Full Title: Zebrafish Polymerase Theta and human Polymerase Theta: orthologues with homologous function.

Short Title: Biochemical analysis of zebrafish DNA Polymerase Theta

Corey Thomas[1], Sydney Green[1], Lily Kimball[2], Isaiah R 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, key loop regions were identified to play an important role in key functions of Pol θ.

Here we present a comparative structure-function study of the polymerase domain of zebrafish and human Pol θ. We show that these two proteins share a large amount of sequence and structural homology. However, we identify differences in the amino acid composition within the key loop areas shown to drive characteristic Pol θ functions. Despite these differences zebrafish Pol θ still displays characteristics identify in human Pol θ, including DNA template extension in the presence of different divalent metals, microhomology-mediated end joining, and translesion synthesis. These results will support future studies looking to gain insight into Pol θ function on the basis of evolutionarily conserved features.

INTRODUCTION

A cell's genome is damaged at a nearly continuous rate due to a combination of internal (i.e. reactive oxygen species and cellular processes) and external factors (i.e. ultra-violet radiation, environmental exposure). Left unresolved, DNA damage has the ability to alter cell function through disruption of genomic stability. To preserve genomic integrity a robust system of DNA repair enzymes has evolved an essential role in fixing

the genome and protecting the cell from aberrant function. However, while some repair mechanisms faithfully preserve sequence (i.e. homologous recombination) some sacrifice fidelity and generate mutations (i.e. nonhomologous end joining) to keep the genome together. Despite this mutagenic behavior, these mechanism help avoid cell death due to genomic disfunction due to fragmentation. Unsurprisingly, this ability of DNA repair enzymes to avoid cell death while also driving mutagenesis has made DNA repair enzymes key factors in carcinogenesis. To date, every known DNA repair polymerase and many of their cofactors have been linked to cancer. This has made understanding how DNA repair polymerases work and the factors that regulate their activity of great interest.

One emerging factor identified in a recent study of patient derived melanoma samples is DNA polymerase-theta (Pol θ or POLQ) (1). An A-family DNA repair enzyme, Pol θ is essential for cell function and organismal development (2). Inherently error prone Pol θ (3,4) plays a predominant role repairing double strand breaks (DSB) in the DNA strand through microhomology-mediated end joining (MMEJ, also known as theta-mediated end joining, (TMEJ)), and translesion nucleotide bypass (5–8). Unlike homologous recombination (HR), the favored DSB repair pathway, TMEJ is highly error prone and is proposed to be activated when HR is overwhelmed (when the genome occurs many double strand breaks) and/or inactive (such as in cancer states). The activity of Pol θ in translesion nucleotide bypass plays a critical role in replication, that while perpetuating genomic mutations, allows replicative DNA polymerases to continue replication while also avoiding more DNA DSBs and potential mutagenesis through replication fork collapse (9). Together this indicates that Pol θ function is intrinsically

mutagenic yet required for cell function. This duality, mutagenic enzymatic behavior while also supporting cell survival (2,10,11), along with aberrant Pol θ activity in cancer cells (12–14) that has led to many hypothesizing that Pol θ activity drives carcinogenesis. However, few models for assessing the function of Pol θ as well as the outcome of Pol θ function in the context of an organism exist.

Zebrafish (*Danio* rerio) have long been employed to model organism and would represent a powerful tool to better understand the function of Pol θ in the context of an organism. With a high degree of similarly to humans, zebrafish have nearly 70% homology in their genes and 85% in human disease-related genes (15). Zebrafish have a predicted Pol θ orthologue of 2576 amino acid residues and has been shown to be essential for fixing double strand breaks during embryonic development (2). However, little is known about the function of zebrafish Pol θ and if it functions similarity to that of its human orthologue.

Here we present the first comparative analysis of protein structure and function of purified zebrafish POLQ (zPol θ or zPOLQ) and human POLQ (hPol θ or hPOLQ) polymerase domains. Protein alignment indicates that many of the residues present in the polymerase domain between the two proteins are conserved resulting in similar folded structures. However, within loop regions (unresolved in the human crystal structure), specific to PolQ relative to other A-type proteins, there is little conservation. Despite this lack of conservation, we observe similar zPOLQ behavior compared to hPOLQ. zPolQ can extend DNA templates even in the presence of conventionally inhibitory Ca^{2+}, perform TMEJ, and bypass DNA lesions, hallmarks of PolQ function in the cell.

RESULTS

Zebrafish and human polymerase domains display high degree of structural similarity

To determine the degree of similarity between the zPol θ and hPol θ PD primary amino acid sequences were aligned (Table 1, Sfig 1). The alignment of the full-length Pol θ protein indicates, zebrafish and human Pol θ share 46% identity. This degree of similarity increases when comparing the predicted polymerase domain (63%), as well as subdomains containing catalytic activity, fingers (75.3%), thumb (74.6%), and palm (66.7%). These data suggest a structurally similar molecule.

Sfig 1 CLUSTAL O(1.2.4) zPol θ and hPol θ alignment.

Red indicates loop insertions

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

To assess the extent of structural similarity we generated a predicted structure for zPol θ PD using ColabFold (16) to compare to the solved crystal structure of hPol θ PD(17). Upon visual inspection the predicted zPol θ PD displays classical DNA polymerase PD structures (Fig 1). The three major subunits, the fingers, thumb, and palm are visible, and when modeled in, a DNA molecule can fit in the presumed catalytic domain. The model also indicates the presence of unstructured loop domains, that have functional importance (8), that were not resolved in the hPol θ PD structure. An over lay

of the hPol θ PD and zPol θ PD show that the structures have a high degree of similarity. As predicted by the amino-acid alignment, these data indicate that much of the structure of the hPol θ PD is conserved in zPol θ.

Fig 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 (17) 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 (18) 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 (Fig 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 similar secondary characteristics and thermal stability with a Tm of about 55°C.

Fig 2. hPol θ and zPol θ display similar secondary characteristics and thermal stability.

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

zPol θ binds to dsDNA substrate

DNA binding by a DNA polymerase is one of the first steps in its catalytic mechanism. To determine the DNA binding capabilities of zPol θ, we titrated zPol θ from 0-1000 nM protein against 10 nM 25/40 dsDNA (Fig 3). Complexed DNA/protein products were separated on a denaturing gel to determine a dissociation constant ($K_{D(DNA)}$) for DNA binding. Similar to hPol θ, zPol θ has a low $K_{D(DNA)}$ value of approximately 19.8 ± 3.1 nM.

Fig 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 products were separated on a 6% non-denaturing gel and quantified using ImageQuant. $K_{D(DNA)}$ was mathematically calculated using Equation 1 and is the midpoint between bound and unbound fractions.

zPol θ can extend dsDNA similar to hPol θ

The second step in the DNA polymerase catalytic pathway is nucleotide binding and formation of the phosphodiester bond. To explore this fundamental step of DNA

158 Polymerase activity, we assayed zPol θ's ability to extend 25/40 dsDNA under varying
159 conditions. Under standard steady-state conditions, 200 nM of zPol θ or hPol θ was pre-
160 incubated with 50 nM 25/40 dsDNA. The reaction was initiated by the addition of 125 nM
161 dNTP as described in Figure 4 along with 20 mM $MgCl_2$, the preferred metal for DNA
162 polymerase (19). We observed under these conditions both hPol θ and zPol θ were able
163 to extend the full 18-mer template with all nucleotides present (Fig 4). Both enzymes
164 were able to incorporate single nucleotides, correct and incorrect as well. Notably, zPol θ
165 was able to incorporate incorrect dGTP to full extension (n+1) compared to only n+6 with
166 hPol θ. DNA polymerases can utilize other metals including Mn^{2+}, and we observed an
167 increase in mutagenesis through misincorporation for both hPol θ and zPol θ. When
168 provided with all dNTP, zPol θ can extend past the end of the template (n+18). Overall,
169 zPol θ experiences more extension products especially with incorrect nucleotides dATP,
170 dGTP, and dTTP compared with hPol θ under similar conditions. Steady-state conditions
171 highlight overall DNA polymerase activity, but because the dsDNA substrate is in excess,
172 activity highlights multiple turnovers (20). Although DNA pol θ has been shown to have
173 robust de novo activity with manganese (21), we wanted to be sure this over extension
174 observed with zPol θ and Mn^{2+} was the result of extension and not an artifact. We
175 changed the ratio of protein to DNA to reflect single-turnover conditions; excess protein
176 over dsDNA substrate. Here we are able to observe polymerization events for
177 theoretically every available DNA substrate. Similar to steady-state conditions, we
178 observe an even more robust de novo extension with not only all nucleotides, but also
179 with dATP, suggesting that zPol θ misincorporation with dATP is preferred (SFig 2).

180 **Fig 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.

SFig 2. Both hWT and zWT Pol θ were assayed under single-turnover conditions at t 4:1 ratio protein:DNA (see Materials and Methods). Pol θ and 25/40 dsDNA were preincubated 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.

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 ability to extend DNA by incorporating the correct nucleotide opposite a templating base. This activity is biphasic in which there is a rapid polymerization step of nucleotide incorporation at the DNA primer's 3'OH and a slower, rate limiting step of product release (20). If biphasic activity is not observed, it suggests a step before nucleotide incorporation

is the rate-limiting step(22). To ensure that purified c-terminal zPol θ follows the traditional DNA polymerase mechanism, 100 nM zPol θ was preincubated with 300 nM 25/40 dsDNA. The DNA/Pol θ complex was rapidly combined with 100 μM correct nucleotide and 10 mM $MgCl_2$ from 0.004-0.6 seconds. DNA products were separated on a denaturing polyacrylamide gel and primer extension of n+1 was quantified and data fit to a full biphasic burst equation. zPol θ fit to a biphasic equation with an observable polymerization rate (k_{obs}) of 15.9 ± 2.5 s^{-1} (Fig 5).

Fig 5. zPol θ experiences biphasic burst activity.

Zebrafish Pol θ (100 nM) was preincubated with 300 nM 25/40 dsDNA. The DNA/Pol θ complex was rapidly combined with 100 μM dCTP (correct nucleotide) and 10 mM $MgCl_2$. Reactions were carried out at 37°C and quenched with 0.5 M EDTA. Products were separated on a denaturing gel and quantified with ImageQuant software. Data 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}.

zPol θ performs MMEJ activity

One of the major functions of DNA Pol θ is its ability to repair double-strand breaks and is the primary DNA polymerase for microhomology-mediated end joining. In doing so, Pol θ utilizes internal homology within the DNA sequence to act as a template. Pol θ aligns these complementary pieces and extends in the 5' to 3' direction(23–25). Truncated hPol θ has been shown to able to perform MMEJ activity on short 12-mer single-stranded DNA, but the full 290 kDa Pol θ with the N-terminal helicase and central

domains are needed to anneal and extend larger segments of DNA(23). We wanted ensure that zPol θ could also perform MMEJ in a similar manner to hPol θ on short fragments of DNA. Figure 6 is a representative gel of hPol θ and zPol θ performing MMEJ on a ssDNA. As indicated in the schematic above, the CCCGGG are aligned through Pol θ in the presence of (+) dNTP and subsequently extended in the opposite direction giving rise to a slower moving double-stranded DNA product. Both hPol θ and zPol θ are able to perform this activity. We hypothesize the smaller product bands are indicative of classic snap-back synthesis in which the DNA substrate anneals onto itself for Pol θ to extend. This behavior has been observed by others on hPol θ and there is little variation between the two species (23).

Fig 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)(26). By being able to extend a DNA primer passed a template containing a contorted Thymine-Thymine lesion, human and mouse Pol θ have been demonstrated to be critical in suppressing DNA damage and preventing skin lesions. On a molecular level, human Pol θ has demonstrated that not only can it insert opposite the initial T in the T-T dimer but is able to mutagenically extend

past this lesion for the remaining DNA template. We hypothesized that zPol θ has the same ability to bypass CPD lesions and we assayed both Pol θ under single-turnover conditions (4:1 protein to DNA) in the presence of a 24/33 CPD damaged DNA template with both Mg^{2+} and Mn^{2+}. As predicted, there was little variance in bypass activity of zPol θ compared to hPol θ (Fig 7). Both enzymes were able to readily insert opposite a T-T dimer as well as extend past this lesion with both all dNTPs present and dATP and to some extent dGTP. Both Pol θs could not incorporate dCTP opposite T-T, but we observed only insertion of dTTP opposite and no extension. In the presence of Mn^{2+}, both Pol θs readily bypassed T-T dimers, again demonstrating de novo synthesis past the template. zPol θ was more robust in extension with the other incorrect nucleotides suggesting Mn^{2+} has an increased mutagenic effect.

Fig 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+}.

To explore the role of divalent metals in DNA polymerase activity for Pol θ, we performed a DNA polymerase extension assay again with either these assays with their specific DNA substrates swapping out the active metals for Ca^{2+} which has traditionally

used as an inert control. Unlike the other divalent metals, Ca^{2+} allows for ternary complex formation, but extension is limited or slow (27,28). Using 50 nm if the 24/33 undamaged and CPD damaged DNA substrate, we performed a primer extension assay with 200 nM hPol θ or z WT Pol θ with Mg^{2+} substituted for $CaCl_2$. Extension products were separated on a denaturing polyacrylamide gel and quantified based on the percent extension. Figure 8A is a representative gel of extension on 24/33 undamaged DNA template. We observe that hPol θ could incorporate every nucleotide to some extent, with an n+3 extension product only observed in the presence of all dNTP or purines. zPol θ was observed to generate full extension product (94%) on this DNA template (n+12) with all nucleotides present and, like hPol θ, could extend with purines as well. Incorporation of dATP led to 92% conversion to product although the enzyme stalled around n+2. Interestingly, zPol θ appears to skip the first thymine in the undamaged sequence for both all nucleotides and dATP. The same experiment was carried out with 24/33 CPD damaged DNA. Under these conditions we report that Ca^{2+} reduced DNA polymerase activity for both hPol θ and zPol θ with incorporation of only one nucleotide irrespective if that nucleotide was 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 (Fig 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 (8). Comparison of the zPol θ and hPol θ sequences indicate that zPol θ also contains these inserted loops however, they contain little homology (SFig 1) to that observed between human and mouse (17) Pol θ. Despite this difference, expression and purification yields of zPol θ were similar to that of hPol θ as were the secondary structural characteristics and thermal stability (Fig 2); an early indication similar protein folding. We also show here that zPol θ still retains the same activity observed in hPol θ.

Zebrafish Pol θ extends dsDNA

DNA polymerase θ primary role in the human cell is the primary DNA polymerase repair enzyme in microhomology-mediated end joining and thus, the major goal of this study was to query if zPol θ retained a similar function. Initially we simplify the activity by asking ‘can zPol θ bind to a primer/template dsDNA substrate and then extend it?’ Our data suggests that yes it can. zPol θ binds tightly to this substrate (Fig 3) (18) similar to values obtained with hPol θ. This is expected because loop 1 which is located in the

thumb domain or DNA binding domain is thought to be involved with contacts to DNA (8) and is the only conserved loop region between zPol θ and hPol θ. zPol θ can in fact extend a DNA substrate, and we show it has robust activity on this particular DNA substrate (Fig 4) especially in the presence of Mn^{2+}. While most of the data presented in this work was qualitative, probing how fast a DNA polymerase makes a phosphodiester bound through biochemical kinetics can provide insight into mechanism of incorporation (20). Like most DNA polymerases, zPol θ performs biphasic burst kinetics which is indicative of a two-step mechanism with a rate limiting step of product release. zPol θ experiences an observed polymerization rate of around 16 s^{-1} (Figure 5) which is almost 4 times slower than its human ortholog (18), but not uncommonly slow as a similar DNA Polymerase β experiences a similar rate (29–31). Why zPol θ might experience a slower rate is unknown. It could be due to the lack of conservation within the looping structures in the palm domain which in human Pol θ may drive substrate alignment for rapid polymerization.

Zebrafish Pol θ is able to perform microhomology-mediated End Joining

Despite being able to extend DNA, it is important that zPol θ also be able to perform MMEJ as in humans it is its primary function. Although a majority of human Pol θ's N-terminal and central domains are critical for this function, studies have shown that the c-terminal polymerase domain of Pol θ does retain limited function for aligning and extending short single-stranded DNA (23). zPol θ was no exception (Figure 6) and was able to complement two single-strands and extend which is perhaps the most compelling evidence of homologous function.

zPol θ can bypass CPD lesions

Another function of human Pol θ is its ability to bypass DNA damage and a more recent study in mice suggest that bypassing UV damage is critical in the prevention of skin cancer (26). Our data provides evidence that zPol θ is able to bypass and extend CPD lesions *in vivo* similar to that of hPol θ (Fig 7). Translesion bypass activity has been highlighted as a function of loops 2 and 3 in human Pol θ (8,17). Surprisingly, zebrafish display very little homology through similar inserts (SFig 1). However, our studies might suggest that the only critical residues for this function are isolated to the c-terminal end of this insert beginning with the sequence GMXFSXSMR. Further studies exploring this insert in zebrafish are needed to determine if these conserved amino acids are truly critical and that the function is either dictated through the overall presence of the loop or that the loop dependent activities require the few conserved amino acids retain in zPol θ.

zPol θ experiences Ca^{2+} dependent polymerization

Our data shows that zPol θ retains all of the DNA polymerase activities of hPol θ, with the exception being the robust Ca^{2+} dependency during DNA extension (Fig 8). While it is unusual to see DNA polymerases extend past the initial insertion event, there have been instances where high-fidelity *Sulfolobus solfataricus* Dpo4 polymerase uses calcium (32). Whether or not swapping Ca^{2+} in zebrafish makes it a faster or mutagenic polymerase has not been explored but evolutionarily it is possible that Ca^{2+} makes for a stable ion swap for structural alignment and catalytic activity. Given that in freshwater

366 contains nearly identical concentrations of the divalent ions (33), its plausible that Ca^{2+}
367 could readily be a co-factor substitute.

368 Our data presented is clear evidence that zebrafish Pol θ is a homolog to human
369 Pol θ and that structurally and enzymatically behave with similar functions. This study is
370 significant as it highlights the availability of zebrafish as a model organism for studying
371 Pol θ and its potential function in DNA repair and disease. In particular given the robust
372 assortment of tools zebrafish offers a powerful, functionally relevant model for human
373 melanoma. Future studies could introduce patient derived mutations by way of germline
374 alterations and study the effects over the life time of the animal. Thus, adding new insights
375 into potential disease markers and mechanisms of disease progression and treatment.
376

377 **MATERIALS AND METHODS**

378 *Materials*

379 All materials were purchased from Sigma-Aldrich (St. Louis, MO), Bio-rad
380 Laboratories (Hercules, CA), AmericanBio (Canton, MA), and Research Products (Mount
381 Prospect, IL). DNA oligonucleotides were purchased from Integrated DNA Technologies
382 (Newark, NJ) and deoxynucleotides from New England Biolabs (Ipswich, MA). All DNA
383 oligos were purified via HPLC with standard desalting from the manufacturer.
384

385 *Zebrafish Pol θ cloning*

386 Total RNA from 4-hour post fertilization embryos was extracted using TRIzol
387 (Invitrogen) following manufactures instructions. A library of cDNAs was generated from
388 the pool of polyA mRNAs using ProtoScript II Reverse Transcriptase (New England

Biolabs, NEB) following manufactures instructions, primed by oligo(dT). The polymerase domain of zebrafish Pol θ was then amplified for cloning into the POLQM1 vector (8) a pSUMO3 based expression vector. This was a two-step cloning process as the POLQM1 vector did not have multiple cloning sites.

First, the polymerase domain of zebrafish Pol θ (residues 1801-2579) was amplified from the cDNA library with primers contain a 5' KpnI site and a 3' BamHI site:

Pol θ RVS BamHI- TATACTGGATCCTTATATGTCCAGGTCTTGAAGGTTACC

Pol θ FWD KpnI- ATTAGGTACCTCAACATCAGTGTTAGGCGCAC

Second, the 6xHIS and SUMO sequences (HIS-SUMO) of POLQM1 were amplified off of the plasmid using primers containing a 5' XbaI site and a 3' KpnI site:

HIS-SUMO RVS KpnI- ATTAGGTACCTCCCGTCTGCTGC

HIS-SUMO FWD XbaI- TTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAG

PCRs reactions used Phusion High-Fidelity DNA Polymerase (NEB) following manufacturer's instructions and were run for 30 cycles. PCR Products were gel isolated from a 1% TAE agarose gel using Freeze 'N Squeeze DNA gel extraction columns (Bio-Rad), following manufacturer's instructions.

Next, the zPol θ, HIS-SUMO, and POLQM1 DNAs were digested with appropriate enzymes (NEB) overnight at 37ºC:

zPol θ PCR - BamHI and KpnI

HIS-SUMO PCR - KpnI and XbaI

POLQM1 - XbaI and BamHI

Digested samples were gel separated on a 1% TAE agarose gel and fragments were isolated using Freeze 'N Squeeze DNA gel extraction columns. zPol θ and HIS-SUMO

412 digested fragments were then ligated using T4 ligase (NEB) incubating at 16ºC overnight
413 and gel isolated from a 1% TAE agarose gel using Freeze 'N Squeeze DNA gel extraction
414 columns. zPol θ-HIS-SUMO fragment was ligated into linearized POLQM1 vector using
415 T4 ligase and incubating at 16ºC overnight. NEB 5-alpha competent *E.coli* (NEB) were
416 transformed by ligated productions using manufacturer's instructions. Bacteria were
417 selected for through ampicillin resistance.
418
419 *zPol θ polymerase domain modeling and alignments*
420 Amino acid sequence alignments were completed using EMBL-EBI Clustal Omega
421 MSA (34) on default settings. Structural rendering of the zPol θ polymerase domain was
422 completed using ColabFold v1.5.5: AlphaFold2 using MMseqs2 (16,35–38) on default
423 settings. The resulting zPol θ structure was compared to the solved hPol θ polymerase
424 domain structure (AX0Q) (17) using the pairwise alignment tool in FATCAT (39) on default
425 setting, and visualized using RCSB PDB visualization tools (40).
426
427 *Expression and Purification of hPol θ and zPol θ*
428 Recombinant pSUMO3 plasmids containing the truncated polymerase domain of
429 hPol θ and zPol θ gene were expressed in *E.coli* and purified as previously described in
430 Thomas et al (18). Briefly, Rosetta2(DE3) cells containing the zPol θ plasmid were
431 inoculated into autoinduction Terrific Broth and grown at 20°C for 60 hours. Cells were
432 harvested through centrifugation and lysed in Lysis buffer. After 6 rounds of sonication,
433 cell fractions were further separated via centrifugation. The fraction containing soluble
434 protein were applied to a 5mL His-Trap FF crude Nickel Column (Cytiva) by FPLC using

a high imidazole gradient. Fractions containing Pol θ were separated again on a HiTrap Heparin HP (Cytiva) column for further purification. Eluate containing Pol θ was incubated overnight at 4°C with SUMO2 Protease (Fisher Scientific) to remove the 6xHIS-SUMO tag. Untagged Pol θ was separated from the 6X-HIS-SUMO on a Hi-Trap Chelating HP column, reserving the flowthrough that contained only Pol θ. A final HiTrap Heparin column removed any remaining non-specific binding proteins and exchanged the imidazole buffer for a high NaCl buffer. Protein purification was verified on 10% denaturing SDS visualized on an Odyssey CL-x IR scanner (LiCOR). Purified protein is highly unstable and was flash-frozen in liquid nitrogen and stored at -80°C for 3 months maximum.

DNA Substrate Generation

Double stranded DNA substrates (dsDNA) were generated using complementary oligodeoxynucleotides from IDT. Templates representing cyclobutane pyrimidine Thymine-Thymine dimer (CPD) damaged and undamaged were synthesized in the Delaney laboratory (Sarah Delaney, Brown University).

The CPD-containing sequence is as follows: 5'-AAG AGT TCG AXX GCC TAC ACT GGA GTA CCG GAG-3' where XX denotes the CPD lesion. The oligonucleotide was synthesized on a MerMade 4 (BioAutomation) using standard phosphoramidite chemistry. All reagents were purchased from Glen Research. The 5'-dimethoxytrityl group was retained for HPLC purification (Agilent PLRP-S column, 250 mm × 4.6 mm; mobile phase A = 1% acetonitrile, 10% triethylammonium acetate (TEAA), 89% water; mobile phase B = 10% TEAA, 90% acetonitrile). The gradient was as follows: 95% A / 5% B to 65% A /

35% B over 35 min at 1 mL/min. Oligonucleotide was subject to detritylation by incubation for 60 min at room temperature in 20% (v/v) aqueous glacial acetic acid. The reaction was quenched by precipitation of the oligonucleotides in room-temperature ethanol. A second HPLC purification was then performed (same column and mobile phases as above) using the following gradient: 100% A / 0% B to 75% A / 25% B over 40 min at 1 mL/min. The purified oligonucleotide was flash-frozen with liquid nitrogen and lyophilized.

The 5'6-FAM primers were annealed to complementary DNA templates with sequence context as previously described (26,41) and are described below:

25/40 undamaged DNA substrate

```
5′-/FAM/ TTT GCCT TGA CCA TGT AAC AGA GAG
             CGGA ACT GGT ACA TTG TCT CTC GCA CTC ACT CTC TTC TCT
```

24/33 CPD damaged DNA substrate

```
5′-/FAM/-TTT CTC CGG TAC TCC AGT GTA GGC
             GAG GCC ATG AGG TCA CAT CCG TTA GCT TGA GAA
```

24/33 undamaged DNA substrate

```
5′-/FAM/-TTT CTC CGG TAC TCC AGT GTA GGC
             GAG GCC ATG AGG TCA CAT CCG TTA GCT TGA GAA
```

Confirmation of annealed substrates was determined 12% Native PAGE and samples scanned on an RB Amersham Typhoon Fluorescent Imager (Cytiva) with a FAM filter. Single oligodeoxynucleotides were purchased from IDT for MMEJ with internal consensus sequence as previously described(23).

```
5′-/FAM/ GGT TAG CCC GGG
```

Circular Dichroism and Melting Temperature

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

Electrophoretic Mobility Shift Assay (EMSA)

The DNA binding affinity constant $K_{D(DNA)}$ was determined as previously described(18). zPol θ was titrated from 0-1000nM against 10nM 25/40 dsDNA substrate in binding buffer and incubated for 1 hour at room temperature. Samples were separated on a 6% Native PAGE and scanned on an RB Typhoon scanner (Cytiva) with the FAM fluorescence filter. Separated bound and unbound products were quantified using ImageQuant. $K_{D(DNA)}$ was determined by equation 1.

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

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

Rapid Chemical Quench Assay

Biphasic burst kinetics were measured as previously described(18). Briefly,100 nM Pol θ was pre-mixed with 300 nM 25/40 dsDNA substrate and rapidly mixed with 100 μM of dCTP (correct nucleotide) with 10 mM $MgCl_2$ using an RQF-3 Rapid Chemical Quench

instrument (KinTek Corporation) at 37°C between 0.004-0.6 seconds. Reactions were quenched by addition of 0.5M EDTA. Products were separated on a 15% Urea-denaturing polyacrylamide gel and scanned using an Amersham Typhoon RB Fluorescent imager (Cytiva). Extended product (n+1) was quantified using ImageQuant software and then plotted to a full biphasic pre-steady state burst equation via non-linear regression using Prism 9 GraphPad software (equation 2). A minimum of three replicates were included for each assay on two independent protein preparations.

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

Primer Extension Assays

Qualitative primer extension assays were performed as previously described(18). Varying conditions were used to explore the primer extension capabilities between the hPol θ and zPol θ. Under single-turnover conditions, excess Pol θ (200 nM) was pre-incubated with 50 nM DNA and incubated for 5 minutes at 37˚C. Nucleotide (125 μM) of either none, all, individual nucleotides were preincubated with 20 mM *x* inorganic salt (x = $MgCl_2$, $CaCl_2$, $MnCl_2$). Under Michaelis-Menton conditions, DNA (200nM) was in excess to Pol θ (50 nM). All reactions were carried out in buffer containing 20 mM Tris HCl, pH 8.0, 25 mM KCl, 4% glycerol, 1 mM βME, and 80 μg/mL BSA. The reaction was initiated by combining Pol-θ/DNA with dNTP/salt. Reactions were incubated at 37˚C for 5 minutes before being stopped by 80% Formamide/EDTA quench. Products were separated out on a 15% urea-denaturing polyacrylamide gel and scanned on an RB Amersham Typhoon fluorescent imager (Cytiva).

MMEJ Assay

Microhomology mediated end-joining assay for both hPol θ and zPol θ were carried out as previously described(23) on a 12-mer FAM labeled oligodeoxynucleotide. Pol θ (20 nM) was preincubated with 30 nM 5'-FAM ssDNA in reaction buffer (25mM Tris-HCl pH8.8, 1mM βME, 0.01% NP-40, 0.1 mg/mL BSA, 10% glycerol, 10 mM $MgCl_2$, 30mM NaCl) for 5 minutes at 37°C. Nucleotides (20 μM) were added and incubated at 37°C for an additional 45 minutes. Reactions were terminated by addition of non-denaturing stop buffer (100 mM Tris-HCl pH 7.5, 10 mg/mL proteinase K, 80 mM EDTA, and 0.5% SDS) for an additional 15 minutes. DNA products were separated on a 12% native polyacrylamide gel and scanned by an RB fluorescent Amersham Typhoon (Cytiva) with a FAM filter.

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. L.Rothwell received support from the Bryant Center of Health and Behavioral Science.

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

554 ## WORK CITED

555 1. Farshidfar F, Rhrissorrakrai K, Levovitz C, Peng C, Knight J, Bacchiocchi A, et al.
556 Integrative molecular and clinical profiling of acral melanoma links focal amplification
557 of 22q11.21 to metastasis. Nat Commun. 2022 Feb 23;13(1):898.

558 2. Thyme SB, Schier AF. Polq-mediated end joining is essential for surviving DNA double-
559 strand breaks during early zebrafish development. Cell Rep. 2016 Apr 26;15(4):707–14.

560 3. Brambati A, Barry RM, Sfeir A. DNA polymerase theta (Polθ) – an error-prone
561 polymerase necessary for genome stability. Current Opinion in Genetics &
562 Development. 2020 Feb 1;60:119–26.

563 4. Yoon JH, McArthur MJ, Park J, Basu D, Wakamiya M, Prakash L, et al. Error-Prone
564 Replication through UV Lesions by DNA Polymerase θ Protects against Skin Cancers.
565 Cell. 2019 Mar 7;176(6):1295-1309.e15.

566 5. Chan SH, Yu AM, McVey M. Dual Roles for DNA Polymerase Theta in Alternative End-
567 Joining Repair of Double-Strand Breaks in Drosophila. PLOS Genetics. 2010 Jul
568 1;6(7):e1001005.

569 6. Kent T, Chandramouly G, McDevitt SM, Ozdemir AY, Pomerantz RT. Mechanism of
570 microhomology-mediated end-joining promoted by human DNA polymerase θ. Nat
571 Struct Mol Biol. 2015 Mar;22(3):230–7.

572 7. Yousefzadeh MJ, Wyatt DW, Takata K ichi, Mu Y, Hensley SC, Tomida J, et al. Mechanism
573 of Suppression of Chromosomal Instability by DNA Polymerase POLQ. PLOS Genetics.
574 2014 Oct 2;10(10):e1004654.

575 8. Hogg M, Seki M, Wood RD, Doublié S, Wallace SS. Lesion Bypass Activity of DNA
576 Polymerase θ (POLQ) Is an Intrinsic Property of the Pol Domain and Depends on Unique
577 Sequence Inserts. Journal of Molecular Biology. 2011 Jan 21;405(3):642–52.

578 9. Yu AM, McVey M. Synthesis-dependent microhomology-mediated end joining accounts
579 for multiple types of repair junctions. Nucleic Acids Res. 2010 Sep;38(17):5706–17.

580 10. Ceccaldi R, Liu JC, Amunugama R, Hajdu I, Primack B, Petalcorin MIR, et al.
581 Homologous-recombination-deficient tumours are dependent on Polθ-mediated
582 repair. Nature. 2015 Feb;518(7538):258–62.

583 11. Goff JP, Shields DS, Seki M, Choi S, Epperly MW, Dixon T, et al. Lack of DNA Polymerase
584 θ (POLQ) Radiosensitizes Bone Marrow Stromal Cells In Vitro and Increases
585 Reticulocyte Micronuclei after Total-Body Irradiation. Radiat Res. 2009 Aug;172(2):165–
586 74.

12. Goullet de Rugy T, Bashkurov M, Datti A, Betous R, Guitton-Sert L, Cazaux C, et al. Excess Polθ functions in response to replicative stress in homologous recombination-proficient cancer cells. Biology Open. 2016 Sep 9;5(10):1485–92.

13. Allera-Moreau C, Rouquette I, Lepage B, Oumouhou N, Walschaerts M, Leconte E, et al. DNA replication stress response involving PLK1, CDC6, POLQ, RAD51 and CLASPIN upregulation prognoses the outcome of early/mid-stage non-small cell lung cancer patients. Oncogenesis. 2012 Oct;1(10):e30–e30.

14. Lemée F, Bergoglio V, Fernandez-Vidal A, Machado-Silva A, Pillaire MJ, Bieth A, et al. DNA polymerase θ up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proceedings of the National Academy of Sciences. 2010 Jul 27;107(30):13390–5.

15. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013 Apr 25;496(7446):498–503.

16. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022 Jun;19(6):679–82.

17. Zahn KE, Averill AM, Aller P, Wood RD, Doublié S. Human DNA polymerase θ grasps the primer terminus to mediate DNA repair. Nat Struct Mol Biol. 2015 Apr;22(4):304–11.

18. Thomas C, Avalos-Irving L, Victorino J, Green S, Andrews M, Rodrigues N, et al. Melanoma-Derived DNA Polymerase Theta Variants Exhibit Altered DNA Polymerase Activity. Biochemistry. 2024 May 7;63(9):1107–17.

19. Steitz TA. DNA Polymerases: Structural Diversity and Common Mechanisms *. Journal of Biological Chemistry. 1999 Jun 18;274(25):17395–8.

20. Johnson KA. Transient-State Kinetic Analysis of Enzyme Reaction Pathways. Enzymes. 1992;20(C):1–61.

21. Kent T, Mateos-Gomez PA, Sfeir A, Pomerantz RT. Polymerase θ is a robust terminal transferase that oscillates between three different mechanisms during end-joining. eLife. 5:e13740.

22. Joyce CM. Techniques used to study the DNA polymerase reaction pathway. Biochimica et biophysica acta. 2010 May;1804(5):1032–40.

23. Black SJ, Ozdemir AY, Kashkina E, Kent T, Rusanov T, Ristic D, et al. Molecular basis of microhomology-mediated end-joining by purified full-length Polθ. Nature Communications. 2019;10(1).

24. Brambati A, Barry RM, Sfeir A. DNA polymerase theta (Polθ) – an error-prone polymerase necessary for genome stability. Current Opinion in Genetics and Development. 2020;60:119–26.

25. Li C, Zhu H, Jin S, Maksoud LM, Jain N, Sun J, et al. Structural basis of DNA polymerase θ mediated DNA end joining. Nucleic Acids Research. 2023 Jan 11;51(1):463–74.

26. Yoon JH, McArthur MJ, Park J, Basu D, Wakamiya M, Prakash L, et al. Error-Prone Replication through UV Lesions by DNA Polymerase θ Protects against Skin Cancers. Cell. 2019;176(6):1295-1309.e15.

27. Freudenthal BD, Beard WA, Shock DD, Wilson SH. Observing a DNA Polymerase Choose Right from Wrong. Cell. 2013;154(1):157–68.

28. Nakamura T, Zhao Y, Yamagata Y, Hua YJ, Yang W. Watching DNA polymerase eta make a phosphodiester bond. Nature. 2012;487(7406):196–201.

29. Towle-Weicksel JB, Dalal S, Sohl CD, Doublie S, Anderson KS, Sweasy JB, et al. Fluorescence resonance energy transfer studies of DNA polymerase β the critical role of fingers domain movements and a novel non-covalent step During nucleotide selection. Journal of Biological Chemistry. 2014;289(23):16541–50.

30. Werneburg BG, Ahn J, Zhong X, Hondal RJ, Kraynov VS, Tsai MD. DNA Polymerase β: Pre-Steady-State Kinetic Analysis and Roles of Arginine-283 in Catalysis and Fidelity†. Biochemistry. 1996 Mar 20;35(22):7041–50.

31. Murphy DL, Kosa J, Jaeger J, Sweasy JB. The Asp285 Variant of DNA Polymerase Beta Extends Mispaired Primer Termini via Increased Nucleotide Binding. Biochemistry. 2008 Aug 6;47(31):8048–57.

32. Ralec C, Henry E, Lemor M, Killelea T, Henneke G. Calcium-driven DNA synthesis by a high-fidelity DNA polymerase. Nucleic Acids Res. 2017 Dec 1;45(21):12425–40.

33. Guerin AJ, Weise AM, Chu JWF, Wilcox MA, Greene ES, Therriault TW. High-resolution freshwater dissolved calcium and pH data layers for Canada and the United States. Sci Data. 2024 Apr 11;11(1):370.

34. Madeira F, Madhusoodanan N, Lee J, Eusebi A, Niewielska A, Tivey ARN, et al. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. Nucleic Acids Res. 2024 Jul 1;52(W1):W521–5.

35. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;

36. Mirdita M, Steinegger M, S"oding J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics. 2019;35(16):2856–8.

37. Mirdita M, von den Driesch L, Galiez C, Martin MJ, S"oding J, Steinegger M. Uniclust databases of clustered and deeply annotated protein sequences and alignments. Nucleic Acids Res. 2017;45(D1):D170–6.

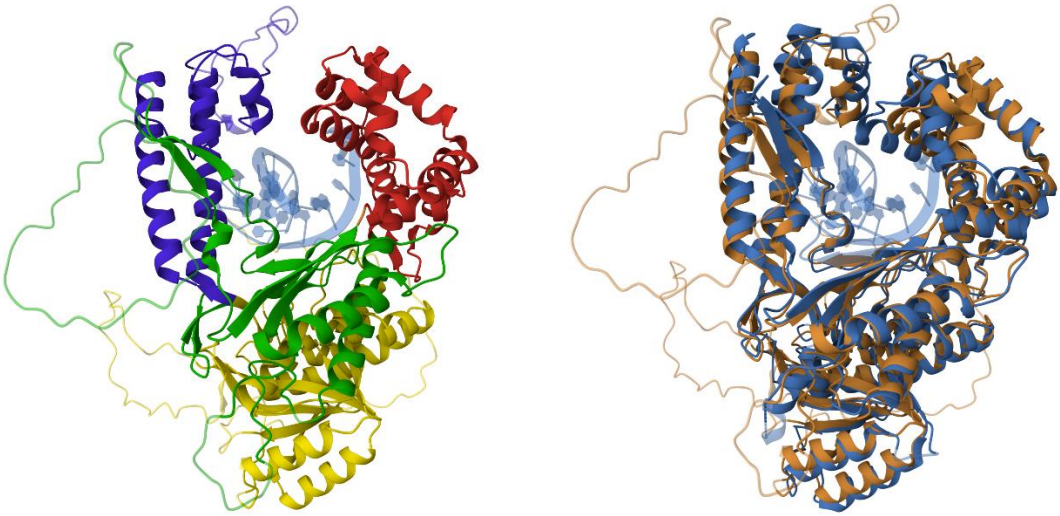
38. Mitchell AL, Almeida A, Beracochea M, Boland M, Burgin J, Cochrane G, et al. MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res. 2019;

39. 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. 2020 Jul 2;48(W1):W60–4.

40. Bittrich S, Segura J, Duarte JM, Burley SK, Rose Y. RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments. Bioinformatics. 2024 Jun 3;40(6):btae370.

41. Prasad R, Longley MJ, Sharief FS, Hou EW, Copeland WC, Wilson SH. Human DNA polymerase θ possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. Nucleic Acids Research. 2009;37(6):1868–77.

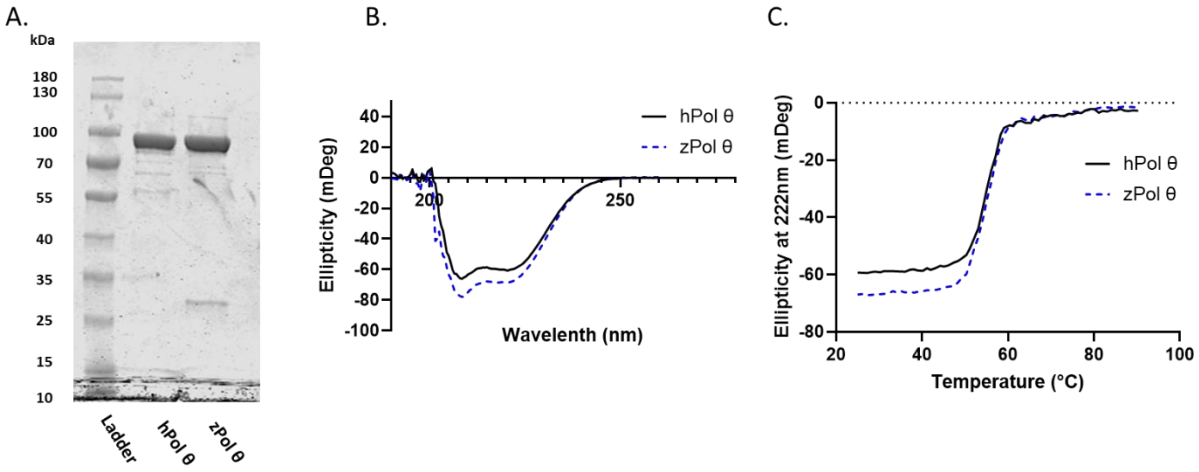
670 Fig 1



671

672

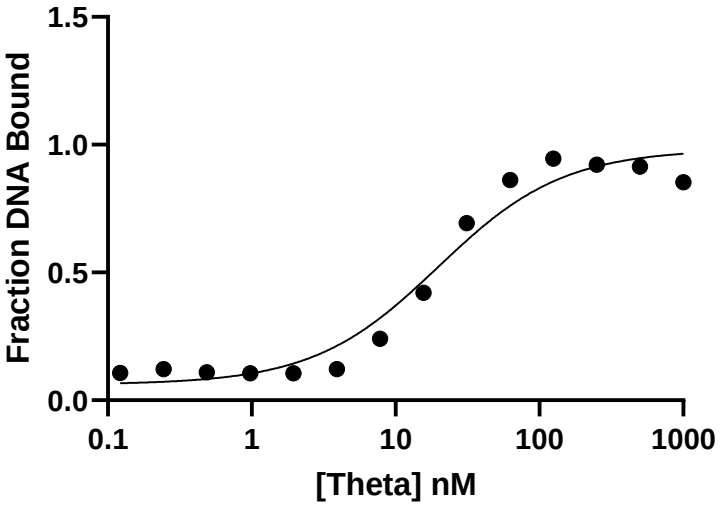
673 Fig 2



674

675

676 Fig 3

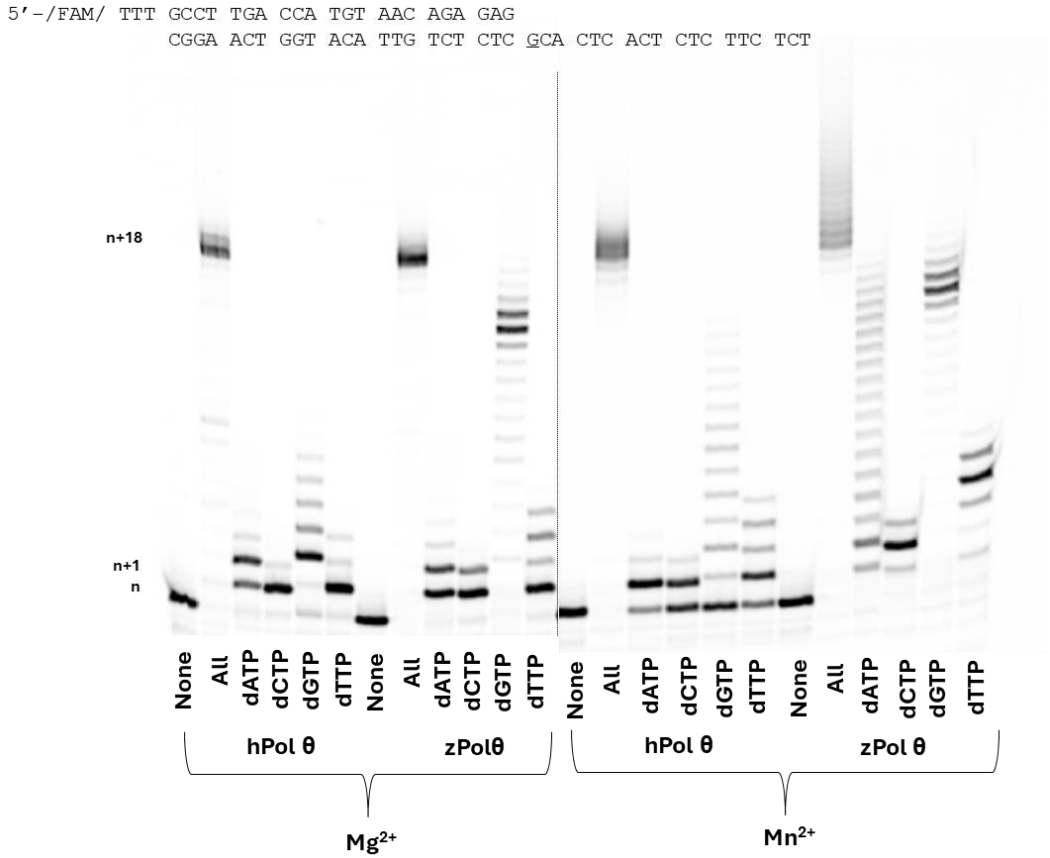


677

678

679

680 Fig 4

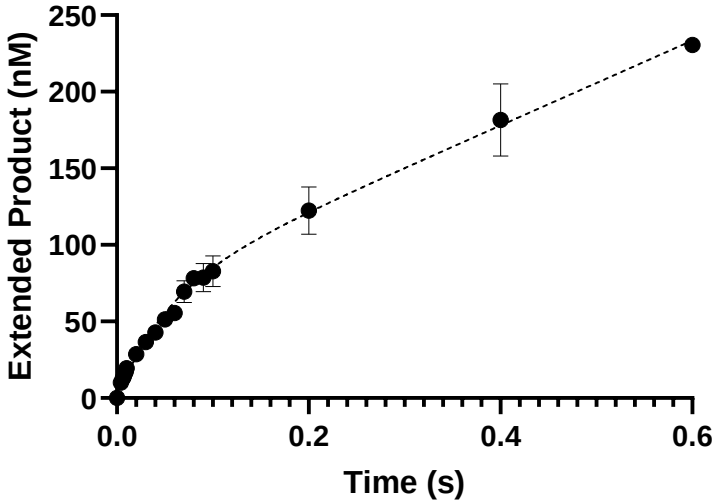


681

682

683 Fig 5

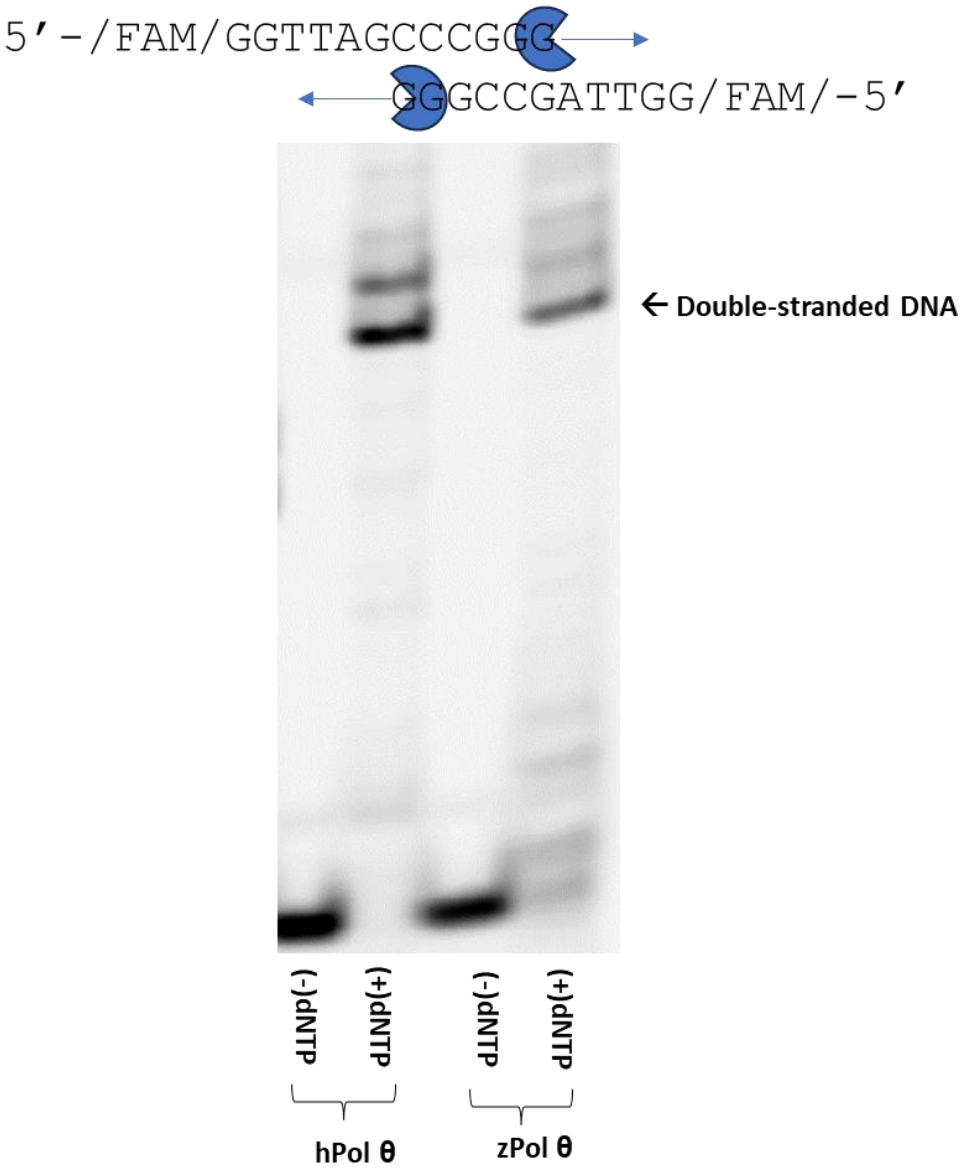
684



685

686

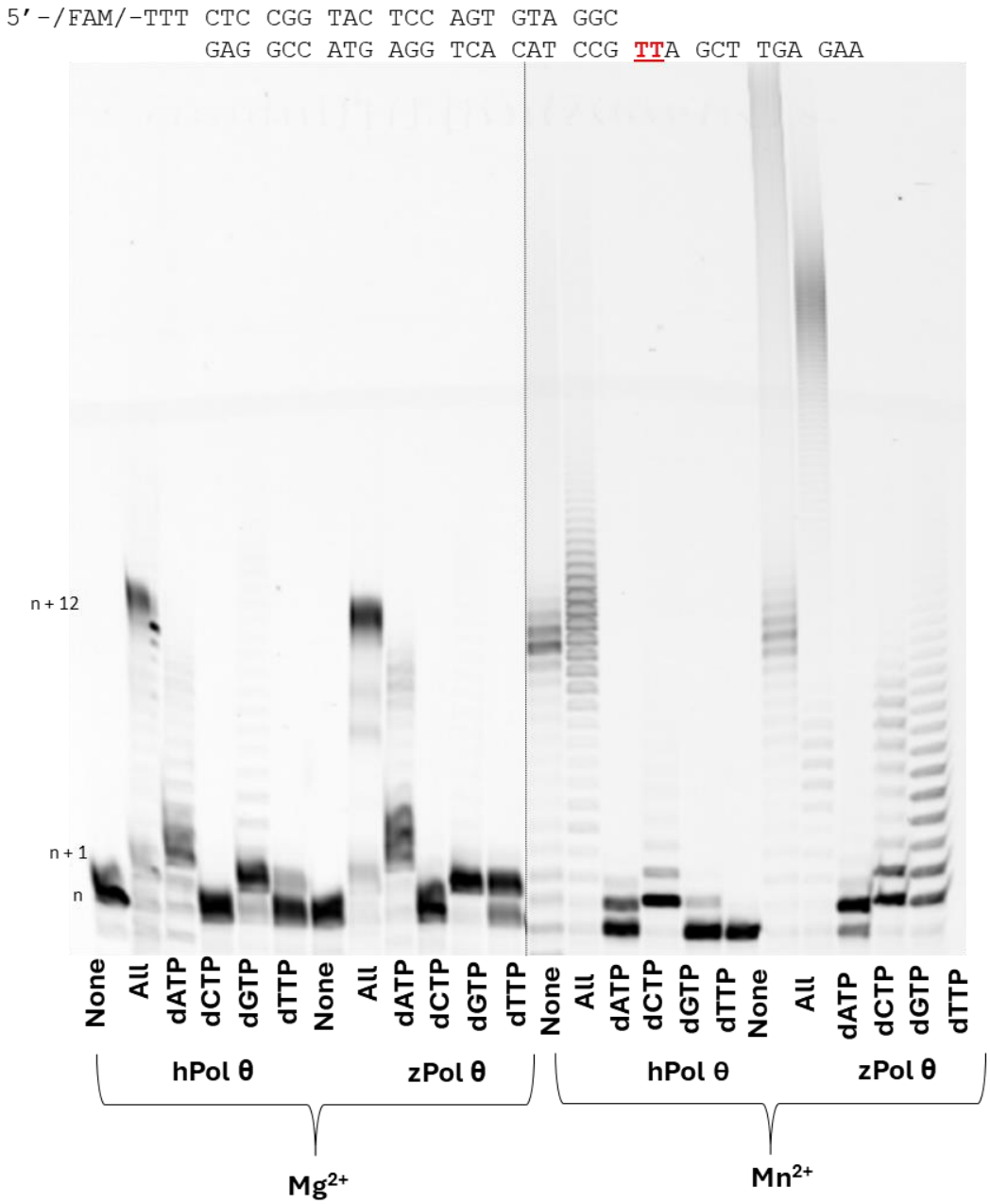
687 Fig 6



688

689

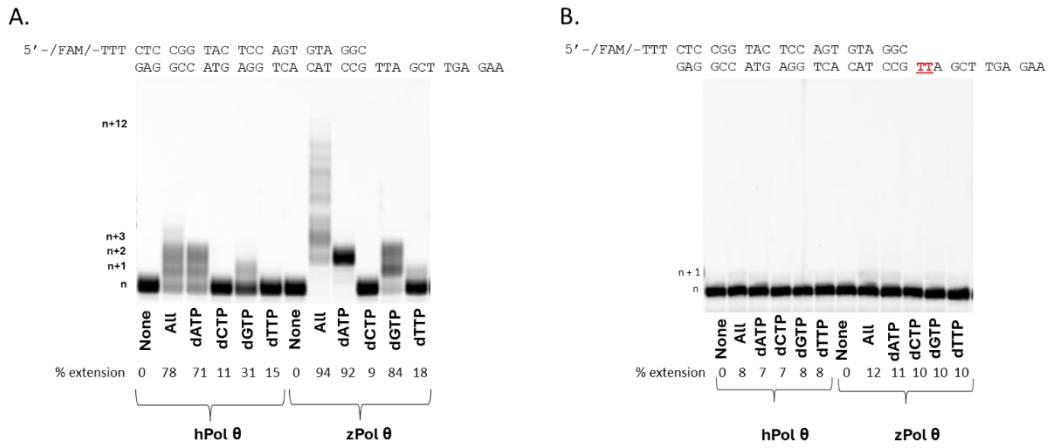
690 Fig 7



691

692

693 Fig 8



694

695

696 SFig 1

697

698

699

SFig2