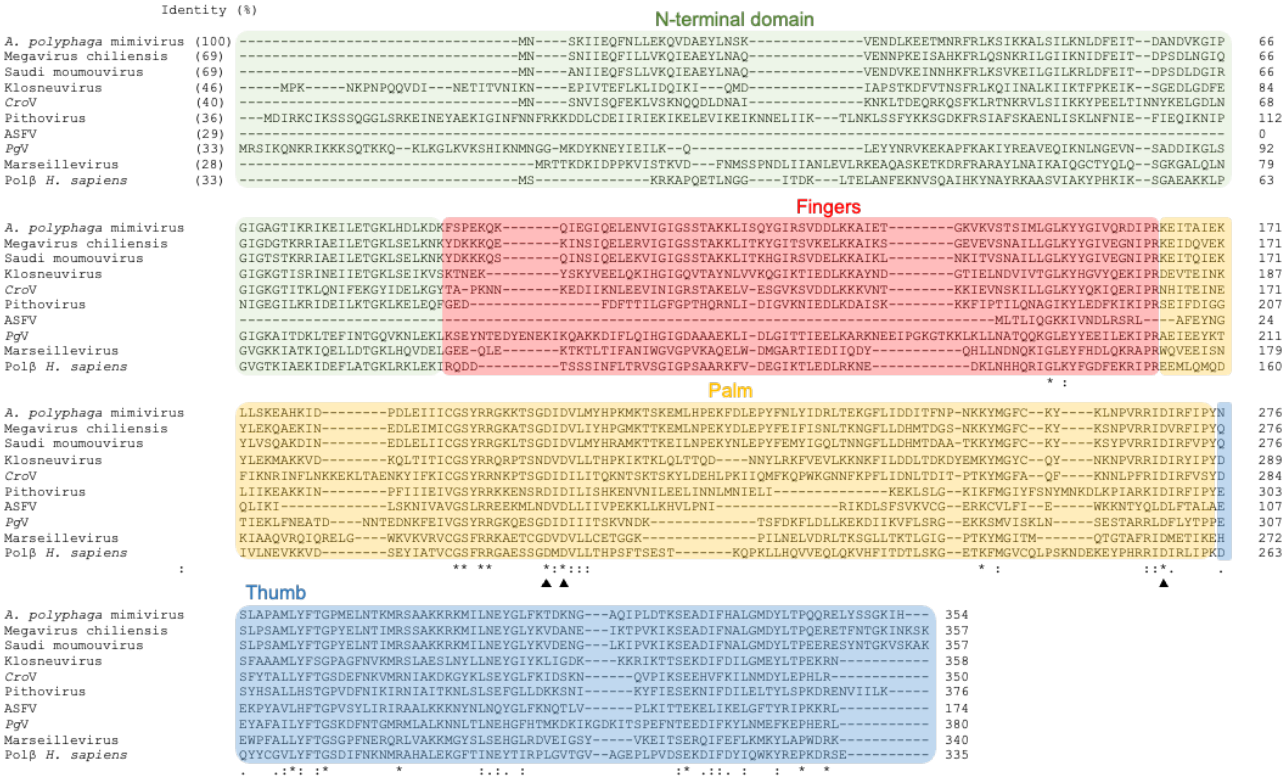


Figure S1 Multiple sequence alignment of mvPolX with known polymerases from other organisms.

The highlighted regions show the different domains namely the 8 kDa N-terminal domain (green), fingers (red), palm domain (yellow), and thumb (blue). The active site residues predicted to be essential for the polymerase activity are indicated by solid triangles.
Accession IDs:
YP_003986821.1, YP_004894637.1, AQN68410.1, ARF11843.1, YP_003970091.1, QBK90626.1, QBQ34698.1, YP_008052712.1, QBK88296.1, AAB60688.1

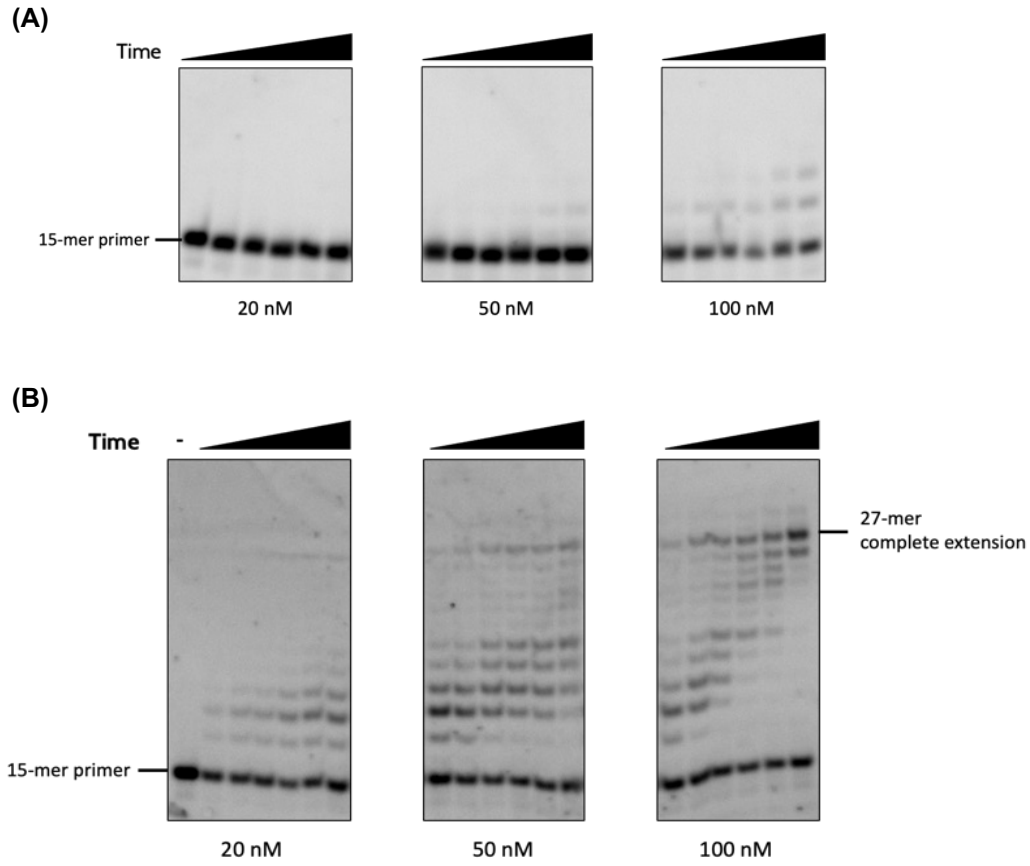


Figure S2 DNA polymerase activity of mvPolX is dependent on divalent cations.

DNA polymerase activity of mvPolX on a primer-template complex in the absence **(A)** and presence **(B)** of 5 mM $MgCl_2$ salt. Reactions were incubated with all four dNTPs and the indicated amounts of mvPolX at 30 °C. Aliquots were taken at 2, 5, 10, 30, 60, and 90 min after initiating the reaction.

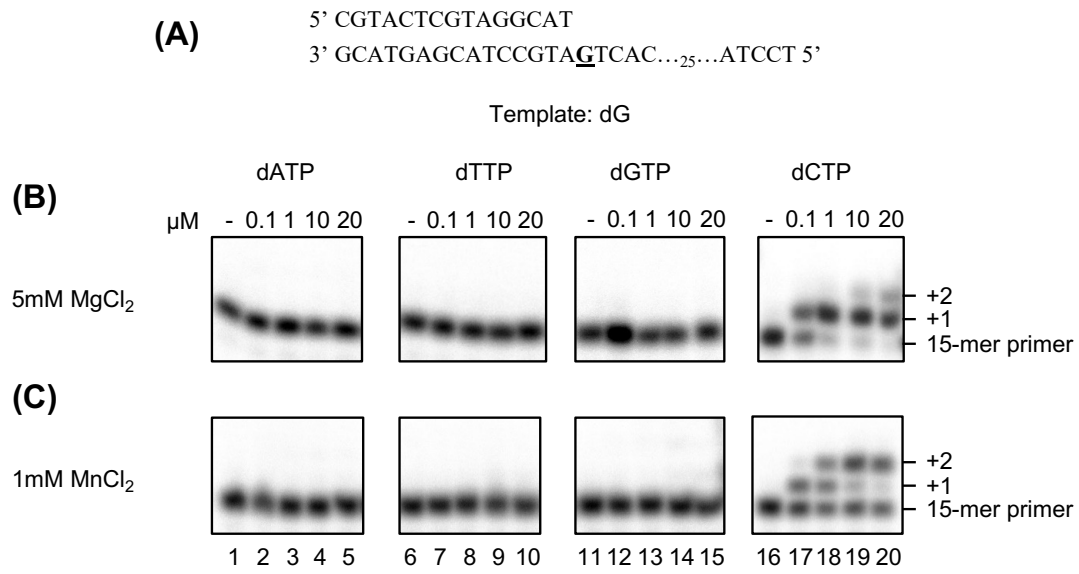


Figure S3: Effect of divalent cations on the activity of mvPolX.

(A) A schematic representation of the substrate used for assessing the fidelity of mvPolX. The template nucleotide is indicated in bold and underlined. (B) and (C) Single nucleotide incorporation against a template dG using 5 mM $MgCl_2$ and 1 mM $MnCl_2$ as cofactors, respectively. All standing start assays were performed using 100 nM mvPolX and incubated at 30 °C for 5 min to maintain near steady-state conditions.

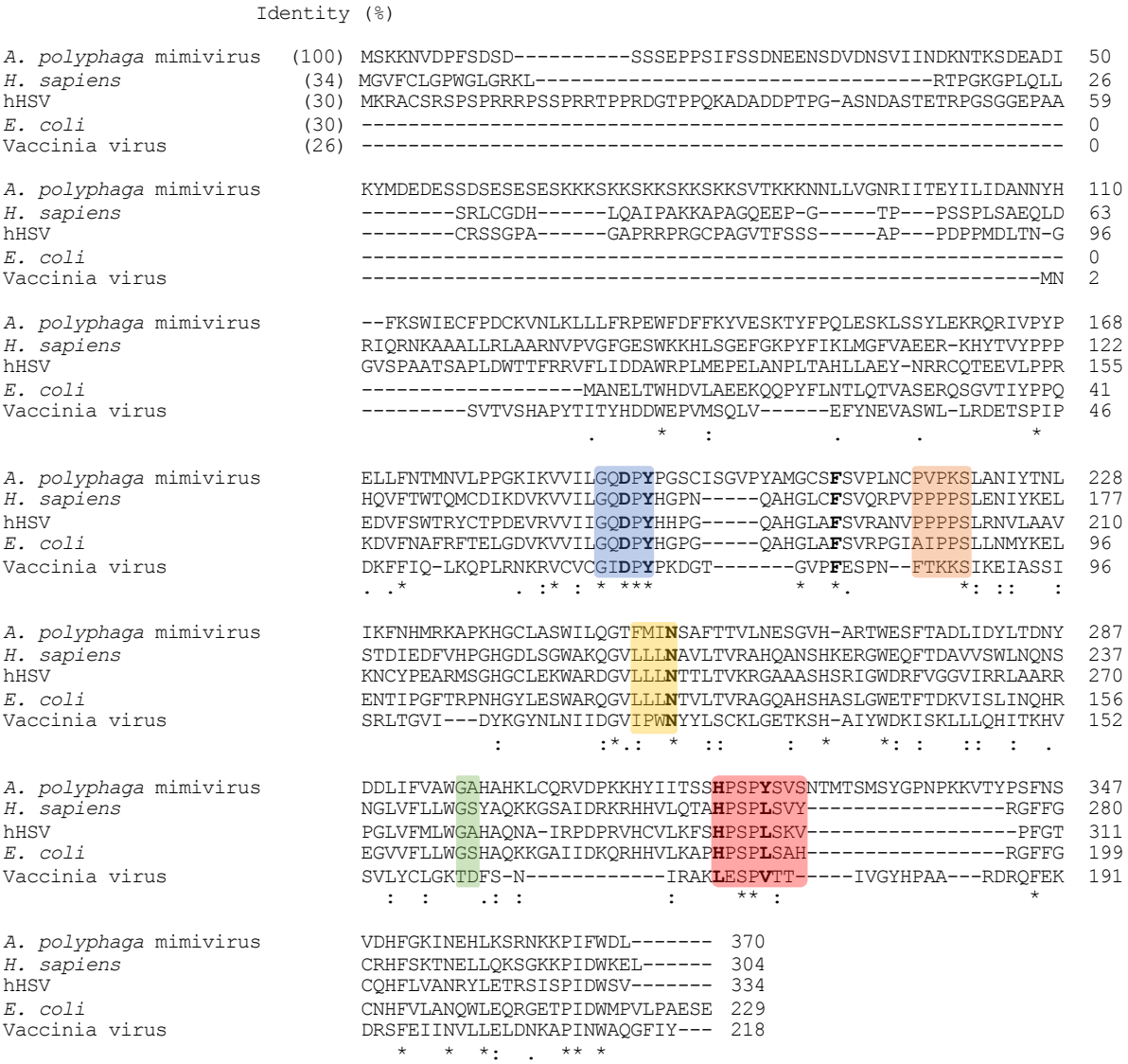


Figure S4 Multiple sequence alignment of mvUDG with known proteins.

Alignment of mvUDG with previously studied proteins from bacteria, viruses and eukaryotes. The catalytically important motifs are boxed: water-activating motif (blue), proline rich motif (orange), uracil recognition motif (yellow), glycine-serine motif (green), and DNA minor groove intercalation motif (red).
Accession IDs: YP_003986745.1, NP_417075.1, NP_003353.1, ABD52579.1, SBS69165.1.

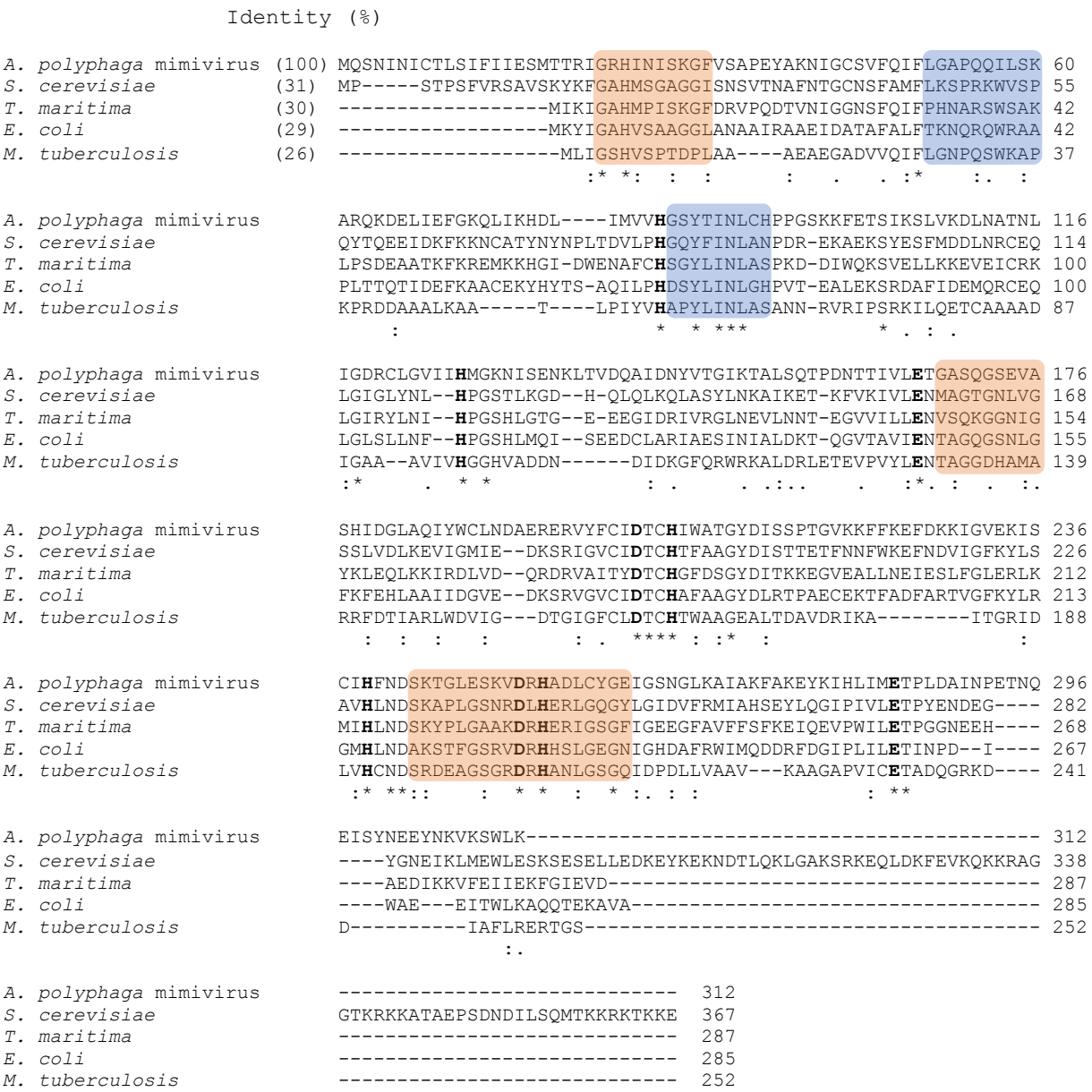


Figure S5 Multiple sequence alignment of mvAPE with known proteins.

The mvAPE protein was aligned with other known AP-endonuclease IV proteins using Clustal Omega. Identity of mvAPE with other proteins is shown in parenthesis. Asterisk signs indicate conserved residues across species. The metal binding active site amino acids are highlighted in bold and the conserved motifs including those involved in phosphate and minor groove binding are boxed in orange and blue, respectively.
Accession IDs: YP_003986798.1, AAA60529.1, NP_215184.1, AKE30034.1, AAA34429.1.