Zebrafish Pol Theta and Human Pol Theta: Orthologues with Homologous Function.

Cory Thomas[1], Sidney Green[1], Lily Kimball[2], Isaiah Schmidtke[2], Makayla Griffin[2], Lauren Rothwell[2], Ivy Par[1], Sophia Schobel[1], Yayleene Palacio[1], Jamie B Towle-Weicksel[1], Steven E Weicksel[2]

[1]Department of Physical Sciences, Rhode Island College, Providence, RI

[2]Department of Biology and Biological Sciences, Bryant University, Smithfield RI

ABSTRACT

DNA polymerase theta (Pol θ) is a conserved an A-family polymerase that plays an essential role in repairing double strand breaks, through micro-homology end joining, and bypassing DNA lesions, through translesion synthesis, to protect genome integrity. Despite its essential role in DNA repair, Pol θ is inherently error prone. Recently, Pol θ has been identified in various cancer types, suggesting Pol θ's promiscuous nature aids in cancer progression.

Here we present a study comparing the structure and function of the polymerase domain of zebrafish and human Pol θ to determine the potential of zebrafish as a model for pol θ function. We show that these two proteins share a large amount of sequence and structural homology, but differ in key loop areas thought to drive defining functions of the enzyme. Despite these differences zebrafish Pol θ still displays characteristics identify in human Pol θ, including DNA template extension, microhomology-mediated end

23 joining, and translesion synthesis. Additionally, we found several important residues within
24 loops of the polymerase domain that support function unique to Pol θ. These results will
25 support future studies with zebrafish as a model to investigate Pol θ function.
26

27 **INTRODUCTION**

28 Melanoma is the rarest but deadliest form of skin cancer, with an estimate in 2020
29 of 325,000 new cases worldwide, increase to nearly 510,000 new cases by 2040[1]. Unlike
30 the expansion of melanocytes horizontally along the epidermis as benign nevi, (commonly
31 referred to a moles) melanoma is characterized by expansion vertically, down, to and
32 through the basement membrane where cells then begin to spread out from the original
33 tumor location. Melanoma cells are highly mobile in part due to the migratory nature of
34 melanocytes[2], making migration at key tropism of the disease. Surgical treatments prior
35 to migration of melanoma to the lymph nodes have a high degree of success, with low
36 rates of tumor reoccurrence and favorable patient outcomes. Patient outcomes, however,
37 become less favorable after migration, as metastatic tumors often are no longer sensitive
38 to radiation or chemotherapy-based approaches. Recent advances in new immune-
39 boosting therapies have shown promise in treating metastatic melanoma, but early
40 detection and tumor resection treatments are still the most effective[3,4]. This indicates that
41 there is still much to know about melanoma carcinogenesis. Particularly, what are the
42 genetic markers for disease and what potential targets exist for novel treatments
43 metastatic melanoma.

44 One emerging factor identified in a recent study of patient derived melanoma
45 samples is DNA polymerase-theta (Pol θ or POLQ)[5]. An A-family DNA repair enzyme,

46 Pol θ is essential for cell function and organismal development[6]. Inherently error prone
47 Pol θ[7,8] plays a predominant role repairing double strand breaks (DSB) in the DNA strand
48 through microhomology-mediated end joining (MMEJ, also known as theta-mediated end
49 joining, (TMEJ)), and translesion nucleotide bypass[9–12]. Unlike homologous
50 recombination (HR), the favored DSB repair pathway, TMEJ is highly error prone and is
51 proposed to be activated when HR is overwhelmed (when the genome occurs many
52 double strand breaks) and/or inactive (such as in cancer states). The activity of Pol θ in
53 translesion nucleotide bypass plays a critical role in replication, that while perpetuating
54 genomic mutations, allows replicative DNA polymerases to continue replication while also
55 avoiding more DNA DSBs and potential mutagenesis through replication fork collapse[13].
56 Pol θ is also a key factor in repairing DBSs during early embryonic development[6].
57 Together this indicates that Pol θ function is intrinsically mutagenic yet required for cell
58 function.

59 This duality, mutagenic enzymatic behavior while also supporting cell survival[6,14,15],
60 along with aberrant Pol θ activity in cancer cells[16–18] that has led to many hypothesizing
61 that Pol θ activity drives carcinogenesis. In the case of melanoma, it is hypothesized
62 carcinogenesis is driven in part by improper DNA repair of UV damage generated by sun
63 exposure[8,19]. UV radiation has the potential to create thymine dimers on the DNA strand
64 which either need to be resolved by base excision repair, a mechanism with low
65 mutagenic potential, or by translesion nucleotide bypass, a mutagenic mechanism which
66 Pol θ can perform. As mentioned before, bypassing or repairing these damaged sites in
67 the genome are important for genomic stability as they will result in replication fork
68 collapse and DSBs during replication. These DSBs could potentially lead to greater

69 mutagenic events if instead of HR, DSBs are repaired though the MMEJ/TMEJ pathways
70 by Pol θ. Interestingly, variant forms of Pol θ identified in patient melanoma samples,
71 display a range of enzymatic properties different from wild-type (WT)[20]. These data
72 suggest enhanced mutagenic capabilities for these versions of Pol θ. However, few
73 models for assessing the function of Pol θ as well as the outcome of Pol θ function in the
74 context of an organism exist.

75 Zebrafish have long been employed to model melanoma for organismal study[21].
76 With nearly 70% homology in their genes and 85% in human disease-related genes[22],
77 zebrafish have a high degree of similarity to humans. Like humans, zebrafish have
78 melanin producing melanocytes which are displayed most prominently in their horizonal
79 stripes. Zebrafish melanocytes produce melanin through pathways homologous to those
80 in mammals and develop from neural crest cells in the embryo[2]. Xenografts of human
81 melanoma tumors into zebrafish have also shown that these tumors have similar
82 metastatic tropism in zebrafish to those observed in human malignancies[23,24]. These
83 studies suggest many of the cellular cues of melanoma spread and disease are similar
84 between human and zebrafish.

85 Here we present the first comparative analysis of protein structure and function of
86 purified zebrafish POLQ (zPol θ or zPOLQ) and human POLQ (hPol θ or hPOLQ)
87 polymerase domains. Protein alignment indicates that many of the residues present in
88 the polymerase domain between the two proteins are conserved resulting in similar folded
89 structures. However, within loop regions (unresolved in the human crystal structure),
90 specific to PolQ relative to other A-type proteins, there is little conservation. Despite this
91 lack of conservation, we observe similar zPOLQ behavior compared to hPOLQ. zPolQ

92 can extend DNA templates, perform TMEJ, and bypass DNA lesions, hallmarks of PolQ
93 function in the cell.
94
95 **MATERIALS AND METHODS**
96 *Materials*
97 All materials were purchased from Sigma-Aldrich (St. Louis, MO), Bio-rad
98 Laboratories (Hercules, CA), AmericanBio (Canton, MA), and Research Products (Mount
99 Prospect, IL). DNA oligonucleotides were purchased from Integrated DNA Technologies
100 (Newark, NJ) and deoxynucleotides from New England Biolabs (Ipswich, MA). All DNA
101 oligos were purified via HPLC with standard desalting from the manufacturer.
102
103 *Zebrafish Pol θ cloning*
104 Total RNA from 4-hour post fertilization embryos was extracted using TRIzol
105 (Invitrogen) following manufactures instructions. A library of cDNAs was generated from
106 the pool of polyA mRNAs using ProtoScript II Reverse Transcriptase (New England
107 Biolabs, NEB) following manufactures instructions, primed by oligo(dT). The polymerase
108 domain of zebrafish Pol θ was then amplified for cloning into the POLQM1 vector[12] a
109 pSUMO3 based expression vector. This was a two-step cloning process as the POLQM1
110 vector did not have multiple cloning sites.
111 First, the polymerase domain of zebrafish Pol θ (residues 1801-2579) was
112 amplified from the cDNA library with primers contain a 5' KpnI site and a 3' BamHI site:
113 Pol θ RVS BamHI- TATACTGGATCCTTATATGTCCAGGTCTTGAAGGTTACC
114 Pol θ FWD KpnI- ATTAGGTACCTCAACATCAGTGTTAGGCGCAC

115 Second, the 6xHIS and SUMO sequences (HIS-SUMO) of POLQM1 were
116 amplified off of the plasmid using primers containing a 5' XbaI site and a 3' KpnI site:
117 HIS-SUMO RVS KpnI- ATTAGGTACCTCCCGTCTGCTGC
118 HIS-SUMO FWD XbaI- TTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAG
119 PCRs reactions used Phusion High-Fidelity DNA Polymerase (NEB) following
120 manufacturer's instructions and were run for 30 cycles. PCR Products were gel isolated
121 from a 1% TAE agarose gel using Freeze 'N Squeeze DNA gel extraction columns (Bio-
122 Rad), following manufacturer's instructions.
123 Next, the zPol θ, HIS-SUMO, and POLQM1 DNAs were digested with appropriate
124 enzymes (NEB) overnight at 37ºC:
125 zPol θ PCR - BamHI and KpnI
126 HIS-SUMO PCR - KpnI and XbaI
127 POLQM1 - XbaI and BamHI
128 Digested samples were gel separated on a 1% TAE agarose gel and fragments were
129 isolated using Freeze 'N Squeeze DNA gel extraction columns. zPol θ and HIS-SUMO
130 digested fragments were then ligated using T4 ligase (NEB) incubating at 16ºC overnight
131 and gel isolated from a 1% TAE agarose gel using Freeze 'N Squeeze DNA gel extraction
132 columns. zPol θ-HIS-SUMO fragment was ligated into linearized POLQM1 vector using
133 T4 ligase and incubating at 16ºC overnight. NEB 5-alpha competent *E.coli* (NEB) were
134 transformed by ligated productions using manufacturer's instructions. Bacteria were
135 selected for through ampicillin resistance.
136
137 *zPol θ polymerase domain modeling and alignments*

138 Amino acid sequence alignments were completed using EMBL-EBI Clustal Omega
139 MSA[25] on default settings. Structural rendering of the zPol θ polymerase domain was
140 completed using ColabFold v1.5.5: AlphaFold2 using MMseqs2[26–30] on default settings.
141 The resulting zPol θ structure was compared to the solved hPol θ polymerase domain
142 structure (AX0Q)[31] using the pairwise alignment tool in FATCAT[32] on default setting, and
143 visualized using RCSB PDB visualization tools[33].
144
145 *Expression and Purification of hPol θ and zPol θ*
146 Recombinant pSUMO3 plasmids containing the truncated polymerase domain of
147 hPol θ and zPol θ gene were expressed in *E.coli* and purified as previously described in
148 Thomas et al[34]. Briefly, Rosetta2(DE3) cells containing the zPol θ plasmid were inoculated
149 into autoinduction Terrific Broth and grown at 20°C for 60 hours. Cells were harvested
150 through centrifugation and lysed in Lysis buffer. After 6 rounds of sonication, cell fractions
151 were further separated via centrifugation. The fraction containing soluble protein were
152 applied to a 5mL His-Trap FF crude Nickel Column (Cytiva) by FPLC using a high
153 imidazole gradient. Fractions containing Pol θ were separated again on a HiTrap Heparin
154 HP (Cytiva) column for further purification. Eluate containing Pol θ was incubated
155 overnight at 4°C with SUMO2 Protease (Fisher Scientific) to remove the 6xHIS-SUMO
156 tag. Untagged Pol θ was separated from the 6X-HIS-SUMO on a Hi-Trap Chelating HP
157 column, reserving the flowthrough that contained only Pol θ. A final HiTrap Heparin
158 column removed any remaining non-specific binding proteins and exchanged the
159 imidazole buffer for a high NaCl buffer. Protein purification was verified on 10% denaturing
160 SDS visualized on an Odyssey CL-x IR scanner (LiCOR). Purified protein is highly

161 unstable and was flash-frozen in liquid nitrogen and stored at -80°C for 3 months
162 maximum.
163
164 *DNA Substrate Generation*
165 Double stranded DNA substrates (dsDNA) were generated using complementary
166 oligodeoxynucleotides from IDT. Templates representing cyclobutane pyrimidine
167 Thymine-Thymine dimer (CPD) damaged and undamaged were synthesized in the
168 Delaney laboratory Sarah Delaney (Brown University). The 5'6-FAM primers were
169 annealed to complementary DNA templates with sequence context as previously
170 described [35,36] and are described below:
171 25/40 undamaged DNA substrate
172 5'-/FAM/ TTT GCCT TGA CCA TGT AAC AGA GAG
173 CGGA ACT GGT ACA TTG TCT CTC GCA CTC ACT CTC TTC TCT
174 24/33 CPD damaged DNA substrate
175 5'-/FAM/-TTT CTC CGG TAC TCC AGT GTA GGC
176 GAG GCC ATG AGG TCA CAT CCG **TT**A GCT TGA GAA
177 24/33 undamaged DNA substrate
178 5'-/FAM/-TTT CTC CGG TAC TCC AGT GTA GGC
179 GAG GCC ATG AGG TCA CAT CCG **TT**A GCT TGA GAA
180 Confirmation of annealed substrates was determined 12% Native PAGE and samples
181 scanned on an RB Amersham Typhoon Fluorescent Imager (Cytiva) with a FAM filter.
182 Single oligodeoxynucleotides were purchased from IDT for MMEJ with internal consensus
183 sequence as previously described[37].
184 5'-/FAM/GGTTAGCCCGGG

185

186 *Circular Dichroism and Melting Temperature*

187 Secondary protein characteristics of hPol θ to zPol θ were determined on a J-815-
188 CD Spectropolarimeter (Jasco, Brown University) with a 0.2 cm quartz cuvette at room
189 temperature (20°C). For each sample, 3 μM of protein in 10 mM K_2HPO_4 buffer were
190 scanned in triplicate from 190-280 nm. The thermal denaturation profile was determined
191 by using the same instrument by heating the same sample from 20-90°C with a 5°C/min
192 temperature rate increase at 222 nm using the same. Data were analyzed on Prism 10
193 GraphPad and the melting temperature (Tm) was estimated using the halfway point of
194 the denaturing curve

195

196 *Electrophoretic Mobility Shift Assay (EMSA)*

197 The DNA binding affinity constant $K_{D(DNA)}$ was determined as previously described[34]. zPol
198 θ was titrated from 0-1000nM against 10nM 25/40 dsDNA substrate in binding buffer and
199 incubated for 1 hour at room temperature. Samples were separated on a 6% Native
200 PAGE and scanned on an RB Typhoon scanner (Cytiva) with the FAM fluorescence filter.
201 Separated bound and unbound products were quantified using ImageQuant. $K_{D(DNA)}$ was
202 determined by equation 1.

203 $$Y = \left[\frac{(mx)}{(x+K_D)}\right] + b \qquad \text{Equation 1}$$

204 Four replicates and two protein preparations were used to generate this data.

205

206 *Rapid Chemical Quench Assay*

207 Biphasic burst kinetics were measured as previously described[34]. Briefly,100 nM
208 Pol θ was pre-mixed with 300 nM 25/40 dsDNA substrate and rapidly mixed with 100 μM
209 of dCTP (correct nucleotide) with 10 mM $MgCl_2$ using an RQF-3 Rapid Chemical Quench
210 instrument (KinTek Corporation) at 37°C between 0.004-0.6 seconds. Reactions were
211 quenched by addition of 0.5M EDTA. Products were separated on a 15% Urea-denaturing
212 polyacrylamide gel and scanned using an Amersham Typhoon RB Fluorescent imager
213 (Cytiva). Extended product (n+1) was quantified using ImageQuant software and then
214 plotted to a full biphasic pre-steady state burst equation via non-linear regression using
215 Prism 9 GraphPad software (equation 2). A minimum of three replicates were included for
216 each assay on two independent protein preparations.

217 $$[Product] = [E]_{app}[\frac{{k_{obs}}^2}{(k_{obs}+k_{ss})^2} \times (1 - e^{-(k_{obs}+k_{ss})t} + \frac{k_{obs}k_{ss}}{(k_{obs}+k_{ss})}t] \quad \text{Equation 2}$$

218

219 *Primer Extension Assays*

220 Qualitative primer extension assays were performed as previously described[34].
221 Varying conditions were used to explore the primer extension capabilities between the
222 hPol θ and zPol θ. Under single-turnover conditions, excess Pol θ (200 nM) was pre-
223 incubated with 50 nM DNA and incubated for 5 minutes at 37˚C. Nucleotide (125 μM) of
224 either none, all, individual nucleotides were preincubated with 20 mM *x* inorganic salt (x
225 = $MgCl_2$, $CaCl_2$, $MnCl_2$). Under Michaelis-Menton conditions, DNA (200nM) was in excess
226 to Pol θ (50 nM). All reactions were carried out in buffer containing 20 mM Tris HCl, pH
227 8.0, 25 mM KCl, 4% glycerol, 1 mM βME, and 80 μg/mL BSA. The reaction was initiated
228 by combining Pol-θ/DNA with dNTP/salt. Reactions were incubated at 37˚C for 5 minutes
229 before being stopped by 80% Formamide/EDTA quench. Products were separated out on

230 a 15% urea-denaturing polyacrylamide gel, and scanned on a RB Amersham Typhoon
231 fluorescent imager (Cytiva).
232
233 *MMEJ Assay*
234 Microhomology mediated end-joining assay for both hPol θ and zPol θ were carried
235 out as previously described[37] on a 12-mer FAM labeled oligodeoxynucleotide. Pol θ (20
236 nM) was preincubated with 30 nM 5'-FAM ssDNA in reaction buffer (25mM Tris-HCl
237 pH8.8, 1mM βME, 0.01% NP-40, 0.1 mg/mL BSA, 10% glycerol, 10 mM $MgCl_2$, 30mM
238 NaCl) for 5 minutes at 37°C. Nucleotides (20 μM) were added and incubated at 37°C for
239 an additional 45 minutes. Reactions were terminated by addition of non-denaturing stop
240 buffer (100 mM Tris-HCl pH 7.5, 10 mg/mL proteinase K, 80 mM EDTA, and 0.5% SDS)
241 for an additional 15 minutes. DNA products were separated on a 12% native
242 polyacrylamide gel and scanned by an RB fluorescent Amersham Typhoon (Cytiva) with
243 a FAM filter.
244
245 **RESULTS**
246
247 *Zebrafish and human polymerase domains display high degree of structural similarity*
248 To determine the degree of similarity between the zPol θ and hPol θ PD primary
249 amino acid sequences were aligned (Table 1, Supplemental figure 1). The alignment of
250 the full-length Pol θ protein indicates, zebrafish and human Pol θ share 46% identity. This
251 degree of similarity increases when comparing the predicted polymerase domain (63%),

252 as well as subdomains containing catalytic activity, fingers (75.3%), thumb (74.6%), and
253 palm (66.7%). These data suggest a structurally similar molecule.

254 To assess the extent of structural similarity we generated a predicted structure for
255 zPol θ PD using ColabFold[26] to compare to the solved crystal structure of hPol θ PD[31].
256 Upon visual inspection the predicted zPol θ PD displays classical DNA polymerase PD
257 structures (Figure 1). The three major subunits, the fingers, thumb, and palm are visible,
258 and when modeled in, a DNA molecule can fit in the presumed catalytic domain. The
259 model also indicates the presence of unstructured loop domains, that have functional
260 importance[12], that were not resolved in the hPol θ PD structure. An over lay of the hPol
261 θ PD and zPol θ PD show that the structures have a high degree of similarity. As predicted
262 by the amino-acid alignment, these data indicate that much of the structure of the hPol θ
263 PD is conserved in zPol θ.

264 Table 1. Sequence alignment analysis

	Polymerase theta comparisons		
	zebrafish	% identity	human
Polymerase Domain	thumb (122 aa)	74.6%	thumb (125 aa)
	fingers (142 aa)	75.3%	fingers (142 aa)
	palm (163 aa)	66.7%	palm (177 aa)
	total (744 aa)	63%	total (799 aa)
Full length	2576 aa	46%	2590 aa

265

266

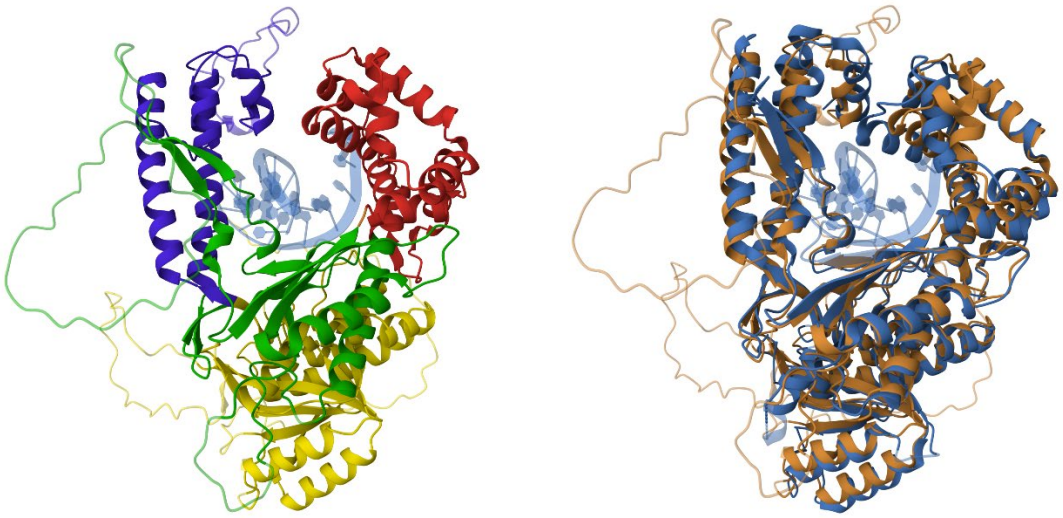


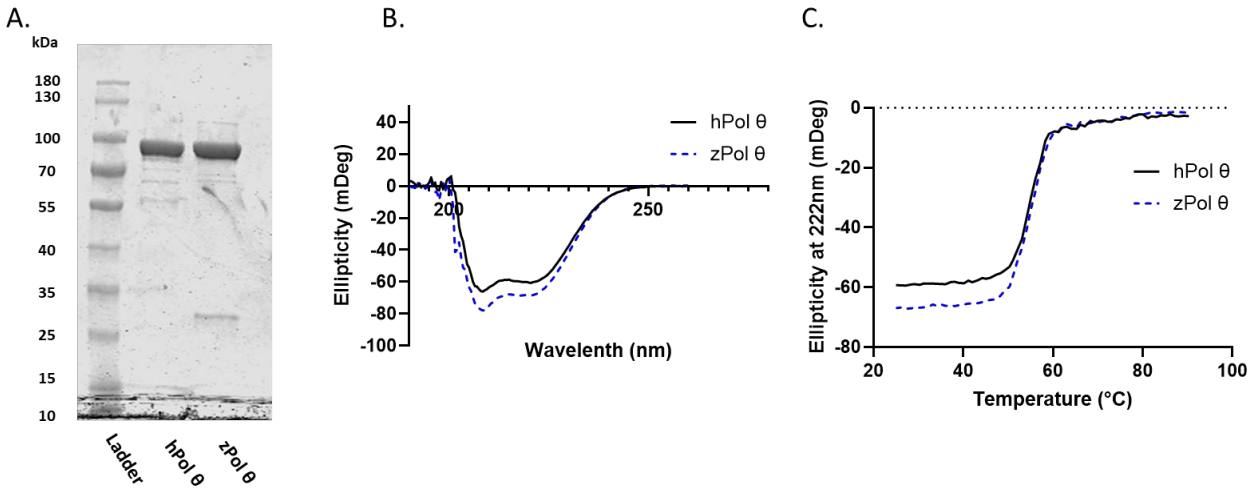
Figure 1. Structural modeling of zPol θ. (A) AlphaFold rendering of zPol θ with subdomains colored, thumb – blue, fingers – red, palm – green, and exo-nuclease in yellow. DNA is colored in light blue. (B) FATCAT overlay of AlphaFold rending of zPol θ (bronze) and hPol θ (navy) crystal structure[31] without loop inserts, DNA is colored in light blue.

hPol θ and zPol θ are structural similar

The plasmid containing the c-terminal recombinant zPol θ was expressed and purified in the same way as hPol θ[34] and as summarized in the Materials and Methods. Similar to hPol θ, one protein preparation yields approximately 5-10 μM and we observed similar expression and purification levels as seen with hPol θ (Figure 2A)

To confirm similarity in secondary structure between hPol θ and zPol θ, circular dichroism spectroscopy (CD) was performed at 20°C. The same sample was heated from 20-90°C in order to determine the thermal denaturation profile. Both spectra were overlayed and indicated minimal variance suggesting that both hPol θ and zPol θ have

281 similar secondary characteristics and thermal stability with a Tm of about 55°C.



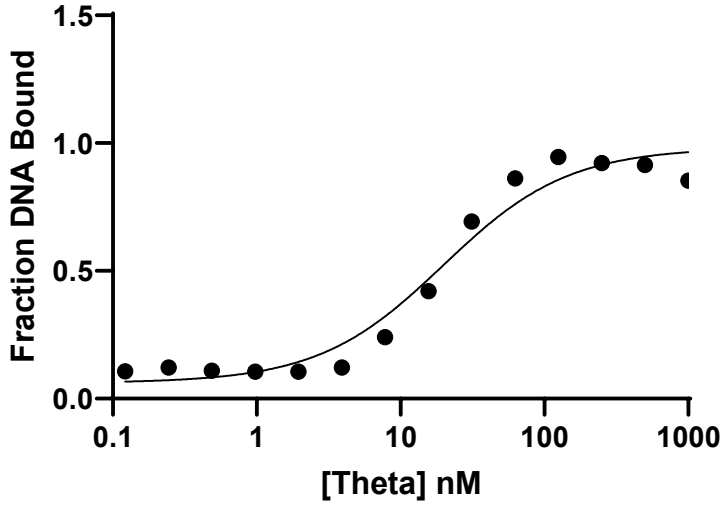
282

283 **Figure 2. hPol θ and zPol θ display similar secondary characteristics and thermal stability.** (A) Expression and purification of
284 zPol θ was the same as hPol θ as described in the Materials and Methods. For each sample, approximately 56-60 pmol of cleaved,
285 purified protein were loaded on a 10% SDS PAGE and Coomassie stained. Both hPol θ and zPol θ migrate to approximately 90 kDa
286 as expected. (B) Circular dichroism spectra of 3 μM hPol θ (solid line) and zPol θ (dashed) proteins in 10 mM Potassium Phosphate
287 buffer. Samples were scanned from 190 to 280 nm. (C)The same samples were heated from 20-90°C and ellipticity measured at 222
288 nm.

289

290 *zPol θ binds to dsDNA substrate*

291 DNA binding by a DNA polymerase is one of the first steps in its catalytic
292 mechanism. To determine the DNA binding capabilities of zPol θ, we titrated zPol θ from
293 0-1000 nM protein against 10 nM 25/40 dsDNA. Complexed DNA/protein products were
294 separated on a denaturing gel to determine a dissociation constant ($K_{D(DNA)}$) for DNA
295 binding. Similar to hPol θ, zPol θ has a low $K_{D(DNA)}$ value of approximately 19.8 ± 3.1 nM.



296

297 **Figure 3. zPol θ binds tightly to ds DNA.** zPol θ was titrated from 0-1000 nM against 10 nM 25/40 dsDNA. Bound and unbound
298 products were separated on a 6% non-denaturing gel and quantified using ImageQuant. $K_{D(DNA)}$ was mathematically calculated using
299 Equation 1 and is the midpoint between bound and unbound fractions.

300

301 *zPol θ can extend dsDNA similar to hPol θ*

302 The second step in the DNA polymerase catalytic pathway is nucleotide binding
303 and formation of the phosphodiester bond. To explore this fundamental step of DNA
304 Polymerase activity, we assayed zPol θ's ability to extend 25/40 dsDNA under varying
305 conditions. Under standard steady-state conditions, 200 nM of zPol θ or hPol θ was pre-
306 incubated with 50 nM 25/40 dsDNA. The reaction was initiated by the addition of 125 nM
307 dNTP as described in Figure 4 along with 20 mM $MgCl_2$, the preferred metal for DNA
308 polymerase [38]. We observed under these conditions both hPol θ and zPol θ were able to
309 extend the full 18-mer template with all nucleotides present (Figure 4). Both enzymes
310 were able to incorporate single nucleotides, correct and incorrect as well. Notably, zPol θ
311 was able to incorporate incorrect dGTP to full extension (n+1) compared to only n+6 with
312 hPol θ. DNA polymerases can utilize other metals including Mn^{2+}, and we observed an

313 increase in mutagenesis through misincorporation for both hPol θ and zPol θ. When
314 provided with all dNTP, zPol θ can extend past the end of the template (n+18). Overall,
315 zPol θ experiences more extension products especially with incorrect nucleotides dATP,
316 dGTP, and dTTP compared with hPol θ under similar conditions. Steady-state conditions
317 highlight overall DNA polymerase activity, but because the dsDNA substrate is in excess,
318 activity highlights multiple turnovers[39]. Although DNA pol θ has been shown to have
319 robust de novo activity with manganese[40], we wanted to be sure this over extension
320 observed with zPol θ and Mn^{2+} was the result of extension and not an artifact. We
321 changed the ratio of protein to DNA to reflect single-turnover conditions; excess protein
322 over dsDNA substrate. Here we are able to observe polymerization events for
323 theoretically every available DNA substrate. Similar to steady-state conditions, we
324 observe an even more robust de novo extension with not only all nucleotides, but also
325 with dATP, suggesting that zPol θ misincorporation with dATP is preferred (Supplement
326 2).

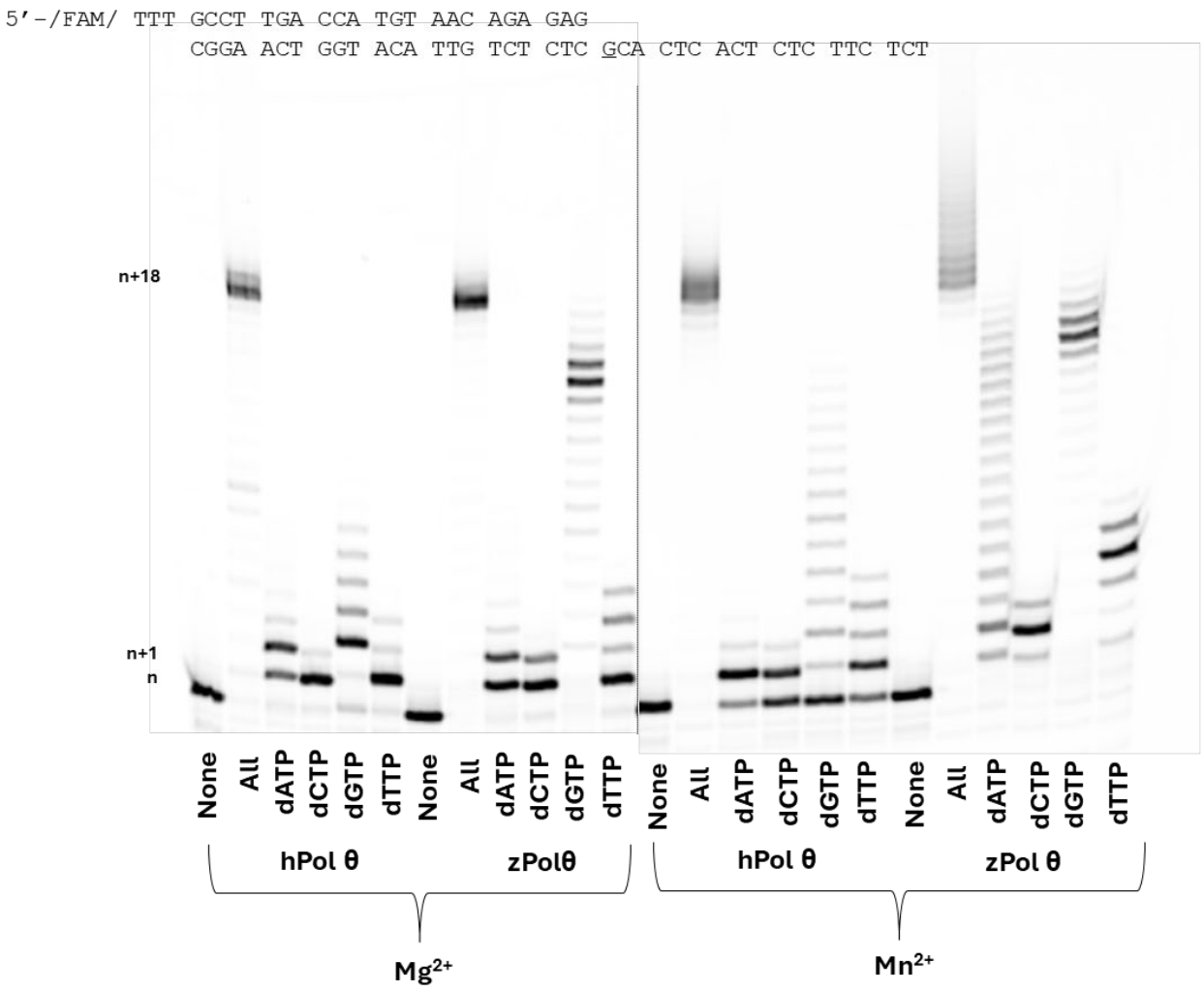
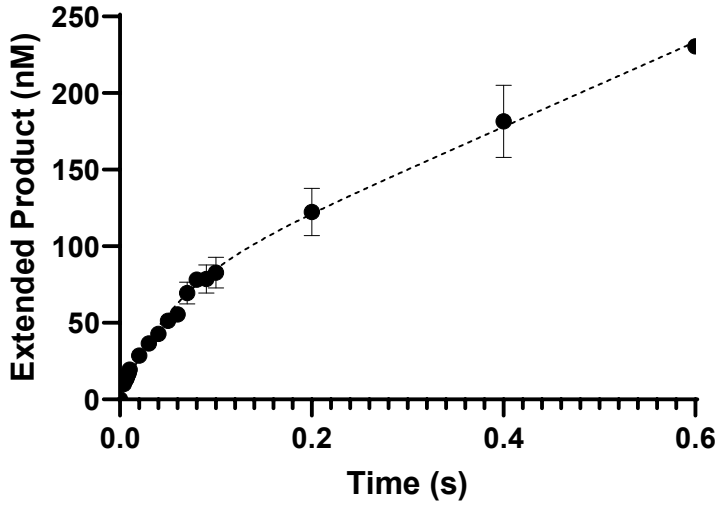


Figure 4. zPol θ experiences greater nucleotide extension activity compared to hPol θ. Under steady-state conditions 50 nM hPol θ or zPol θ were preincubated with 200 nM 25/40 dsDNA and combined with either 10 mM $MgCl_2$ or $MnCl_2$ for 5 minutes and 37°C. DNA extension products were separated on a denaturing gel and visualized on a Typhoon scanner. Each n+1 band represents an extension of one nucleotide following the DNA template as described above. N+1 would represent either correct nucleotide incorporation of dCTP opposite a templating G (underlined) or a misincorporation event of dATP, dGTP, or dTTP opposite templating G. Each subsequent band is another nucleotide extension with a maximum of template-dependent extension of n+18. Bands migrating higher than n+18 represent de novo synthesis.

zPol θ catalytic activity similar to other DNA polymerases

To further explore the mechanism of nucleotide incorporation of zebrafish Pol θ, we assayed zPol θ under presteady-state conditions in which there is an excess DNA substrate to enzyme with correct nucleotide. This assay focuses on the DNA polymerase

340 ability to extend DNA by incorporating the correct nucleotide opposite a templating base.
341 This activity is biphasic in which there is a rapid polymerization step of nucleotide
342 incorporation at the DNA primer's 3'OH and a slower, rate limiting step of product
343 release[39]. If biphasic activity is not observed, it suggests a step before nucleotide
344 incorporation is the rate-limiting step[41]. To ensure that purified c-terminal zPol θ follows
345 the traditional DNA polymerase mechanism, 100 nM zPol θ was preincubated with 300
346 nM 25/40 dsDNA. The DNA/Pol θ complex was rapidly combined with 100 μM correct
347 nucleotide and 10 mM $MgCl_2$ from 0.004-0.6 seconds. DNA products were separated on
348 a denaturing polyacrylamide gel and primer extension of n+1 was quantified and data fit
349 to a full biphasic burst equation. zPol θ fit to a biphasic equation with an observable
350 polymerization rate (k_{obs}) of 15.9 ± 2.5 s^{-1} (Figure 5).



351

352 **Figure 5. zPol θ experiences biphasic burst activity.** Zebrafish Pol θ (100 nM) was preincubated with 300 nM 25/40 dsDNA. The
353 DNA/Pol θ complex was rapidly combined with 100 μM dCTP (correct nucleotide) and 10 mM $MgCl_2$. Reactions were carried out at
354 37°C and quenched with 0.5 M EDTA. Products were separated on a denaturing gel and quantified with ImageQuant software. Data
355 were fit to a biphasic burst equation to obtain observed k_{obs} rates 15.9 ± s^{-1}. The slower rate k_{ss} was calculated to be 3.4 ± 0.46 s^{-1}.

356 *zPol θ performs MMEJ activity*

357 One of the major functions of DNA Pol θ is its ability to repair double-strand breaks
358 and is the primary DNA polymerase for microhomology-mediated end joining. In doing
359 so, Pol θ utilizes internal homology within the DNA sequence to act as a template. Pol θ
360 aligns these complementary pieces and extends in the 5' to 3' direction[37,42,43]. Truncated
361 hPol θ has been shown to able to perform MMEJ activity on short 12-mer single-stranded
362 DNA, but the full 290 kDa Pol θ with the N-terminal helicase and central domains are
363 needed to anneal and extend larger segments of DNA[37]. We wanted ensure that zPol θ
364 could also perform MMEJ in a similar manner to hPol θ on short fragments of DNA. Figure
365 6 is a representative gel of hPol θ and zPol θ performing MMEJ on a ssDNA. As indicated
366 in the schematic above, the CCCGGG are aligned through Pol θ in the presence of (+)
367 dNTP and subsequently extended in the opposite direction giving rise to a slower moving
368 double-stranded DNA product. Both hPol θ and zPol θ are able to perform this activity.
369 We hypothesize the smaller product bands are indicative of classic snap-back synthesis
370 in which the DNA substrate anneals onto itself for Pol θ to extend. This behavior has
371 been observed by others on hPol θ and there is little variation between the two species[37].
372
373
374
375
376
377

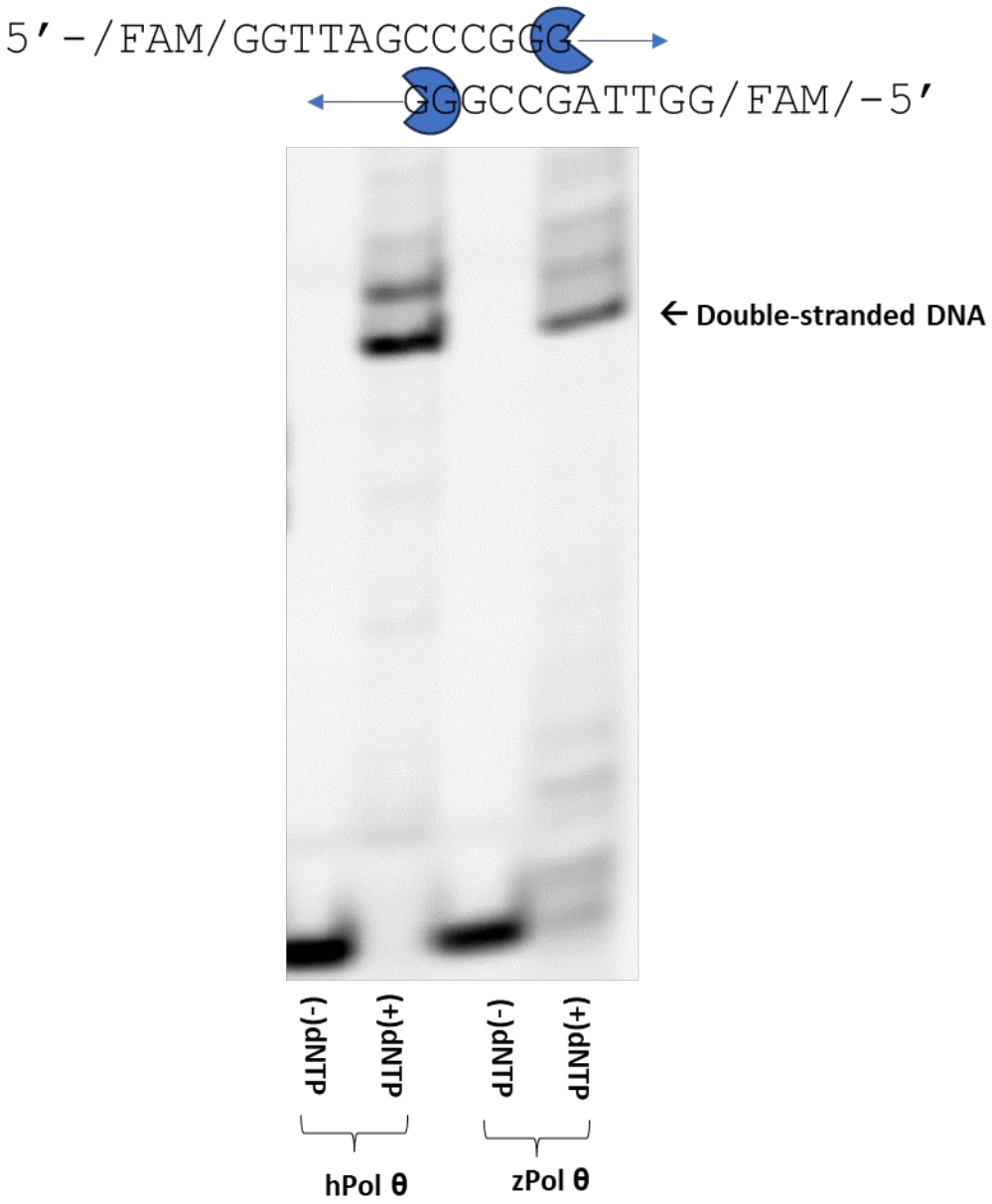


Figure 6. zPol θ is able to perform MMEJ activity of short DNA fragments. Pol θ (20 nM) was preincubated with 30 nM 5'-FAM ssDNA in reaction buffer. All nucleotides (+ dNTP) were added and the ternary complex was incubated for 45 minutes at 37°C. A no dNTP (-dNTP) control was carried out in the same manner. Reactions were stopped and products separated on a 12% Native PAGE. The gel was visualized on a Typhoon scanner.

zPol θ is able to bypass CPD lesion DNA.

Pol θ is a versatile DNA polymerase in not only can it perform MMEJ, it has also been shown to bypass cyclobutane pyrimidine dimers (CPD)[36]. By being able to extend

386 a DNA primer passed a template containing a contorted Thymine-Thymine lesion, human
387 and mouse Pol θ have been demonstrated to be critical in suppressing DNA damage and
388 preventing skin lesions. On a molecular level, human Pol θ has demonstrated that not
389 only can it insert opposite the initial T in the T-T dimer but is able to mutagenically extend
390 past this lesion for the remaining DNA template. We hypothesized that zPol θ has the
391 same ability to bypass CPD lesions and we assayed both Pol θ under single-turnover
392 conditions (4:1 protein to DNA) in the presence of a 24/33 CPD damaged DNA template
393 with both Mg^{2+} and Mn^{2+}. As predicted, there was little variance in bypass activity of zPol
394 θ compared to hPol θ. Both enzymes were able to readily insert opposite a T-T dimer as
395 well as extend past this lesion with both all dNTPs present and dATP and to some extent
396 dGTP. Both Pol θs could not incorporate dCTP opposite T-T, but we observed only
397 insertion of dTTP opposite and no extension. In the presence of Mn^{2+}, both Pol θs readily
398 bypassed T-T dimers, again demonstrating de novo synthesis past the template. zPol θ
399 was more robust in extension with the other incorrect nucleotides suggesting Mn^{2+} has
400 an increased mutagenic effect.

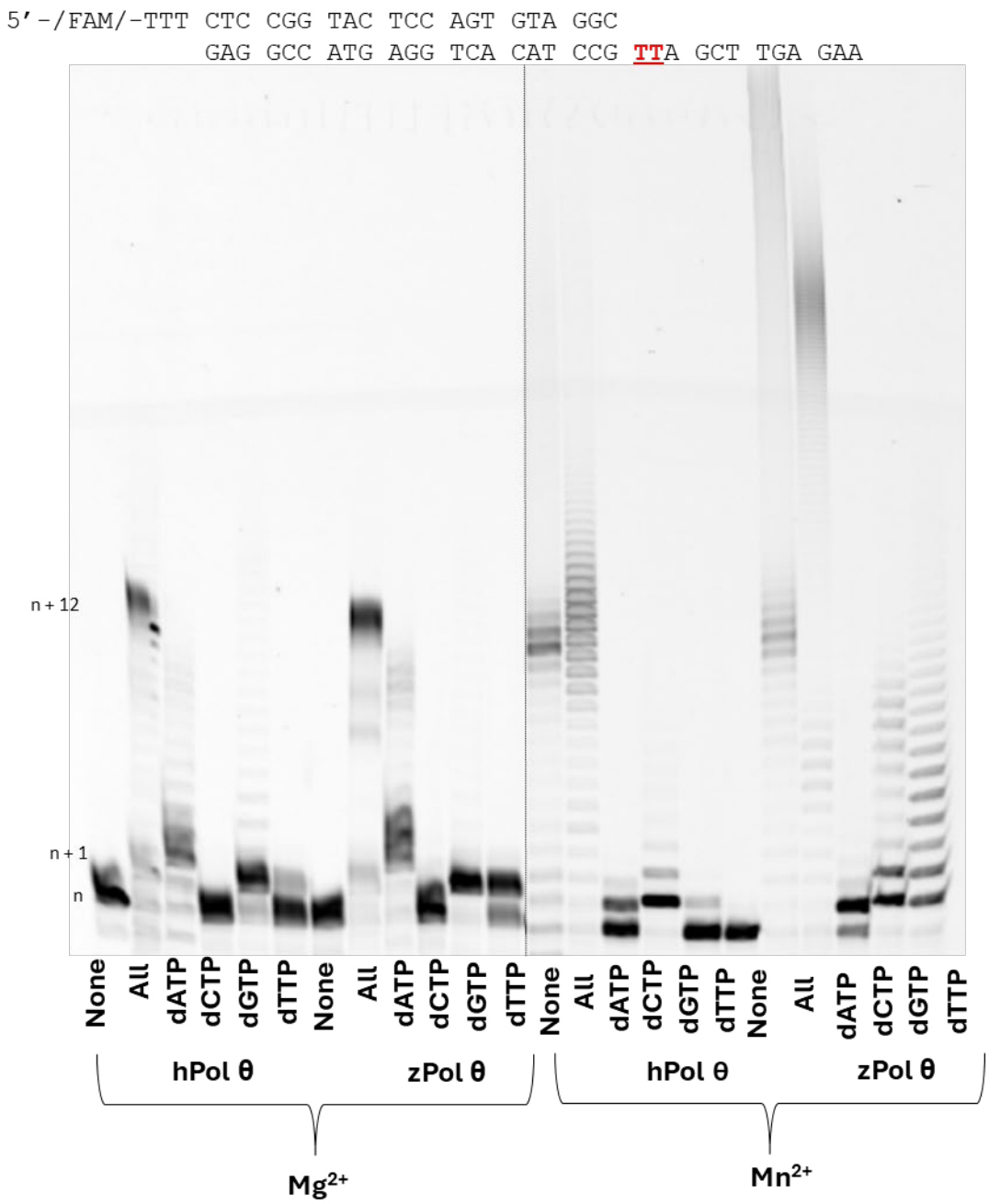


Figure 7. zPol θ is able to bypass CPD lesions. As described in the Materials and Methods, 200 nM Pol θ (human or zebrafish) was preincubated with 50 nM CPD Damaged DNA substrate. Reactions were initiated by the addition of 125 nM nucleotides as described and either 10 mM $MgCl_2$ or $MnCl_2$. Reactions were carried out at 37°C for 5 minutes and products visualized on a 12% denaturing gel. Higher migrating products are indicative of full extension (n+12) with smeared bands representing de novo synthesis with extension past n+12.

zPol θ experiences unusual extension of DNA substrates in the presence of Ca^{2+}.

409 To explore the role of divalent metals in DNA polymerase activity for Pol θ, we
410 performed a DNA polymerase extension assay again with either these assays with their
411 specific DNA substrates swapping out the active metals for Ca^{2+} which has traditionally
412 used as an inert control. Unlike the other divalent metals, Ca^{2+} allows for ternary complex
413 formation, but extension is limited or slow[44,45]. Using 50 nm if the 24/33 undamaged and
414 CPD damaged DNA substrate, we performed a primer extension assay with 200 nM hPol
415 θ or z WT Pol θ with Mg^{2+} substituted for $CaCl_2$. Extension products were separated on
416 a denaturing polyacrylamide gel and quantified based on the percent extension. Figure
417 8A is a representative gel of extension on 24/33 undamaged DNA template. We observe
418 that hPol θ could incorporate every nucleotide to some extent, with an n+3 extension
419 product only observed in the presence of all dNTP or purines. zPol θ was observed to
420 generate full extension product (94%) on this DNA template (n+12) with all nucleotides
421 present and, like hPol θ, could extend with purines as well. Incorporation of dATP led to
422 92% conversion to product although the enzyme stalled around n+2. Interestingly, zPol
423 θ appears to skip the first thymine in the undamaged sequence for both all nucleotides
424 and dATP. The same experiment was carried out with 24/33 CPD damaged DNA. Under
425 these conditions we report that Ca^{2+} reduced DNA polymerase activity for both hPol θ and
426 zPol θ with incorporation of only one nucleotide irrespective if that nucleotide was
427 matched or mismatched with the templating base.

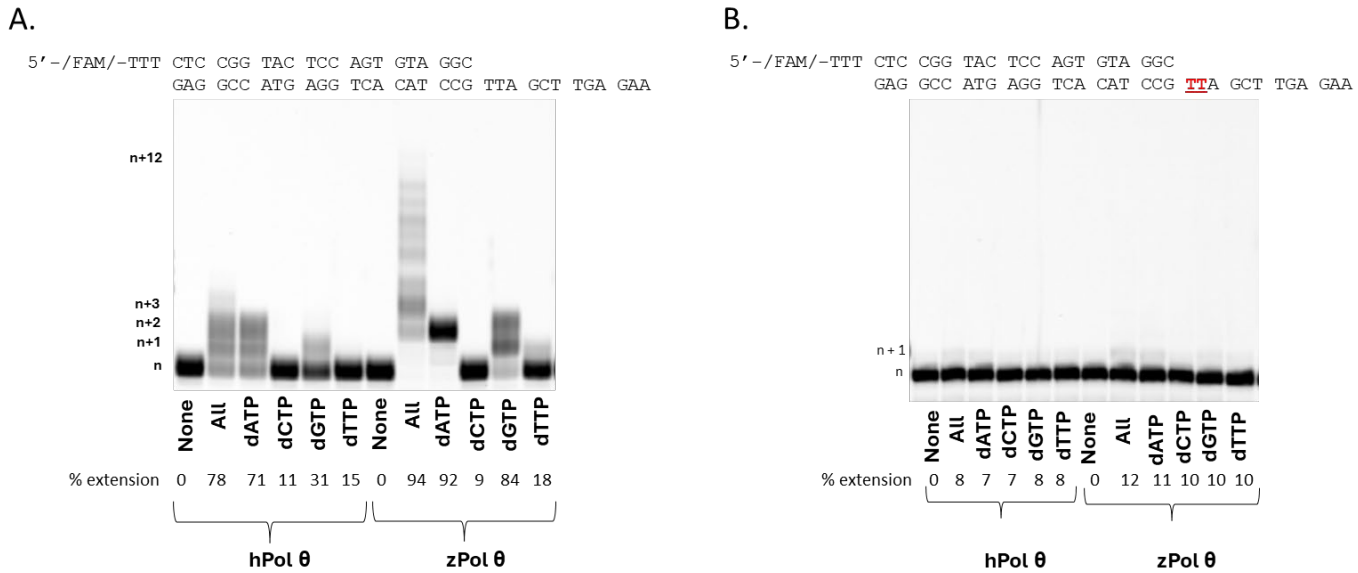


Figure 8. DNA Pol θ is able incorporate and extend dsDNA in the presence of Ca^{2+}. Under single turnover conditions 200 nM Pol θ (human or zebrafish) was preincubated with 50 nM 24/33 undamaged (A) or damaged (B) DNA substrate and reacted with 125 nM nucleotides in 10 mM $CaCl_2$. Reactions were carried out at 37°C for 5 minutes and products visualized on a 12% denaturing gel. Percent extended was calculated using ImageQuant software by quantifying the intensity of the extended products (n+1 and higher) divided by the intensity of the total amount of DNA.

DISCUSSION

Zebrafish and human Pol θ structures have a high degree of similarity

Comparisons of the amino acid sequence (table 1) and the structures (Figure 1) of zPol θ and hPol θ reveal that the two proteins share a high degree of similarity. Importantly, and perhaps not surprisingly, the areas of greatest similarity are around the catalytic subdomains, the fingers, thumb, and palm, of the polymerase domain. These sites of activity would be important to the protein function across evolutionary time. Interestingly, unlike other A-type polymerase family members Pol θ has three loop structures within the PD which have been identified to be important for function[12]. Comparison of the zPol θ and hPol θ sequences indicate that zPol θ also contains these inserted loops however, they contain little homology (Supplement 1) to that observed

446 between human and mouse[31] Pol θ. Despite this difference, expression and purification
447 yields of zPol θ were similar to that of hPol θ as were the secondary structural
448 characteristics and thermal stability (Figure 2); an early indication similar protein folding.
449 We also show here that zPol θ still retains the same activity observed in hPol θ.
450
451 *Zebrafish Pol θ extends dsDNA*
452 DNA polymerase θ primary role in the human cell is the primary DNA polymerase
453 repair enzyme in microhomology-mediated end joining and thus, the major goal of this
454 study was to query if zPol θ retained a similar function. Initially we simplify the activity by
455 asking 'can zPol θ bind to a primer/template dsDNA substrate and then extend it?' Our
456 data suggests that yes it can. zPol θ binds tightly to this substrate (Figure 3)[34] similar to
457 values obtained with hPol θ. This is expected because loop 1 which is located in the
458 thumb domain or DNA binding domain is thought to be involved with contacts to DNA[12]
459 and is the only conserved loop region between zPol θ and hPol θ. zPol θ can in fact
460 extend a DNA substrate, and we show it has robust activity on this particular DNA
461 substrate (Figure 4) especially in the presence of Mn^{2+}. While most of the data presented
462 in this work was qualitative, probing how fast a DNA polymerase makes a phosphodiester
463 bound through biochemical kinetics can provide insight into mechanism of
464 incorporation[39]. Like most DNA polymerases, zPol θ performs biphasic burst kinetics
465 which is indicative of a two-step mechanism with a rate limiting step of product release.
466 zPol θ experiences an observed polymerization rate of around 16 s^{-1} (Figure 5) which is
467 almost 4 times slower than its human ortholog[34], but not uncommonly slow as a similar
468 DNA Polymerase β experiences a similar rate[46–48]. Why zPol θ might experience a slower

469 rate is unknown. It could be due to the lack of conservation within the looping structures
470 in the palm domain which in human Pol θ may drive substrate alignment for rapid
471 polymerization.
472
473 *Zebrafish Pol θ is able to perform microhomology-mediated End Joining*
474 Despite being able to extend DNA, it is important that zPol θ also be able to perform
475 MMEJ as in humans it is its primary function. Although a majority of human Pol θ's N-
476 terminal and central domains are critical for this function, studies have shown that the c-
477 terminal polymerase domain of Pol θ does retain limited function for aligning and
478 extending short single-stranded DNA[37]. zPol θ was no exception (Figure 6) and was able
479 to complement two single-strands and extend which is perhaps the most compelling
480 evidence of homologous function.
481
482 *zPol θ can bypass CPD lesions*
483 Another function of human Pol θ is its ability to bypass DNA damage and a more
484 recent study in mice suggest that bypassing UV damage is critical in the prevention of
485 skin cancer[36]. Our data provides evidence that zPol θ is able to bypass and extend CPD
486 lesions *in vivo* similar to that of hPol θ (Figure 7). Translesion bypass activity has been
487 highlighted as a function of loops 2 and 3 in human Pol θ[12,31]. Surprisingly, zebrafish
488 display very little homology through similar inserts (Supplement 1). However, our studies
489 might suggest that the only critical residues for this function are isolated to the c-terminal
490 end of this insert beginning with the sequence GMXFSXSMR. Further studies exploring
491 this insert in zebrafish are needed to determine if these conserved amino acids are truly

492 critical and that the function is either dictated through the overall presence of the loop or
493 that the loop dependent activities require the few conserved amino acids retain in zPol θ.
494
495 *zPol θ experiences Ca^{2+} dependent polymerization*
496 Our data shows that zPol θ retains all of the DNA polymerase activities of hPol θ,
497 with the exception being the robust Ca^{2+} dependency during DNA extension (Figure 8).
498 While it is unusual to see DNA polymerases extend past the initial insertion event, there
499 have been instances where high-fidelity *Sulfolobus solfataricus* Dpo4 polymerase uses
500 calcium [49]. Whether or not swapping Ca^{2+} in zebrafish makes it a faster or mutagenic
501 polymerase has not been explored but evolutionarily it is possible that Ca^{2+} makes for a
502 stable ion swap for structural alignment and catalytic activity. Given that in freshwater
503 contains nearly identical concentrations of the divalent ions[50], its plausible that Ca^{2+} could
504 readily be a co-factor substitute.
505 Our data presented is clear evidence that zebrafish Pol θ is a homolog to human
506 Pol θ and that structurally and enzymatically behave with similar functions. This study is
507 significant as it highlights the availability of zebrafish as a model organism for studying
508 Pol θ and its potential function in DNA repair and disease. In particular given the robust
509 assortment of tools zebrafish offers a powerful, functionally relevant model for human
510 melanoma. Future studies could introduce patient derived mutations by way of germline
511 alterations and study the effects over the life time of the animal. Thus, adding new insights
512 into potential disease markers and mechanisms of disease progression and treatment.
513
514 **ACKNOWLEDGEMENTS**

We would like to thank Sarah Delaney and Mary Tarantino from Brown University for the generation of the CPD damaged DNA. Thank you to Sylvie Doublié from University of Vermont for the human pol θ plasmid.

FUNDING

Research reported in this publication was supported by the Rhode Island Institutional Development Award (IDeA) Network of Biomedical Research Excellence under P20GM103430 and in part by the National Institute of General Medical Sciences of the National Institutes of Health under grant number R15GM144903-01.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

527 **WORK CITED**

528 1. Arnold, M. *et al.* Global Burden of Cutaneous Melanoma in 2020 and Projections to
529 2040. *JAMA Dermatol* **158**, 495–503 (2022).
530 2. Mort, R. L., Jackson, I. J. & Patton, E. E. The melanocyte lineage in development
531 and disease. *Development* **142**, 620–632 (2015).
532 3. Jilani, S. *et al.* CAR-T cell therapy targeting surface expression of TYRP1 to treat
533 cutaneous and rare melanoma subtypes. *Nat Commun* **15**, 1244 (2024).
534 4. Larkin, J. *et al.* Combined Nivolumab and Ipilimumab or Monotherapy in Previously
535 Untreated Melanoma. *N Engl J Med* **373**, 23–34 (2015).
536 5. Farshidfar, F. *et al.* Integrative molecular and clinical profiling of acral melanoma
537 links focal amplification of 22q11.21 to metastasis. *Nat Commun* **13**, 898 (2022).
538 6. Thyme, S. B. & Schier, A. F. Polq-mediated end joining is essential for surviving DNA
539 double-strand breaks during early zebrafish development. *Cell Rep* **15**, 707–714
540 (2016).
541 7. Brambati, A., Barry, R. M. & Sfeir, A. DNA polymerase theta (Polθ) – an error-prone
542 polymerase necessary for genome stability. *Current Opinion in Genetics &*
543 *Development* **60**, 119–126 (2020).
544 8. Yoon, J.-H. *et al.* Error-Prone Replication through UV Lesions by DNA Polymerase θ
545 Protects against Skin Cancers. *Cell* **176**, 1295-1309.e15 (2019).
546 9. Chan, S. H., Yu, A. M. & McVey, M. Dual Roles for DNA Polymerase Theta in
547 Alternative End-Joining Repair of Double-Strand Breaks in Drosophila. *PLOS*
548 *Genetics* **6**, e1001005 (2010).

549 10. Kent, T., Chandramouly, G., McDevitt, S. M., Ozdemir, A. Y. & Pomerantz, R. T.
550 Mechanism of microhomology-mediated end-joining promoted by human DNA
551 polymerase θ. *Nat Struct Mol Biol* **22**, 230–237 (2015).
552 11. Yousefzadeh, M. J. *et al.* Mechanism of Suppression of Chromosomal Instability by
553 DNA Polymerase POLQ. *PLOS Genetics* **10**, e1004654 (2014).
554 12. Hogg, M., Seki, M., Wood, R. D., Doublié, S. & Wallace, S. S. Lesion Bypass Activity
555 of DNA Polymerase θ (POLQ) Is an Intrinsic Property of the Pol Domain and
556 Depends on Unique Sequence Inserts. *Journal of Molecular Biology* **405**, 642–652
557 (2011).
558 13. Yu, A. M. & McVey, M. Synthesis-dependent microhomology-mediated end joining
559 accounts for multiple types of repair junctions. *Nucleic Acids Res* **38**, 5706–5717
560 (2010).
561 14. Ceccaldi, R. *et al.* Homologous-recombination-deficient tumours are dependent on
562 Polθ-mediated repair. *Nature* **518**, 258–262 (2015).
563 15. Goff, J. P. *et al.* Lack of DNA Polymerase θ (POLQ) Radiosensitizes Bone Marrow
564 Stromal Cells In Vitro and Increases Reticulocyte Micronuclei after Total-Body
565 Irradiation. *Radiat Res* **172**, 165–174 (2009).
566 16. Goullet de Rugy, T. *et al.* Excess Polθ functions in response to replicative stress in
567 homologous recombination-proficient cancer cells. *Biology Open* **5**, 1485–1492
568 (2016).
569 17. Allera-Moreau, C. *et al.* DNA replication stress response involving PLK1, CDC6,
570 POLQ, RAD51 and CLASPIN upregulation prognoses the outcome of early/mid-
571 stage non-small cell lung cancer patients. *Oncogenesis* **1**, e30–e30 (2012).

572 18. Lemée, F. *et al.* DNA polymerase θ up-regulation is associated with poor survival in
573 breast cancer, perturbs DNA replication, and promotes genetic instability.
574 *Proceedings of the National Academy of Sciences* **107**, 13390–13395 (2010).
575 19. Seki, M. *et al.* High-efficiency bypass of DNA damage by human DNA polymerase
576 Q. *The EMBO Journal* **23**, 4484–4494 (2004).
577 20. Thomas, C. *et al.* Melanoma-Derived DNA Polymerase Theta Variants Exhibit
578 Altered DNA Polymerase Activity. *Biochemistry* **63**, 1107–1117 (2024).
579 21. Patton, E. E. *et al.* BRAF mutations are sufficient to promote nevi formation and
580 cooperate with p53 in the genesis of melanoma. *Curr Biol* **15**, 249–254 (2005).
581 22. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the
582 human genome. *Nature* **496**, 498–503 (2013).
583 23. Groenewoud, A. *et al.* Patient-derived zebrafish xenografts of uveal melanoma
584 reveal ferroptosis as a drug target. *Cell Death Discov.* **9**, 1–11 (2023).
585 24. Fazio, M., Ablain, J., Chuan, Y., Langenau, D. M. & Zon, L. I. Zebrafish patient
586 avatars in cancer biology and precision cancer therapy. *Nat Rev Cancer* **20**, 263–
587 273 (2020).
588 25. Madeira, F. *et al.* The EMBL-EBI Job Dispatcher sequence analysis tools framework
589 in 2024. *Nucleic Acids Res* **52**, W521–W525 (2024).
590 26. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods*
591 **19**, 679–682 (2022).
592 27. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*
593 (2021) doi:10.1038/s41586-021-03819-2.

28. Mirdita, M., Steinegger, M. & S"oding, J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics* **35**, 2856–2858 (2019).
29. Mirdita, M. *et al.* Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* **45**, D170–D176 (2017).
30. Mitchell, A. L. *et al.* MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz1035.
31. Zahn, K. E., Averill, A. M., Aller, P., Wood, R. D. & Doublié, S. Human DNA polymerase θ grasps the primer terminus to mediate DNA repair. *Nat Struct Mol Biol* **22**, 304–311 (2015).
32. Li, Z., Jaroszewski, L., Iyer, M., Sedova, M. & Godzik, A. FATCAT 2.0: towards a better understanding of the structural diversity of proteins. *Nucleic Acids Res* **48**, W60–W64 (2020).
33. Bittrich, S., Segura, J., Duarte, J. M., Burley, S. K. & Rose, Y. RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments. *Bioinformatics* **40**, btae370 (2024).
34. Thomas, C. *et al.* Melanoma-Derived DNA Polymerase Theta Variants Exhibit Altered DNA Polymerase Activity. *Biochemistry* **63**, 1107–1117 (2024).
35. Prasad, R. *et al.* Human DNA polymerase θ possesses 5′-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Research* **37**, 1868–1877 (2009).
36. Yoon, J. H. *et al.* Error-Prone Replication through UV Lesions by DNA Polymerase θ Protects against Skin Cancers. *Cell* **176**, 1295-1309.e15 (2019).

616 37. Black, S. J. *et al.* Molecular basis of microhomology-mediated end-joining by purified
617 full-length Polθ. *Nature Communications* **10**, (2019).
618 38. Steitz, T. A. DNA Polymerases: Structural Diversity and Common Mechanisms *.
619 *Journal of Biological Chemistry* **274**, 17395–17398 (1999).
620 39. Johnson, K. A. Transient-State Kinetic Analysis of Enzyme Reaction Pathways.
621 *Enzymes* **20**, 1–61 (1992).
622 40. Kent, T., Mateos-Gomez, P. A., Sfeir, A. & Pomerantz, R. T. Polymerase θ is a robust
623 terminal transferase that oscillates between three different mechanisms during end-
624 joining. *eLife* **5**, e13740.
625 41. Joyce, C. M. Techniques used to study the DNA polymerase reaction pathway.
626 *Biochimica et biophysica acta* **1804**, 1032–1040 (2010).
627 42. Brambati, A., Barry, R. M. & Sfeir, A. DNA polymerase theta (Polθ) – an error-prone
628 polymerase necessary for genome stability. *Current Opinion in Genetics and*
629 *Development* **60**, 119–126 (2020).
630 43. Li, C. *et al.* Structural basis of DNA polymerase θ mediated DNA end joining. *Nucleic*
631 *Acids Research* **51**, 463–474 (2023).
632 44. Freudenthal, B. D., Beard, W. A., Shock, D. D. & Wilson, S. H. Observing a DNA
633 Polymerase Choose Right from Wrong. *Cell* **154**, 157–168 (2013).
634 45. Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. J. & Yang, W. Watching DNA
635 polymerase eta make a phosphodiester bond. *Nature* **487**, 196–201 (2012).
636 46. Towle-Weicksel, J. B. *et al.* Fluorescence resonance energy transfer studies of DNA
637 polymerase β the critical role of fingers domain movements and a novel non-

638 covalent step During nucleotide selection. *Journal of Biological Chemistry* **289**,
639 16541–16550 (2014).
640 47. Werneburg, B. G. *et al.* DNA Polymerase β: Pre-Steady-State Kinetic Analysis and
641 Roles of Arginine-283 in Catalysis and Fidelity†. *Biochemistry* **35**, 7041–7050
642 (1996).
643 48. Murphy, D. L., Kosa, J., Jaeger, J. & Sweasy, J. B. The Asp285 Variant of DNA
644 Polymerase Beta Extends Mispaired Primer Termini via Increased Nucleotide
645 Binding. *Biochemistry* **47**, 8048–8057 (2008).
646 49. Ralec, C., Henry, E., Lemor, M., Killelea, T. & Henneke, G. Calcium-driven DNA
647 synthesis by a high-fidelity DNA polymerase. *Nucleic Acids Res* **45**, 12425–12440
648 (2017).
649 50. Guerin, A. J. *et al.* High-resolution freshwater dissolved calcium and pH data layers
650 for Canada and the United States. *Sci Data* **11**, 370 (2024).
651