2013

Fidelity of Plasmodium falciparum apicoplast DNA polymerase

Eric Evan Parrott
Iowa State University

Fidelity of *Plasmodium falciparum* apicoplast DNA polymerase

by

Eric Evan Parrott

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Biochemistry

Program of Study Committee:
Scott W. Nelson, Major Professor
Richard Honzatko
Steven M. Lonergan

Iowa State University

Ames, Iowa

2013

TABLE OF CONTENTS

ABSTRACT

Members of the phylum *Apicomplexa*, including pathogens such as *Plasmodium falciparum*, *Toxoplasma gondii*, and *Babesia bovis* contain a non-photosynthetic plastid called an apicoplast. It has been implicated in the biosynthesis of fatty acids, iron-sulfur clusters, heme groups, and isoprenoid synthesis for the pathogen. This unique plastid was immediately recognized as a potential target for drug discovery as it is essential for the survival and reproduction of the parasite. With this in mind, our group has initiated characterization of the DNA replication system within the apicoplast. Here we report our efforts to determine the fidelity of the DNA polymerase (apPOL) using pre-steady state kinetics. The apPOL exhibits error rates typical of a high fidelity family A polymerase in the 10^4 to the 10^6 range and has an average processivity of approximately 2 nucleotides incorporated per binding event. The data indicates that the rate of chemistry or a conformational change preceding the chemical step is the rate limiting step for sequential polymerization as inclusion of a translocation step was not necessary to explain the product time courses. Additionally, with some mismatches, stalling of the primer extension reactions was observed, indicating that like the misincorporation reaction, the degree of mismatch extension is highly variable. This stalling may indicate that severe distortions of the active site following misincorporation significantly slows mismatch extension, which could increase the excision of the mismatch by the 3'-5' exonuclease activity of the wild-type apPOL and increase overall fidelity. Finally, it was observed that the apPOL will catalyze non-template nucleotide additions to the primer strand with a preference for dGTP, followed by dATP, dCTP, and dTTP.

CHAPTER I: INTRODUCTION

General Introduction

Despite the advances of modern medicine in treating and preventing infectious disease, many diseases including malaria remain a serious threat, particularly to those living in the "third world." The World Health Organization estimates that in 2010 approximately 219 million people contracted malaria with approximately 660 thousand cases resulting in death (World Health Organization, 2013). By far the most vulnerable people are under 5 years old, pregnant, or suffering from human immunodeficiency virus (HIV). While there are currently effective anti-malarial drug treatments, e.g. Artemisinin-combination therapies (Dondorp et al., 2009), the history of malaria drugs indicates that resistance develops in as little as ten years after the drug becomes widely used. For example, resistance for the drugs cholorquine, sulfadoxine-pyrimethamine, mefloquine, and artemisinin have all developed with the last being the current drug of choice (Shetty, 2012). Artemisinin resistance was documented nine years after is wide spread use began (Dondorp et al., 2009). For this reason current drugs are often given in combination to reduce the likelihood of resistance to a single drug occurring (Shetty, 2012). While many other approaches are being taken to eliminate malaria, including nets, insecticides, vaccines, etc., new drugs must be developed to combat current and future resistant strains (World Health Organization, 2013) (Miller et al., 2013).

As a result of this critical need for novel treatments, many aspects of malarial disease are being explored with a particular interest in the causative agents of malaria, protozoa of the genus *Plasmodium*. Malaria and diseases like it are spread through infected mosquitoes. Five species of *Plasmodium* are responsible for human malaria including *Plasmodium malariae*, *Plasmodium*

vivax, *Plasmodium ovale*, *Plasmodium knowlesi*, and *Plasmodium falciparum*. *P. falciparum* in the most widespread and dangerous of the malaria causing species and is the source for much of the drug resistance as well as 91% of the reported malaria cases (Kalanon and McFadden, 2010). The pathogen's life cycle is dependent on the alternation between mosquito and host. *Plasmodium* is part of a larger group of protozoa called the Apicomplexans. Apicomplexans are obligate intracellular parasites responsible for several important animal and human disease worldwide including Toxoplasma gondii (cats and humans), Cryptosporidium (humans), Theileria (cattle), and Bebesia (cattle) (Kalanon and McFadden, 2010).

Particular to the Apicomplexans is the presence of a small 3 or 4 membraned organelle called the apicoplast. Current evolutionary thought posits that the apicoplast is derived from a secondary endosymbiosis of a photosynthetic red algae by a eukaryotic protozoan, followed by the loss of photosynthetic elements and transfer of apicoplast genes to the nucleus(Lim and McFadden, 2010). Recent research has revealed the apicoplast is essential and involved in the biosynthesis of fatty acids, iron-sulfur clusters, heme groups, and isoprenoid units. Yeh and DeRisi recently showed that the *Plasmodium* parasite will not survive without the apicoplast but can be rescued by supplementation of the isoprenoid pathway products. Interestingly, the parasite will survive until upon division its progeny fail to remain viable displaying a now characteristic "delayed death" syndrome. With its biologically unique and essential role in the pathogen, the apicoplast has been identified as an attractive potential drug target for the treatment of malaria (Yeh and DeRisi, 2011).

Of interest to our group is the maintenance and replication of the apicoplast genome. The small 35 kb apicoplast genome is circular in nature and is thought to be replicated by two different methods. For the first, replication is initiated at a large inverted repeat and single

stranded unidirectional replication begins with the formation of the twin D-loops. Secondly, replication has been shown to proceed in a rolling circle mechanism initiated in as yet unknown location. Similar processes are used by chloroplasts in plants and algae (Williamson et al., 2002). The number and function of proteins involved in the replisome varies from organism to organism. One of the simplest systems, T7 phage, uses 4 core enzymes including a polymerase (polymerization of deoxyribonucleotide monomers), helicase (unwinding and separation of duplex DNA into single strands), primase (synthesis of short RNA primers for initiation of lagging strand DNA synthesis), and a single-stranded binding protein (prevents premature annealing and protection of single stranded DNA from other enzymes). In addition, the T7 phage systems use the host thioredoxin as a processivity factor to greatly increase the processivity of the polymerase (Benkovic et al., 2001).

The apicoplast encodes 68 genes including rRNAs, tRNAs and other housekeeping proteins, with the rest of the approximately 550 apicoplast proteins being encoded within the nuclear genome. Nuclear encoded proteins are targeted to the apicoplast through N-terminal signal peptide followed by a transit peptide which is thought to be involved in targeting the gene product to the apicoplast and through the membranes (Waller et al., 1998). Many putative apicoplast-targeted replisome enzymes have been identified including the polyprotein Pfprex (containing predicted polymerase, helicase, and primase activities), a topoisomerase II enzyme (necessary for relieving tension in the DNA coil for circular genomes), and a single-stranded binding protein. No processivity factor for the apicoplast DNA polymerase has been discovered as yet (Lindner et al., 2011; Prusty et al., 2010; Seow et al., 2005; Weissig et al., 1997).

Previous research has shown that inhibition of the apicoplast-targeted topoisomerase II with ciprofloxacin showed anti-malarial activity (Weissig et al., 1997). In recent years, work by

Seow, et al. identified and began characterization of the apicoplast-targeted polyprotein: Pfprex. Evidence suggests that the polyprotein is translated in the cytoplasm of the parasite and then shuttled to the apicoplast where the apicoplast DNA polymerase (apPOL) is cleaved from the DNA helicase and DNA primase (Seow et al., 2005). Given its central role in DNA replication, the apPOL represents an attractive target for anti-malarial drug development. In addition, other polymerases are currently being successfully targeted for treatment of diseases such as HIV, hepatitis B, and herpes simplex virus. Research has shown that several different strategies are possible for inhibition of polymerase activity including DNA damaging agents, modulation of DNA enzymes (such as topoisomerase II), inhibiting substrate metabolism, and finally substrate analogs or small molecules (Berdis, 2008). Of interest to us is the last strategy and an understanding of the enzymatic activity of apPOL is necessary to inform the development of inhibitors. In addition, the apPOL is part of the PolA family polymerases, and particular to a group of PolA polymerases which are found only in the bacterial phylum Aquificae, some thermophilic viruses, and the Apicomplexans (Schoenfeld et al., 2013). The nearest homolog outside the Apicomplexans to the apPOL is in the from the cyanobacteria *Cyanothece* sp. PCC 8802 with 35% identity. The closest human polymerases are lesion bypass polymerases theta and nu (23 and 22% respectively). Within the malaria causing *parasites* the *P. vivax* apicoplast targeted polymerases has 84% identity, suggesting that P. falciparum apPOL drugs may be affective against other malaria pathogens (Wingert et al., 2013).

Our focus with this investigation to characterize the nucleotide incorporation kinetics that govern replication by the apPOL. Specifically, our aim is to use pre-steady state kinetics to elucidate the fidelity of the polymerase in regards to replication-induced mutations (i.e., how faithfully the polymerase inserts the correct nucleotide opposite the template base and the effects

of the incorrect base being inserted). Although some work has been done in this area already, in light of our bioinformatic analysis of the Pfprex gene product, and subsequent addition of 38 amino acids to the N-terminus of the apPOL from the previously published construct, further investigation is warranted. Our group has shown a 750-fold increase in steady state activity over previously published results and fidelity measurements that are more in line with a replicative DNA polymerase, which was in some doubt in previous publications (Kennedy et al., 2011; Wingert et al., 2013). The lack of a complete understanding of the kinetic mechanism and fidelity, combined with its importance in human health makes this an ideal system to study. The desire to gain and deeper understanding of the biochemical characteristics of the apicoplast targeted polymerase serves as the basis for this thesis project.

Thesis Outline

For my thesis project I examined the apicoplast-targeted polymerase of *Plasmodium falciparum* (apPOL), which is thought to replicate the apicoplast genome. I performed initial characterization assays and pre-steady state experiments to elucidate the kinetics that influence fidelity of the polymerase. The properties of this polymerase have proven to be quite interesting with some surprising findings.

This thesis is setup to highlight the progress of my research project and contains a manuscript in preparation that will be submitted for publication to the journal, Biochemistry. Chapter II contains the manuscript which discusses the initial biochemical characterization of apPOL, burst kinetics, pre-steady kinetics analysis and fidelity measurements, multiple turnover mechanisms, and analysis of the mismatch stalling. Chapter III will be a general conclusion

discussing the accomplishments of my research project, future directions, and

acknowledgements.

References

Berdis, A.J. (2008). DNA polymerases as therapeutic targets. Biochemistry (Mosc.) *47*, 8253–8260.

Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyo, A.P., Tarning, J., Lwin, K.M., Ariey, F., Hanpithakpong, W., and Lee, S.J. (2009). Artemisinin resistance in Plasmodium falciparum malaria. N. Engl. J. Med. *361*, 455–467.

Kalanon, M., and McFadden, G.I. (2010). Malaria, Plasmodium falciparum and its apicoplast. Biochem. Soc. Trans. *38*, 775–782.

Kennedy, S.R., Chen, C.-Y., Schmitt, M.W., Bower, C.N., and Loeb, L.A. (2011). The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite Plasmodium falciparum. J. Mol. Biol. *410*, 27–38.

Lim, L., and McFadden, G.I. (2010). The evolution, metabolism and functions of the apicoplast. Philos. Trans. R. Soc. B Biol. Sci. *365*, 749–763.

Lindner, S.E., Llinás, M., Keck, J.L., and Kappe, S.H.I. (2011). The primase domain of PfPrex is a proteolytically matured, essential enzyme of the apicoplast. Mol. Biochem. Parasitol. *180*, 69–75.

Miller, L.H., Ackerman, H.C., Su, X., and Wellems, T.E. (2013). Malaria biology and disease pathogenesis: insights for new treatments. Nat. Med. *19*, 156–167.

Prusty, D., Dar, A., Priya, R., Sharma, A., Dana, S., Choudhury, N.R., Rao, N.S., and Dhar, S.K. (2010). Single-stranded DNA binding protein from human malarial parasite Plasmodium falciparum is encoded in the nucleus and targeted to the apicoplast. Nucleic Acids Res. *38*, 7037–7053.

Schoenfeld, T.W., Murugapiran, S.K., Dodsworth, J.A., Floyd, S., Lodes, M., Mead, D.A., and Hedlund, B.P. (2013). Lateral Gene Transfer of Family A DNA Polymerases between Thermophilic Viruses, Aquificae, and Apicomplexa. Mol. Biol. Evol. *30*, 1653–1664.

Seow, F., Sato, S., Janssen, C.S., Riehle, M.O., Mukhopadhyay, A., Phillips, R.S., Wilson, R.J.M. (Iain), and Barrett, M.P. (2005). The plastidic DNA replication enzyme complex of Plasmodium falciparum. Mol. Biochem. Parasitol. *141*, 145–153.

Shetty, P. (2012). The numbers game. Nature *484*, S14–S15.

Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., and McFadden, G.I. (1998). Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. Proc. Natl. Acad. Sci. *95*, 12352–12357.

Weissig, V., Vetro-Widenhouse, T.S., and Rowe, T.C. (1997). Topoisomerase II inhibitors induce cleavage of nuclear and 35-kb plastid DNAs in the malarial parasite Plasmodium falciparum. Dna Cell Biol. *16*, 1483–1492.

Williamson, D.H., Preiser, P.R., Moore, P.W., McCready, S., Strath, M., and Wilson, R.J.M. (2002). The plastid DNA of the malaria parasite Plasmodium falciparum is replicated by two mechanisms. Mol. Microbiol. *45*, 533–542.

Wingert, B., Parrott, E., and Nelson, S.W. (2013). Misincorporation, mismatch extension, and proofreading activity of the Plasmodium falciparum apicoplast DNA polymerase. Biochemistry (Mosc.).

World Health Organization (2013). World malaria report 2012. ([S.l.]: World Health Organization).

Yeh, E., and DeRisi, J.L. (2011). Chemical Rescue of Malaria Parasites Lacking an Apicoplast Defines Organelle Function in Blood-Stage Plasmodium falciparum. Plos Biol. *9*, e1001138.

CHAPTER II: PRE-STEADY STATE KINETICS AND FIDELITY OF PLASMODIUM FALCIPARUM APICOPLAST DNA POLYMERASE

Eric E. Parrott[1], Bentley Wingert[2], and Scott W. Nelson[3]

[1]BBMB BS/MS Program at Iowa State University; [2]BBMB Undergraduate Biochemistry Program at Iowa State University; [3]BBMB Department at Iowa State University and Corresponding Author

A paper to be submitted to the journal *Biochemistry*

Abstract

Plasmodium falciparum, the most virulent of several species that cause malaria, is a part of the phylum *Apicomplexa* whose members contain a non-photosynthetic plastid called an apicoplast. This unique plastid was immediately recognized as a potential target for drug discovery as it is essential for the survival and reproduction of the parasite. With this in mind, our group has initiated characterization of the DNA replication system within the apicoplast. Here we report our efforts to determine the fidelity of the DNA polymerase (apPOL) using pre-steady state kinetics. We find that there is a 500 to 8000-fold decrease in k_{pol} for incorrect incorporation compared to correct. We similarly find that $K_{d,app}$ values are inflated, but not to the same degree, with increases ranging from 1.1 to 46-fold for incorrect versus correct incorporation. The apPOL exhibits error rates typical of a high fidelity family A polymerase in the 10^4 to the 10^6 range. Additionally, with some mismatches we observed stalling of the primer extension reactions, indicating that like the misincorporation reaction, the degree of mismatch extension is highly variable. Finally, it was observed that the apPOL will catalyze non-template nucleotide additions to the primer strand with a preference for dGTP, followed by dATP, dCTP, and dTTP.

Introduction

Plasmodium falciparum, responsible for causing 91% of the reported 219 million malaria cases in 2010, contains a non-photosynthetic plastid called an apicoplast (World Health Organization, 2013). Organisms containing an apicoplast are part of the phylum *Apicomplexa*. Many of its members are obligate intracellular parasites responsible for several animal and human diseases worldwide including *Toxoplasma gondii* and *Cryptosporidium* which cause human disease and *Theileria*, and *Babesia*, which cause economically important diseases in poultry and cattle (Kalanon and McFadden, 2010). Research has revealed the apicoplast is essential for pathogen growth and survival. It has been recognized to be involved in the biosynthesis of fatty acids, iron-sulfur clusters, heme groups, and isoprenoid synthesis. The pathogen will not survive following drug induced knockout of the apicoplast but can be rescued by supplementation of the isoprenoid pathway products. For that reason the apicoplast has been considered a potential Achilles heel for malaria (Yeh and DeRisi, 2011).

The small 35 kb apicoplast genome is circular in nature and thought to be replicated through D-loop unidirectional replication and rolling circle replication similar to chloroplasts (Williamson et al., 2002). Essential to the maintenance and replication of the apicoplast, the enzymes of the apicoplast replisome are possible drug targets (Seow et al., 2005). 68 genes including rRNAs, tRNAs and other housekeeping proteins are encoded with in the apicoplast, with the rest of the approximately 595 other apicoplast required proteins being encoded within the nuclear genome (Waller et al., 1998). Many putative apicoplast targeted replication enzymes have been found including a polyprotein Pfprex (containing polymerase, helicase, and primase activities), a topoisomerase II enzyme (necessary for relieving tension in the DNA coil for circular genomes), and a single-stranded binding protein (Lindner et al., 2011; Prusty et al.,

2010; Seow et al., 2005; Weissig et al., 1997). Anti-malarial activity has been shown through inhibition of topoisomerase II with ciprofloxacin, validating the replisome has as a potential drug target (Weissig et al., 1997).

The apicoplast targeted polymerase (apPOL) is thought to be cleaved upon transport from the cytoplasm to the apicoplast from the DNA helicase and primase domains of Pfprex (Seow et al., 2005). The apPOL is part of the PolA family polymerases, and particular to a group of PolA polymerases which are found only in the bacterial phylum Aquificae, some thermophilic viruses, and the Apicomplexans (Schoenfeld et al., 2013). The nearest homolog outside the Apicomplexans to the apPOL is in the from the cyanobacteria *Cyanothece* sp. PCC 8802 with 35% identity. The closest human polymerases are lesion bypass polymerases theta and nu (23 and 22% respectively). Within the malaria causing *Plasmodium spp.*, the *Plasmodium vivax* apicoplast targeted polymerases has 84% identity suggesting that *P. falciparum* apPOL drugs may be affective against other malarial pathogens (Wingert et al., 2013). Given its central role in DNA replication, the apPOL represents an attractive target for anti-malarial drug development. In addition, other polymerases are currently being successfully targeted for treatment of diseases such as HIV, hepatitis B, and herpes simplex virus (Berdis, 2008).

Faithful and rapid reproduction of DNA during DNA replication or repair has long been recognized as the key component of the maintenance of the information encoded by the genome. The frequency of mutations is an important causal element for genetically influenced diseases such as cancer (Kunkel, 2004). Three mechanisms govern the mutation frequency during replication including, DNA polymerase fidelity (the likelihood of the incorrect rather than the correct nucleotide incorporation), proofreading through 3’to 5’exonuclease activity (including

the rate of mismatch extension) and post-replication mismatch repair mechanisms (Echols and Goodman, 1991).

Previous work by our group has explored fidelity measurements through steady state kinetics as well as mismatch extension and exonuclease activity of the apPOL (Wingert et al., 2013). In this investigation we will explore fidelity measurements through rapid quench single turnover experiments, which are thought to be the most rigorous technique for quantification of the kinetics of nucleotide incorporation (Johnson, 2010). One method for the measurement of fidelity adapted for pre-steady state experiments by Johnson compares the frequency of incorporation of the incorrect nucleotide to the frequency of the incorporation of the correct nucleotide. Such that the frequency of incorporation is defined as $(k_{pol}/K_{d,app})_{\text{incorrect}}/(k_{pol}/K_{d,app})_{\text{correct}}$, where k_{pol} equals maximum rate of incorporation and $K_{d,app}$ is an approximate measure of the K_m. In this way the frequency of correct and incorrect incorporation is measured as a function of nucleotide concentration, and a rectangular hyperbola fit of the single turnover experimental data to find the k_{obs} (single-turnover rate constant) in order to evaluate k_{pol} and $K_{d,app}$ (Bertram et al., 2010; Johnson, 2010).

Although some work by Kennedy et al. has been done in this area already, in light of our bioinformatic analysis of the Pfprex gene, and subsequent addition of 38 amino acids to the N-terminus of the apPOL from the previously published construct, further investigation is warranted. Our group has shown a 750-fold increase in steady state activity for our construct over previously published results from an apPOL fragment and fidelity measurements more in line with a replicative DNA polymerase (Wingert et al., 2013).

The fidelity of the apicoplast was estimated to be $(5.5 \pm 2.9) \times 10^{-6}$ for a β-lactamase reversion frequency for the apPOL and $(2.3 \pm 1.0) \times 10^{-4}$ for the exonuclease negative

polymerase (apPOL^{exo-}). Error rates for apPOL^{exo-} of 3.9 x 10^{-5} were calculated for single-base substitution using the M13mp2 *Lac*Z*α* forward mutation assay. In addition, steady state values for fidelity of the form $(k_{cat}/K_m)_{incorrect}/(k_{cat}/K_m)_{correct}$ were reported, but under unusual conditions such that the DNA substrate concentration was very low and the polymerase concentration exceeded by the substrate concentration by two-fold (suggesting something more similar to pre-steady state conditions) (Kennedy et al., 2011). It is our hypothesis that the sole known apicoplast polymerase is in fact a high fidelity polymerase typical of other A family high fidelity polymerases. Due to the 750-fold steady state increase in activity for our apPOL construct and the published concerns of steady state fidelity measurements, we here report the pre-steady state fidelity measurements of 12 of the possible 16 nucleotide insertions by our apPOL^{exo-} construct (four incorrect nucleotide incorporations are too slow for accurate quantitation) (Johnson, 2010; Tsai and Johnson, 2006). We also explore the role of stalling of the apPOL following the incorporation of an incorrect nucleotide and an unusual terminal transferase activity by the apPOL.

Material and Methods

Materials– Oligodeoxynucleotides (Table 1) used for mutagenesis were purchased from either Integrated DNA Technologies or the Iowa State University DNA Facility. DNA sequencing was performed at the Iowa State University DNA Facility. Nickel-agarose was purchased from Sigma-Aldrich Chemical Company. Deoxyribonucleotides were purchased from Invitrogen.

Cloning of the P. falciparum apicoplast DNA polymerase and creation of the exonuclease negative mutant– The open reading frame containing the apicoplast DNA polymerase encodes

for a poly-protein made up a DNA primase, helicase, and polymerase (Seow et al., 2005). A linker region between the primase and helicase is proteolytically cleaved and it is assumed that the protein is cleaved between the helicase and polymerase as well (Lindner et al., 2011). Based on protein sequence alignments of POM1 from the Plasmodia genus, we identified a likely boundary for the polymerase protein spanning amino residues 1470 through 2016. This protein sequence was then converted to DNA sequence using optimal *E. coli* codons and synthesized (Genescript). The synthesized gene was sub-cloned from the puc18 vector into the pet28b expression vector using the NdeI and BamHI restriction sites. The Quickchange method of mutagenesis was employed to produce the exonuclease negative mutant (apPOL^{exo-}). The sequence of the forward primer used for mutagenesis is as follows: 5'-gatattaaatattgcggcctgaatatccaaaccacgggtctggaagtg-3' with the mutagenic codons for positions 1552 and 1554 are shown in bold. The reverse primer was the reverse complement of the forward.

Protein Expression, and Purification– The purification protocol for the wild-type (apPOL) and exonuclease deficient (apPOL^{exo-}) polymerases were identical. Either the pet28-apPOL or pet28-apPOL^{exo-} vectors were transformed into BL21(DE3) *E. coli* cells and a single colony was used to inoculate 100-ml flasks of LB-kanamycin that were shaken for 16 h at 37 °C. 10 ml of starter culture was used to inoculate two 1-liter flasks of LB-kanamycin per protein, which were shaken at 225 rpm at 37 °C to an A600 of 0.8. The flasks were then cooled to 18 °C, and expression was induced by the addition of 0.2 mM (final) isopropyl 1-thio-β-d-galactopyranoside. After 16 h the cells were collected by centrifugation at 4000 × g for 20 min, and pellets were frozen at −20 °C.

Target protein purification relied on a hexahistidine tag provided by the pet28 vector. Cell pellets containing expressed apPOL or apPOL^{exo-} (2 liters) were resuspended in 100 ml of Loading Buffer (20 mM Tris-HCl, 500 mM NaCl, 5 mM imidazole, 20% glycerol at pH 8.0 [4°C]) and stored frozen at -20°C. Following thawing, Lysis was accomplished by passage thru an EmulsiFlex-C5 (Avestin, Inc.) at ~16 kpsi. The lysate was clarified by centrifugation at ~32,500 x g and the supernatant loaded onto ~ 3 mL of Ni-Agarose resin. The column was washed with 100 mL of Loading Buffer, followed by 100 mL of High Salt Buffer (5 mM imidazole, 1 M NaCl, 20 mM Tris-HCl, 20% glycerol at pH 8.0 [4°C]) and then 30 mL (10 column volumes) of Ni Wash Buffer (20 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, 20% glycerol at pH 8.0 [4°C]). The protein was eluted with Elution Buffer (20 mM TRIS-Cl, 200 mM NaCl, 150 mM imidazole, 20% glycerol at pH 8.0 [4°C]). The fractions containing protein were pooled and loaded onto 320 mL HiLoad™ 26/60 Superdex™ 200 prep grade column equilibrated with 20 mM TRIS-Cl, 400 mM NaCl, 20% glycerol at pH 8.0 (4°C). The column was washed with 400 mL of 20 mM TRIS-Cl, 400 mM NaCl, 20% glycerol at pH 8.0 (4°C). The fractions containing protein were pooled and the concentration determined spectrophotometrically using an extinction coefficient (λ280 nm = 56,750 $M^{-1}cm^{-1}$) calculated from the deduced protein composition.

Burst Kinetics Experiment- The DNA template was created by annealing by annealing a 5'-Hex labeled 20-nucleotide primer (P1) (Table 1) to a 26-nucleotide fragment (T5) (Table 1). Reactions were performed at pH 7.9 and 25^{o}C in 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitiol, and 0.1 mg/ml bovine serum albumin. The polymerase was diluted to 20 mM Tris-HCl, 50 mM potassium acetate, and 0.1 mg/ml BSA (standard buffer) prior to addition to the reaction. The polymerase and DNA were preincubated

without magnesium and the reaction started with 100 μM dATP and 20 mM magnesium. The assay was performed with 2 μM DNA substrate, 1 μM apPOL^{exo-} , at 100 μM dATP with a timescale of 5-2100 milliseconds. The experiment was performed on a Biologic QFM-400 instrument. Reaction products were resolved with 16% 19:1Acrylamide/Bisacrylamide denaturing PAGE containing 7.5 M urea in 1X Tris-Borate-EDTA (TBE) buffer. Gels were run for 3.5 – 4 h at a constant power of 60 W, visualized using a Typhoon Fluorescence imager, and analyzed using the ImageJ software (NIH). The data were then plotted and fitted to the following equation:

$$y = a(1 - \exp(-k_{obs}x)) + k_{ss}x \qquad \text{Equation 1.}$$

where a, k_{obs}, k_{ss}, and x represent the burst amplitude, the single exponential rate constant, the steady-state rate constant, and time, respectively. The errors given for each parameter are standard errors to the fit.

Pre-Steady state Polymerase Extension Kinetics– DNA templates were made by annealing a 5'-Hex labeled 20-nucleotide primer (P1 in Table 1) to a 26-nucleotide template (T1-T16). Reactions were performed at pH 7.9 and 25^{o}C in 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitiol, and 0.1 mg/ml bovine serum albumin. The polymerase was diluted to 20 mM Tris-HCl, 50 mM potassium acetate, and 0.1 mg/ml BSA (standard buffer) prior to addition to the reaction. For correct nucleotide incorporations, pre-steady-state assays were performed with 200 nM DNA substrate, 2 μM apPOL^{exo-} , and varying amounts of the correct dNTP ranging from 0.5-1024 μM. The polymerase and DNA were preincubated in standard buffer without magnesium and the reaction started with equal volume amount of varying concentrations of dNTPs, 20 mM magnesium, and standard buffer.

Pre-steady-state assays were performed with 200 nM DNA substrate, 2 μM apPOL^{exo-} , and varying amounts of the incorrect dNTP ranging from 20-8000 μM to evaluate misincorporations reactions. Reactions were quenched at varying timescales (5 milliseconds to 2.1 seconds for correct incorporation and 5 seconds to 4 hours for misincorporation) with 0.5 M EDTA. Correct incorporations experiments were performed using a Biologic QFM-400 rapid quench apparatus. Misincorporations were performed by hand using by a stop-time method (pulling smaller volume out of master mix and quenching at each time point). Reaction products were resolved with 16% 19:1Acrylamide/Bisacrylamide denaturing PAGE containing 7.5 M urea in 1X Tris-Borate-EDTA (TBE) buffer. Gels were run for 3.5 – 4 h at a constant power of 60 W, visualized using a Typhoon Fluorescence Imager, and analyzed using the ImageJ software (NIH). The k_{obs} rates were fitted for each nucleotide concentration using an exponential rise to max equation:

$$y = a * (1 - exp(-k_{obs}x)) \quad \text{Equation 2.}$$

where a, k_{obs}, and x represent the amplitude, single exponential rate constant, and time, respectively. Each condition was repeated at least 3 times for correct and incorrect incorporation and repeats were fitted globally to the following equation:

$$\left(k_{obs} = \frac{k_{pol}[dNTP]}{K_{d,app}+[dNTP]}\right) \quad \text{Equation 3.}$$

where k_{pol}, $K_{d,app}$, and x represent the single-exponential rate constant at saturating nucleotide, the dissociation constant for the polymerase-nucleotide complex, and the concentration of nucleotide, respectively. The errors given for each parameter are standard errors to the fit.

Mechanism of Multiple Nucleotide Incorporations– For a more accurate determination of substrate and product amounts we used peak deconvolution software (MagicPlot 2.5, student

edition) to determine the relative amounts of the substrate (n) and products (n+1 through n+6). A plot profile for each lane was generated using NIH ImageJ and then MagicPlot was used to determine peak areas assuming a Gaussian distribution for each peak. The relative areas for each peak were converted to amounts of DNA and time course was globally fitted using Dynafit software. The script file can be found in the Supplemental Material.

Results and Discussion

P. falciparum apPol^{exo-} Burst Kinetics- All misincorporation kinetics were performed using the exonuclease negative mutant of apPOL (apPOL^{exo-}). This mutant purifies in a similar manner to the wild-type construct to greater than 95% purity according to SDS-PAGE analysis. The use of an exonuclease negative mutant is standard for the field and enables the fidelity of the polymerase active site to be isolated without the confounding activity of the separate exonuclease domain contributing to the observed product profile. To initiate these studies, a pre-steady state kinetic analysis for the correct incorporation of dATP across template dTMP was performed to look for the presence of biphasic kinetics. apPOL^{exo-} was preincubated with a 2:1 molar excess of DNA-(P1/T5) in reaction buffer (no magnesium) and placed into syringe 1 and mixed an equal volume of 100 μM dATP, 20 mM magnesium, and reaction buffer from syringe 2. The reaction was quenched with 0.5 M EDTA at varying time points ranging from 5 to 2100 ms at 25°C. Each time point was analyzed by sequencing gel and visualized using a Typhoon Fluorescence Scanner. Figure 1 shows the resulting time course demonstrating biphasic kinetics (a rapid burst followed by linear phase). The burst phase rate (k_{obs}) was found to be 17.4 ± 2.1 s^{-1} with a subsequent slower steady state rate (k_{ss}) of 0.0977 ± 0.003 s^{-1}. For many enzymes the rate limiting step for the first turnover (the burst) is chemistry or a conformational step receding

catalysis (Fiala and Suo, 2004; Johnson and Johnson, 2001). Following this initial burst phase, DNA polymerases are can be limited by the rate of the slow dissociation of the polymerase from the replicated product or by the slow binding/assembly of the polymerase onto the primer-template. Preliminary steady-state experiments suggest that in the case of apPOLexo-, it is the slow assembly of the polymerase onto the primer-template that determines the steady-state rate. Under certain kinetic regimes the burst amplitude is equal to active enzyme, which was determined to be 0.0951 ± 0.0042, indicating that the apPOL^{exo-} is at least 19.02% active. This represents the minimum for active protein since the rate of the reverse reaction and the dissociate rate of the pyrophosphate is currently unknown (Pryor and Washington, 2011).

P. falciparum apPOl^{exo-} Misincorporation Kinetics– Since apPOL^{exo-} displays burst kinetics, pre-steady-state kinetics can be used to isolate the steps up to and including the chemical step of the polymerization reaction. This ability to exclude events subsequent to catalysis (e.g., product dissociation) is the major advantage of a pre-steady state kinetic analysis as compared to a steady-state analysis. *In vivo* substitution error rates for high fidelity polymerases with intrinsic exonuclease activity (proofreading) are commonly in the range of 10^{-7} to 10^{-8}. This extreme high fidelity is governed by the selectivity of the polymerase active site (10^4 to 10^6 for correct over incorrect incorporation) and the intrinsic proof-reading activity of the exonuclease active site that removes 90 to 99% of the misincorporation errors (Kunkel, 2004). In order to isolate the polymerase active site, an exonuclease deficient construct was generated and used in all misincorporation experiments (apPOL^{exo-}).

To eliminate the potentially slow binding of apPOL^{exo-} to the DNA substrate, a 10-fold excess of the enzyme was preincubated with the Hex-labeled primer-template DNA for several minutes to ensure that the enzyme-DNA complex was formed at t = 0 ms. This is done to ensure

that primer extension follows single-turnover kinetics and the data indicate that the rate observed in these time-courses were essentially identical to the burst rate observed in the [DNA] > [POL] experiment .

Figure 2 illustrates representative gel and schematic for the correct incorporation of T:dAMP (DNA-P1/T1) and the incorrect incorporation of G:dAMP (DNA-P1/T4). As a point of clarification we will the shorthand notation for a primer/template base pair as T:dAMP (e.g., where dTTP is the incoming base and dAMP is the template base). Figure 3 illustrates the representative data defining the calculation of (k_{pol}) and ($K_{d,app}$) for the correct and incorrect incorporation of the above examples. We have designed our DNA templates to contain a five nucleotide homopolymeric extension in order to observe continued polymerase turnover and to increase the separation between the substrate and product on the gel. For the calculation of k_{obs}, the percentage primer extended was calculated as [total product/ (total product + substrate)]. Total product represents any extension of the primer (fluorescence of bands 21 and above). Substrate was measured as the fluorescence of band 20 (the unextended primer).

For the correct incorporation (T:dAMP), values for k_{pol} , $K_{d,app}$, and substrate specificity ($k_{pol}/K_{d,app}$) are 46 ± 3 s^{-1}, 87 ± 17 μM, and 0.5 $\mu M^{-1}s^{-1}$ respectively (Figure 3C and Table 2). For the incorrect incorporation (G:dAMP), values for k_{pol} , $K_{d,app}$, and substrate specificity ($k_{pol}/K_{d,app}$) of 0.0074 ± 0.0006 s^{-1}, 224 ± 56 μM, and 3.3 x 10^{-5} $\mu M^{-1}s^{-1}$ respectively (Figure 3D and Table 2). Data in Table 2 represent *at least* the global fit of 3 repeats for each condition. We find there is a wide range among the overall rates of misincorporation with the slowest observable (G:dGMP) mismatch being approximately ~20 times slower than the fastest mismatch (T:dGMP) (Table 2). Correct incorporation shows a much smaller range with G:dCMP correct incorporation being 2.3 times slower than A:dTTP correct incorporation (26 and 59 s^{-1}

respectively). In addition, there is a wide range $K_{d,app}$ observed with C:dAMP being ~ 20 times greater than A:dAMP, which is the lowest $K_{d,app}$ (1906 and 97 μM respectively) (Figure2). Correct incorporations showed a similar difference in $K_{d,app}$ with the G:dCMP correct incorporation being ~16 fold lower than the T:dAMP correct incorporation (5.6 and 87.2 μM respectively)(Table 2). A comparison of k_{pol}, $K_{d,app}$, and substrate specificity is presented visually in Figure 4A, 4B, and 4C.

There is a large discrepancy between the polymerase rates (k_{pol}) for apPOL^{exo-} by Kennedy et al. and that we are reporting here (Table 2) (Kennedy et al., 2011; Wingert et al., 2013). There was virtually no difference in k_{cat} values between correct and incorrect incorporation, with one condition in particular (G:dTMP), with a greater k_{cat} than correct (A:dTMP) (Kennedy et al., 2011). To the contrary, however; we have found a ~550-fold difference in *kpol* between correct and incorrect incorporations for these conditions. Indeed for all misincorporations (A, G, C) across template dAMP there was greater than 5000-fold decrease in *kpol* from the correct incorporation. While the reason for the discrepancy is not clear, as was mentioned in the introduction, our construct has an additional 38 residues at the N-terminus. These residues are well conserved within the genus *Plasmodia*, but not found outside of it. It is possible that the lack of these residues is responsible for the much less reported activity.

The fidelity of correct nucleotide insertion can then be determined by the ratio of the specificity for the incorrect substrate over the specificity for the correct substrate (eq 3).

$$fidelity = \frac{\left(\frac{k_{pol}}{K_{d,app}}\right)_{incorrect}}{\left(\frac{k_{pol}}{K_{d,app}}\right)_{correct}} \qquad \text{Equation 4}$$

Polymerase selectivity of correct over incorrect measured by the fidelity is a useful parameter for exploring the base substitution error spectrum of the apPOl^{exo-} (Boosalis et al., 1987; Johnson, 2010). Fidelity data are presented in Table 2 and visually illustrated in Figure 4D. The fidelity value for a correct incorporation is defined as unity. Misincorporation frequency (fidelity) ranged from 2.4 x 10^{-4} for G:dTMP mismatch to 9.0 x 10^{-6} for the C:dAMP mismatch. This represents a 26.7 fold specificity preference for dGMP across dTMP as opposed to dCTP across dAMP. We find no clear trends in the misincorporation data, other than pyrimidine-pyrimidine misincorporations seem to be strongly disfavored. The rate of polymerization was found to be very slow for pyrimidine-pyrimidine, (on the order of hours to see *any* primer extension) to the point of being too difficult to quantify and analyze in a rigorous manner. We have estimated a minimum for the fidelity of 3 of the 4 pyrimidine-pyrimidine mismatches (C:dTMP, T:dCMP, and C:dCMP) by estimating a k_{pol} based on the maximum k_{obs} for those experiments and a lower bound for the $K_{d,app}$ as 100 μM. These values represent a conservative estimate of the lower bound of Fidelity for these mismatches. As we have selected a $K_{d,app}$ of 100 μM (majority of the $K_{d,app}$'s are greater than this), it is assumed that the fidelities for pyrimidine-pyrimidine mismatches are greater than we have estimated. For the T:dTMP mismatch we discovered a problem with dATP nucleotide contamination that did not allow us to estimate any kinetic constants. At incorrect dNTP concentrations in the low mM range, even a 0.001% contamination of the correct nucleotide is enough to interfere with the kinetics of this assay (e.g. at a concentration of 5 mM dTTP, a 0.001% dATP contamination provides enough correct nucleotide to extend 25% of the DNA substrate). With that in mind we were careful to monitor our data for possible correct incorporation bursts at the beginning of our time courses. Similar precautions against contamination have been noted by others but could not be carried out by

Kennedy et al. due to the use of single-time points and a very low DNA concentration (Boosalis et al., 1987; Kennedy et al., 2011).

In Figure 5 we have compared the reported fidelity measurements of apPOL here and previously to that of several family-A polymerases including the Klenow fragment of *Escherichia coli* polymerase I and the human mitochondrial DNA polymerase Pol γ. apPOL^{exo-} appears to display a similar pattern of fidelity measurements between each of the 5 reported polymerases. G:TMP has the lowest calculated fidelity measurements across all five polymerases. Additionally, it appears that pyrimidine-pyrimidine misincorporations are strongly disfavored (Bertram et al., 2010; Kennedy et al., 2011; Lee and Johnson, 2006; Wingert et al., 2013). We have noticed, as well, that despite the large differences in the rate of polymerization between Kennedy et al. and the data reported here, the relative ranking of misincorporations is very similar save for two differences. We found the A:dAMP mismatch to be more highly favored, consistent with the error spectrum of the Klenow fragment. Secondly, our fidelity data are approximately 5-20 fold less at all conditions (except for A:dAMP) than previously reported (Kennedy et al., 2011). It had been speculated by Schoenfeld et al., that the high mutation spectrum reported in Kennedy et al., may indicate that apPOL had more of "role in diversity generation than accurate replication" for the pathogen as a part of method to evade host immune systems (Schoenfeld et al., 2013). While that may still be the case, the error rates reported here (10^{-4}-10^{-6}) are typical of high fidelity polymerases. Additionally, the apPOL has intrinsic exonuclease activity which, for other polymerases typically removes 90-99% of mismatches (Kunkel, 2004). As the apPOL is the only known polymerase targeted to the apicoplast, its role in maintaining genetic integrity is important.

Mechanism of Multiple Nucleotide Incorporations– In and effort to understand better the mechanism for multiple successive nucleotide incorporations we fit the product amount at each possible product position (n+1 through n+6) as a function of time to several proposed mechanisms (Figure 6). We discriminated between two possible mechanisms (Scheme 1 or Scheme 2) using Dynafit 3 as described in the Supplemental Material. One mechanism included a translocation step (k_t) and the other mechanism had only a polymerization step. The resulting kinetic constants are found in Table 3 and Table 4.

$$ED_n + dNTP \underset{k_{on},k_{off}}{\Longleftrightarrow} ED_n dNTP \xrightarrow[k_{pol}]{} ED_{n+1} + PP_i \qquad \text{Scheme 1}$$

$$ED_n + dNTP \underset{k_{on},k_{off}}{\Longleftrightarrow} ED_n dNTP \xrightarrow[k_{pol}]{} E^*D_{n+1} + PP_i \xrightarrow[k_t]{} ED_{n+1} \qquad \text{Scheme 2}$$

The data fit best to Scheme 1, indicating that the rate of catalysis or a conformational change preceding catalysis is the rate limiting for sequential polymerization since inclusion of a translocation step was not necessary to explain the product time courses. The rate of translocation must be at least 10-fold greater than k_{pol}. The fitted rate constants were as follows: $k_{pol,}$ 24.5±2.5 s^{-1}; k_{off1}, 13.47±4.23 s^{-1}, k_{on1}, 0.575±0.002 $\mu M^{-1}s^{-1}$; and [ED], 0.14 μM. k_{on1} is slow and k_{off1} is fast, suggesting why the first products (n+1, n+2) do not go to zero during the time of the experiment. The processivity (k_{pol}/k_{off}) is approximately 2, which means that for every successful nucleotide incorporation, there is a ~75% the next nucleotide will be incorporated as well.

Stalling following mismatch and terminal transferase activity- During the course of the pre-steady state misincorporation experiments we observed two interesting phenomena, stalling

of primer extension following insertion of a mismatch and the extension of non-template single stranded DNA. First, as mentioned above, we have designed our DNA templates to contain a five nucleotide homopolymeric extension in order to observe continued polymerase turnover. This is a useful construct for steady state misincorporation studies as it helps to eliminate problems with slow product dissociation when a processive polymerase is stalled at the n+1 DNA product (Hacker and Alberts, 1994; Johnson, 2010). While that is not a concern for pre-steady state experiments because we are only interested in the first turnover from a kinetic standpoint, this particular DNA substrate does allow for added separation between the primer and primer extension products. As the turnover for the correct incorporation (bands 22-26, Figure 7) is much faster than the incorrect, often the misincorporation band (band 21, Figure 6) is not present allowing for greater separation for measurement between primer extension and primer. For three conditions (Figure 7D, 7E,7F), however; we observed stalling of primer extension at either the misincorporation (band 21) (A:AMP, A:GMP) or at the two subsequent correct incorporations (band 22 and 23) (T:GMP). This was seen at all concentrations of dNTPs that were measured for these conditions. For all other nucleotide insertions, this pattern was not observed.

Work by Johnson and Beese may shine some light on to these observations. They crystallized the thermophilic *Bacillus* DNA Polymerase I fragment with each of the possible mismatches and observed the effect of the mismatch on the structure of the active site. For the A:dAMP mismatch they observed that bases were not paired within the active site because of steric clashes with residues due to rearrangement of the active typical of mismatches. We observed that the A:dAMP was stalled at the n+1 position suggesting that the active site was distorted enough to prevent release and translocation of the mismatch. The observable effect of

the mismatch disappears once the translocation does occur however. For the A:dGMP mismatch the crystal structure show that the active site was severely distorted with the finger domain blocking the pre-insertion site for the next nucleotide. This is a likely explanation for the stalling of the n+1 site for A:dGMP position as well. Finally, we observed that for the T:dGMP mismatch stalling was observed at the first two correct incorporations. Johnson and Beese observed distortions throughout the active site including the placement of primer and template strands. It appears that contrary to the previous described stalling, the polymerase is able to catalyze the mismatch and move to the next incorporation. At that point, however; it appears that the mismatch distorts the DNA duplex binding region enough to stall the normally very fast correct incorporation of dTTP into dAMP. This is appears to be a form of "memory" for the mismatch that has been observed in other polymerases (Johnson and Beese, 2004).

Severe reduction of mismatch extension because of active site distortion may be a way for polymerases with intrinsic exonuclease activity (or encoded by a separate gene), to increase the likelihood of mismatch being excised (Goodman and Fygenson, 1998; Kunkel, 2004). It is possible that while two of these mismatches (A:dAMP and T:dGMP) have lower fidelities as measured by the our pre-steady state data (ranked 2 and 3 respectively), *in vivo*, the presences of these mutations may be more unlikely then predicted because the effect of stalling observed here on mismatch extension. This could also affect the interpretation of steady state fidelity measurements if the rate of product dissociation becomes rate limiting as result. More work on the kinetics of this finding will need to be done to understand its effect, on the overall fidelity of the polymerase.

The second phenomenon we observed was non-template nucleotide additions to the primer strand by the apPOL^{exo-} beyond the end of the template. This pattern was seen for all 4

deoxyribonucleotides at sufficiently high nucleotide concentration, though not to the same extent. After 1 hour and with a dGTP concentration of 1.28 mM, a 10 base single-stranded overhangs can be seen (Figure 7A). To a lesser extent, the apPOL^{exo-} will extend dATP (4 bases, Figure 7B), dCTP (3 bases, Figure 7C), and dTTP (1 base, Figure 7D) as well. 3' blunt-end nucleotide addition has been observed in a many different polymerases from all three domains but to our knowledge this is the first example of multiple nucleotide additions being made by an A family DNA polymerase (Fiala et al., 2007). While we are unsure if these observations have any physiological relevance at this time, it is intriguing to speculate that this could have a role in non-homologous end joining observed in polymerases involved in DNA repair (e.g. human DNA polymerase μ) (Andrade et al., 2009). Of note, for most polymerases, dATP is strongly favored for extension in contrast to the preferential extension of dGTP found here (Obeid et al., 2010).

The apPOL is a potential target for drug discovery, and this report provides essential data on the polymerase for further work. These experiments have established the fidelity data needed for evaluating the effect of any nucleotide analogs or allosteric small-molecule inhibitors on fidelity.

Acknowledgements

This work was financially supported by The Roy J. Carver Charitable Trust and Iowa State University

References

Andrade, P., Martín, M.J., Juárez, R., López de Saro, F., and Blanco, L. (2009). Limited terminal transferase in human DNA polymerase mu defines the required balance between accuracy and efficiency in NHEJ. Proc. Natl. Acad. Sci. U. S. A. *106*, 16203–16208.

Berdis, A.J. (2008). DNA polymerases as therapeutic targets. Biochemistry (Mosc.) *47*, 8253–8260.

Bertram, J.G., Oertell, K., Petruska, J., and Goodman, M.F. (2010). DNA Polymerase Fidelity: Comparing Direct Competition of Right and Wrong dNTP Substrates with Steady State and Pre-Steady State Kinetics. Biochemistry (Mosc.) *49*, 20–28.

Boosalis, M.S., Petruska, J., and Goodman, M.F. (1987). DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. J. Biol. Chem. *262*, 14689–14696.

Echols, H., and Goodman, M.F. (1991). Fidelity mechanisms in DNA replication. Annu. Rev. Biochem. *60*, 477–511.

Fiala, K.A., and Suo, Z. (2004). Mechanism of DNA Polymerization Catalyzed by *Sulfolobus solfataricus* P2 DNA Polymerase IV †. Biochemistry (Mosc.) *43*, 2116–2125.

Fiala, K.A., Brown, J.A., Ling, H., Kshetry, A.K., Zhang, J., Taylor, J.-S., Yang, W., and Suo, Z. (2007). Mechanism of template-independent nucleotide incorporation catalyzed by a template-dependent DNA polymerase. J. Mol. Biol. *365*, 590–602.

Goodman, M.F., and Fygenson, D.K. (1998). DNA polymerase fidelity: from genetics toward a biochemical understanding. Genetics *148*, 1475–1482.

Hacker, K.J., and Alberts, B.M. (1994). The slow dissociation of the T4 DNA polymerase holoenzyme when stalled by nucleotide omission. An indication of a highly processive enzyme. J. Biol. Chem. *269*, 24209–24220.

Johnson, K.A. (2010). The kinetic and chemical mechanism of high-fidelity DNA polymerases. Biochim. Biophys. Acta Bba - Proteins Proteomics *1804*, 1041–1048.

Johnson, A.A., and Johnson, K.A. (2001). Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase. J. Biol. Chem. *276*, 38090–38096.

Johnson, S.J., and Beese, L.S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. Cell *116*, 803–816.

Kalanon, M., and McFadden, G.I. (2010). Malaria, Plasmodium falciparum and its apicoplast. Biochem. Soc. Trans. *38*, 775–782.

Kennedy, S.R., Chen, C.-Y., Schmitt, M.W., Bower, C.N., and Loeb, L.A. (2011). The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite Plasmodium falciparum. J. Mol. Biol. *410*, 27–38.

Kunkel, T.A. (2004). DNA Replication Fidelity. J. Biol. Chem. *279*, 16895–16898.

Lee, H.R., and Johnson, K.A. (2006). Fidelity of the Human Mitochondrial DNA Polymerase. J. Biol. Chem. *281*, 36236–36240.

Lindner, S.E., Llinás, M., Keck, J.L., and Kappe, S.H.I. (2011). The primase domain of PfPrex is a proteolytically matured, essential enzyme of the apicoplast. Mol. Biochem. Parasitol. *180*, 69–75.

Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Welte, W., and Marx, A. (2010). Replication through an abasic DNA lesion: structural basis for adenine selectivity. Embo J. *29*, 1738–1747.

Prusty, D., Dar, A., Priya, R., Sharma, A., Dana, S., Choudhury, N.R., Rao, N.S., and Dhar, S.K. (2010). Single-stranded DNA binding protein from human malarial parasite Plasmodium falciparum is encoded in the nucleus and targeted to the apicoplast. Nucleic Acids Res. *38*, 7037–7053.

Pryor, J.M., and Washington, M.T. (2011). Pre-steady state kinetic studies show that an abasic site is a cognate lesion for the yeast Rev1 protein. Dna Repair *10*, 1138–1144.

Schoenfeld, T.W., Murugapiran, S.K., Dodsworth, J.A., Floyd, S., Lodes, M., Mead, D.A., and Hedlund, B.P. (2013). Lateral Gene Transfer of Family A DNA Polymerases between Thermophilic Viruses, Aquificae, and Apicomplexa. Mol. Biol. Evol. *30*, 1653–1664.

Seow, F., Sato, S., Janssen, C.S., Riehle, M.O., Mukhopadhyay, A., Phillips, R.S., Wilson, R.J.M. (Iain), and Barrett, M.P. (2005). The plastidic DNA replication enzyme complex of Plasmodium falciparum. Mol. Biochem. Parasitol. *141*, 145–153.

Tsai, Y.-C., and Johnson, K.A. (2006). A New Paradigm for DNA Polymerase Specificity †. Biochemistry (Mosc.) *45*, 9675–9687.

Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., and McFadden, G.I. (1998). Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. Proc. Natl. Acad. Sci. *95*, 12352–12357.

Weissig, V., Vetro-Widenhouse, T.S., and Rowe, T.C. (1997). Topoisomerase II inhibitors induce cleavage of nuclear and 35-kb plastid DNAs in the malarial parasite Plasmodium falciparum. Dna Cell Biol. *16*, 1483–1492.

Williamson, D.H., Preiser, P.R., Moore, P.W., McCready, S., Strath, M., and Wilson, R.J.M. (2002). The plastid DNA of the malaria parasite Plasmodium falciparum is replicated by two mechanisms. Mol. Microbiol. *45*, 533–542.

Wingert, B., Parrott, E., and Nelson, S.W. (2013). Misincorporation, mismatch extension, and proofreading activity of the Plasmodium falciparum apicoplast DNA polymerase. Biochemistry (Mosc.).

World Health Organization (2013). World malaria report 2012. ([S.l.]: World Health Organization).

Yeh, E., and DeRisi, J.L. (2011). Chemical Rescue of Malaria Parasites Lacking an Apicoplast Defines Organelle Function in Blood-Stage Plasmodium falciparum. Plos Biol. *9*, e1001138.

Tables

Table 1. Primer/Template sequences for incorporation studies [a]

Name	Primer/Template
P1[b]	5'-/5HEX/CAGGTGTCAGTCAGCTAGTG-3'
T1	5'-AAAAAACACTAGCTGACTGACACCTG-3'
T2	5'-TTTTTACACTAGCTGACTGACACCTG-3'
T3	5'-GGGGGACACTAGCTGACTGACACCTG-3'
T4	5'-CCCCCACACTAGCTGACTGACACCTG-3'
T5	5'-TTTTTTCACTAGCTGACTGACACCTG-3'
T6	5'-AAAAATCACTAGCTGACTGACACCTG-3'
T7	5'-GGGGGTCACTAGCTGACTGACACCTG-3'
T8	5'-CCCCCTCACTAGCTGACTGACACCTG-3'
T9	5'-TTTTGGCACTAGCTGACTGACACCTG-3'
T10	5'-AAAAAGCACTAGCTGACTGACACCTG-3'
T11	5'-TTTTTGCACTAGCTGACTGACACCTG-3'
T12	5'-CCCCCGCACTAGCTGACTGACACCTG-3'
T13	5'-CCCCCCCACTAGCTGACTGACACCTG-3'
T14	5'-AAAAACCACTAGCTGACTGACACCTG-3'
T15	5'-TTTTTCCACTAGCTGACTGACACCTG-3'
T16	5'-GGGGGCCACTAGCTGACTGACACCTG-3'

[a] example: (DNA-P1/T1) Primer/template pair for T:dAMP

5'-/5HEX/ CAGGTGTCAGTCAGCTAGTG-3'

3'- GTCCACAGTCAGTCGATCACAAAAAA-5'

[b] Primer is 5' labeled with Hexachlorofluorescein

Table 2. Fidelity of single-nucleotide insertion by apPol^{exo-}

DNA	dNTP/template	$K_{d,app}$ (μM)[a]	k_{pol} (s^{-1})[a]	$k_{pol}/K_{d,app}$ ($\mu M^{-1} \cdot s^{-1}$)	Fidelity[h]
P1/T2	A·dAMP	96.5 ± 28.4	0.0085 ± 0.0008	8.8×10^{-5}	1.7×10^{-4}
P1/T4	G·dAMP	223.5 ± 56.4	0.0074 ± 0.0006	3.3×10^{-5}	6.3×10^{-5}
P1/T3	C·dAMP	1906.2 ± 339.3	0.009 ± 0.0008	4.7×10^{-6}	9.0×10^{-6}
P1/T1	T·dAMP	87.2 ± 16.7	45.7 ± 3.1	5.2×10^{-1}	1
P1/T12	G·dGMP	112.5 ± 37.0	0.0053 ± 0.0004	4.7×10^{-5}	3.2×10^{-5}
P1/T11	A·dGMP	1290.5 ± 180.2	0.0211 ± 0.0011	1.6×10^{-5}	1.1×10^{-5}
P1/T10	T·dGMP	586.5 ± 125.3	0.0799 ± 0.0064	1.4×10^{-4}	9.3×10^{-5}
P1/T9	C·dGMP	28.3 ± 10.8	41.4 ± 4.2	1.5×10^{0}	1
P1/T16	C·dCMP	100[b]	0.001[d]	1.0×10^{-5}[g]	2.2×10^{-6}[i]
P1/T15	A·dCMP	219.9 ± 55.8	0.01304 ± 0.0010	5.9×10^{-5}	1.3×10^{-5}
P1/T14	T·dCMP	100[b]	0.0019[e]	1.9×10^{-5}[g]	4.1×10^{-6}[i]
P1/T13	G·dCMP	5.57 ± 1.24	25.7 ± 1.5	4.6×10^{0}	1
P1/T6	T·dTMP	ND[c]	ND[c]	--	--
P1/T7	C·dTMP	100[b]	0.0005[f]	5×10^{-6}[g]	2.4×10^{-6}[i]
P1/T8	G·dTMP	209.9 ± 60.9	0.105 ± 0.0087	5.0×10^{-4}	2.4×10^{-4}
P1/T5	A·dTMP	27.9 ± 2.5	59.2 ± 1.5	2.1×10^{0}	1

[a] Standard error of the data fit to the Michaelis-Menten equation

[b] $K_{d,app}$ set at 100 μM (estimate)

[c] Not Detectable

[d] k_{pol} estimate bases on max k_{obs} value at 4 mM dCTP

[e] k_{pol} estimate bases on max k_{obs} value at 320 μM dTTP

[f] k_{pol} estimate bases on max k_{obs} value at 12 mM dCTP

[g] Estimated lower bound for selectivity

[h] Fidelity is calculated by $(k_{pol}/K_{d,app})_{incorrect}/(k_{pol}/K_{d,app})_{correct}$

[i] Estimated lower bound for Fidelity

Table 3. Kinetic Parameters from Global Fit of Multiple Nucleotide Incorporations with infinitely fast translocation step

Parameter	Initial	Fit	Error	%
k_{pol}	10	23.6[a]	2.17	9.2
k_{off1}	6	11.41[b]	3	26.3
k_{on1}	0.1	0.09725[c]	0.01865	19.2
[ED]	0.15	0.13745[d]	0.00237	1.7

[a] s^{-1}, [b] s^{-1}, [c] $\mu M^{-1}s^{-1}$, [d] μM

Table 4. Kinetic Parameters from Global Fit of Multiple Nucleotide Incorporations with fitted translocation step

Parameter	Initial	Fit	Error	%
k_{pol}	10	22.43[a]	2.003	8.9
k_{t}	10	10000000[b]	7.16E+09	71610
k_{off1}	6	9.787[c]	2.564	26.2
k_{on1}	0.1	0.08804[d]	0.01823	20.7
[ED]	0.15	0.13718[e]	0.002444	1.8

[a] s^{-1}, [b] s^{-1}, [c] s^{-1}, [d] $\mu M^{-1}s^{-1}$, [e] μM

Figures

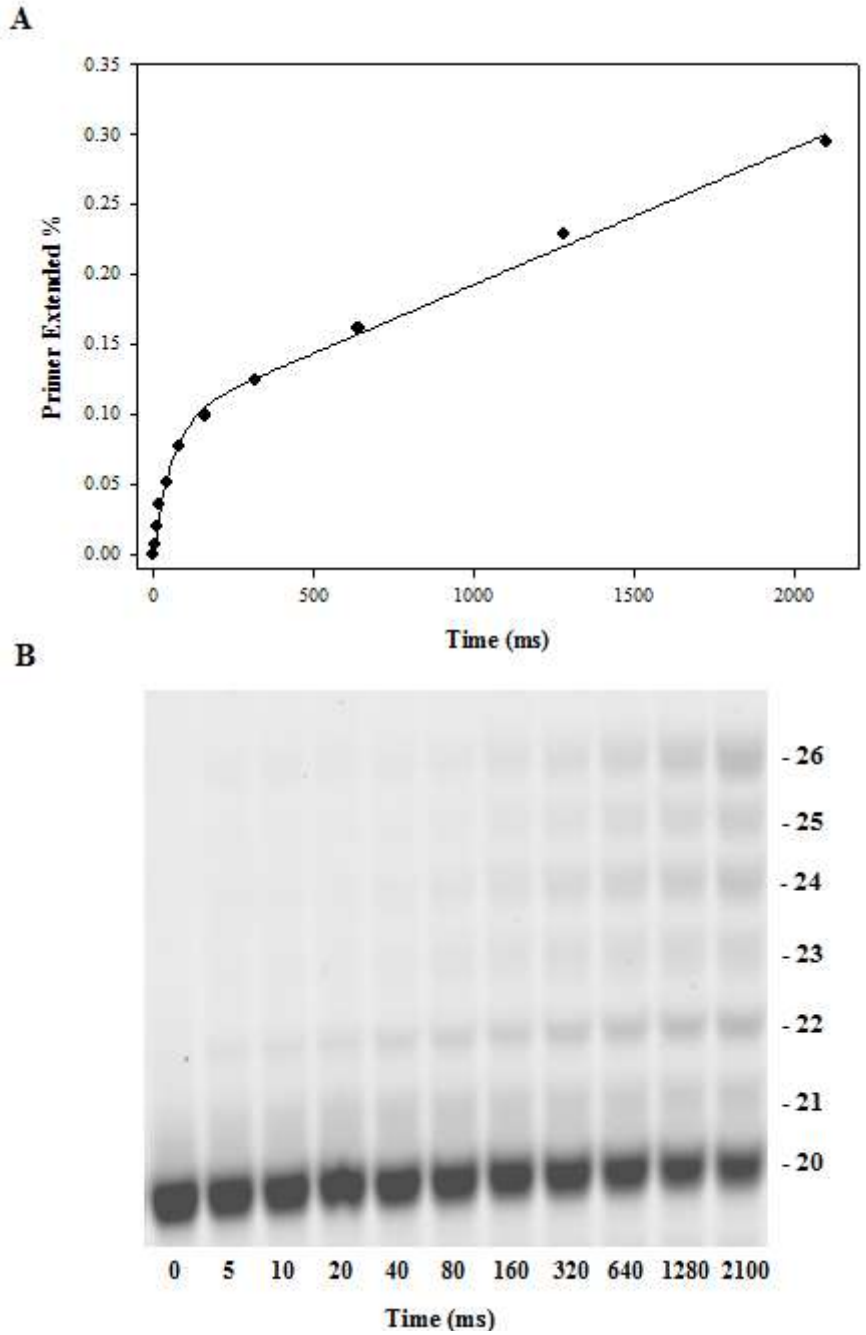

Figure 1: Nucleotide incorporation by apPOL^{exo-}. (A) 1 μM apPOL^{exo-} preincubated with 2 μM DNA (P1/T5) and then mixed with 100 μM dATP for varying lengths of time (5-2100 ms). The amount of primer extension was plotted as a function of time to equation 1. Amplitude was equal to 0.0951±0.0042, k_{obs} was equal to 17.4±2.1 s^{-1}, and k_{ss} was equal to 0.0977±0.0003 s^{-1}. (B) Gel image of primer extension

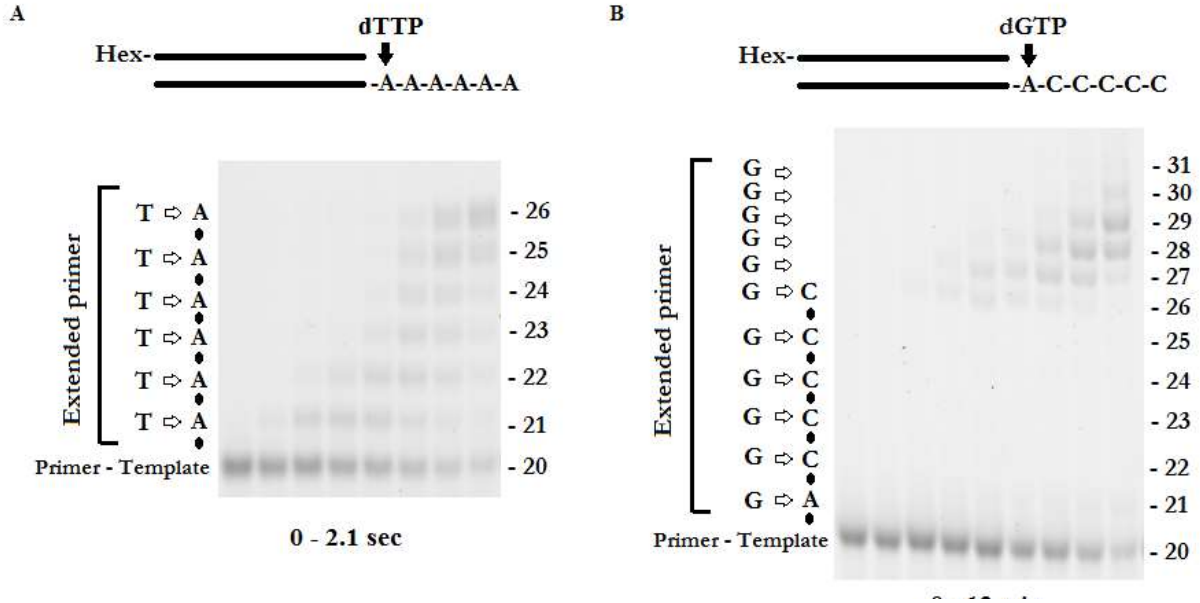

Figure 2. Pre-steady state kinetics for correct and incorrect incorporation of nucleotides. (A) Correct incorporation of dTTP across dAMP. The band labeled 20 is the primer-template. Bands 21-26 are the correct nucleotide extension of the primer template. (B) Incorrect incorporation of dGTP across dAMP. The band labeled 20 is the primer-template. Band 21 is position of the mismatch. Bands 22-26 are the correct nucleotide extension of the primer template. Bands 27-31 are non-template single-stranded extension of dGTP.

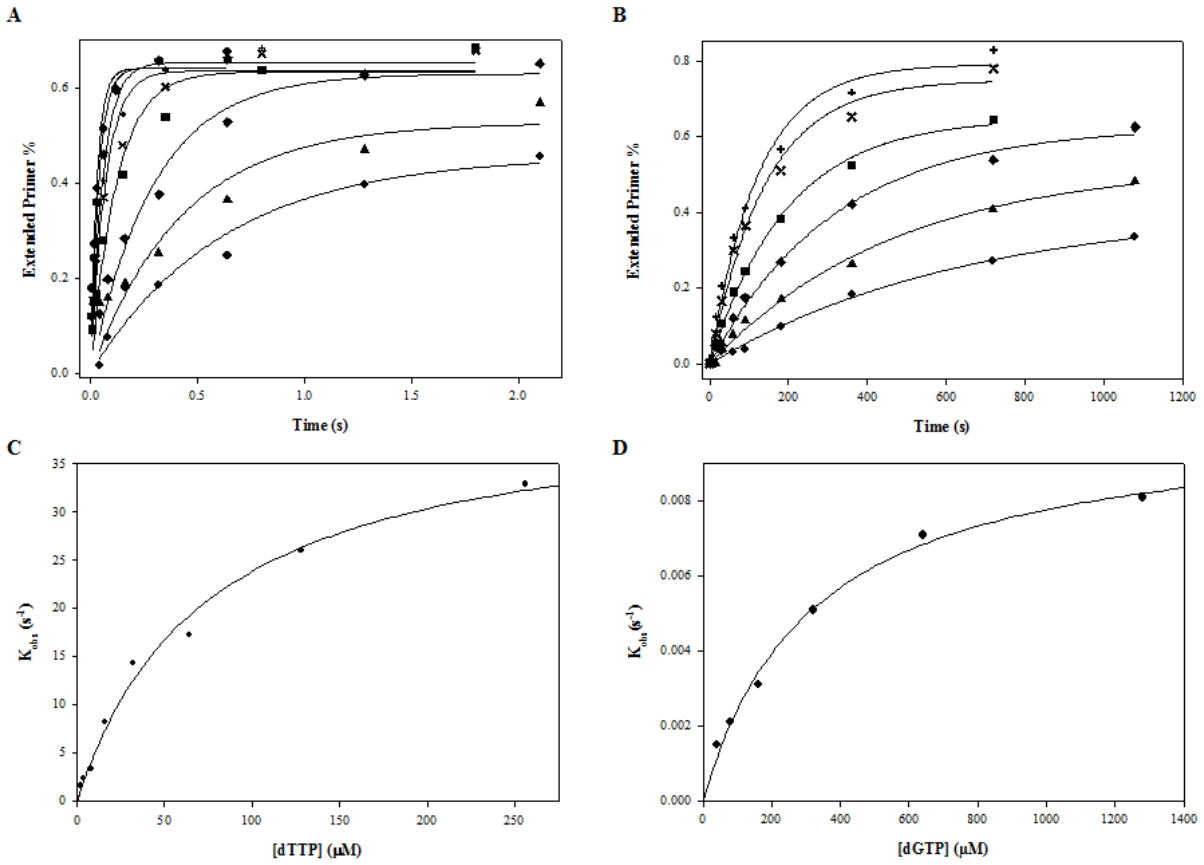

Figure 3. Kinetic assays for the deducing of apPOL fidelity. (A) Single exponential fits of Figure 2 dTTP across A data. ● 2 μM dATP, k_{obs} (1.6±0.4 s^{-1}) ▲ 4 μM dATP, k_{obs} (2.3±0.7 s^{-1}) ♦ 8 μM dATP, k_{obs} (3.4±0.5 s^{-1}) ■ 16 μM dATP, k_{obs} (8.2±1.5 s^{-1}) **x** 32 μM dATP, k_{obs} (14.3±3.0 s^{-1}) + 64 μM dATP, k_{obs} (17.2±2.5 s^{-1}) ► 128 μM dATP, k_{obs} (26.0±3.0 s^{-1}) ◄ 256 μM dATP, k_{obs} (32.9±5.1 s^{-1}) (B) Single Exponential fits of Figure 2 dGTP across A data. ● 40 μM dGTP, k_{obs} (0.0015±0.0003 s^{-1}) ▲ 80 μM dGTP, k_{obs} (0.0021±0.0002 s^{-1}) ♦ 160 μM dGTP, k_{obs} (0.0031±0.0003 s^{-1}) ■ 320 μM dGTP, k_{obs} (0.0051±0.0003 s^{-1}) **x** 640 μM dGTP, k_{obs} (0.0071±0.0007 s^{-1}) + 1280 μM dGTP, k_{obs} (0.0081±0.0008 s^{-1}) (C) Michaelis-Menten fit of k_{obs} (correct incorporation) values plotted as a function of dATP concentrations with k_{pol} (41.6±2.6 s^{-1}) and $K_{d,app}$ (74.3±11.3). (D) Michaelis-Menten fit of k_{obs} (incorrect incorporation) values plotted as a function of dGTP concentrations with k_{pol} (0.01027±0.0005 s^{-1}) and $K_{d,app}$ (320.0±42.1).

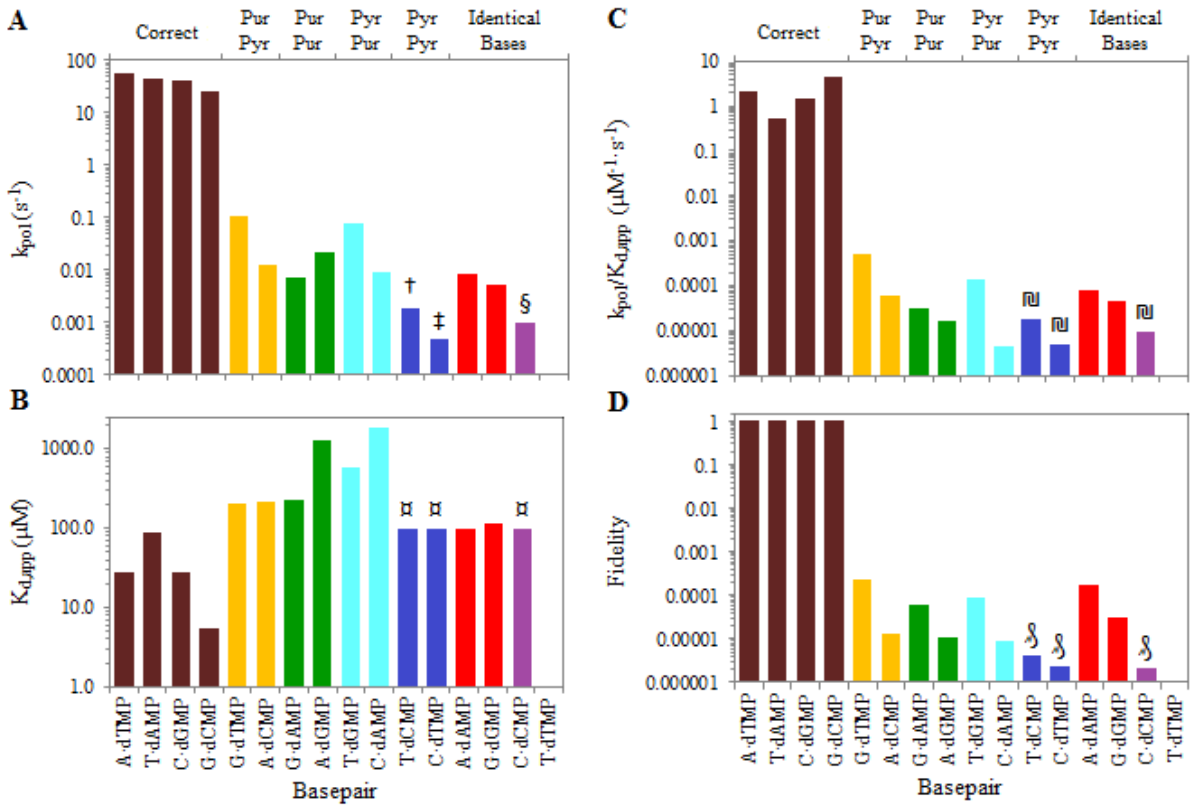

Figure 4. Fidelity of single-nucleotide insertion by apPOL^{exo-}. (A) Variation in the maximum rate of polymerization, k_{pol} (s^{-1}) † k_{pol} estimate bases on max k_{obs} value at 320 μM dTTP, ‡ k_{pol} estimate bases on max k_{obs} value at 12 mM dCTP, § k_{pol} estimate bases on max k_{obs} value at 4 mM dCTP. (B) Variation in the dissociation constant, $K_{d,app}$ (μM) ¤ $K_{d,app}$ set at 100 μM (estimate). (C) Variation in the specificity constant, ($\mu M^{-1} \cdot s^{-1}$) ≈ Estimated lower bound for selectivity. (D) Variation in Fidelity, $((k_{pol}/K_{d,app})_{incorrect}/(k_{pol}/K_{d,app})_{correct})$ § Estimated lower bound for fidelity.

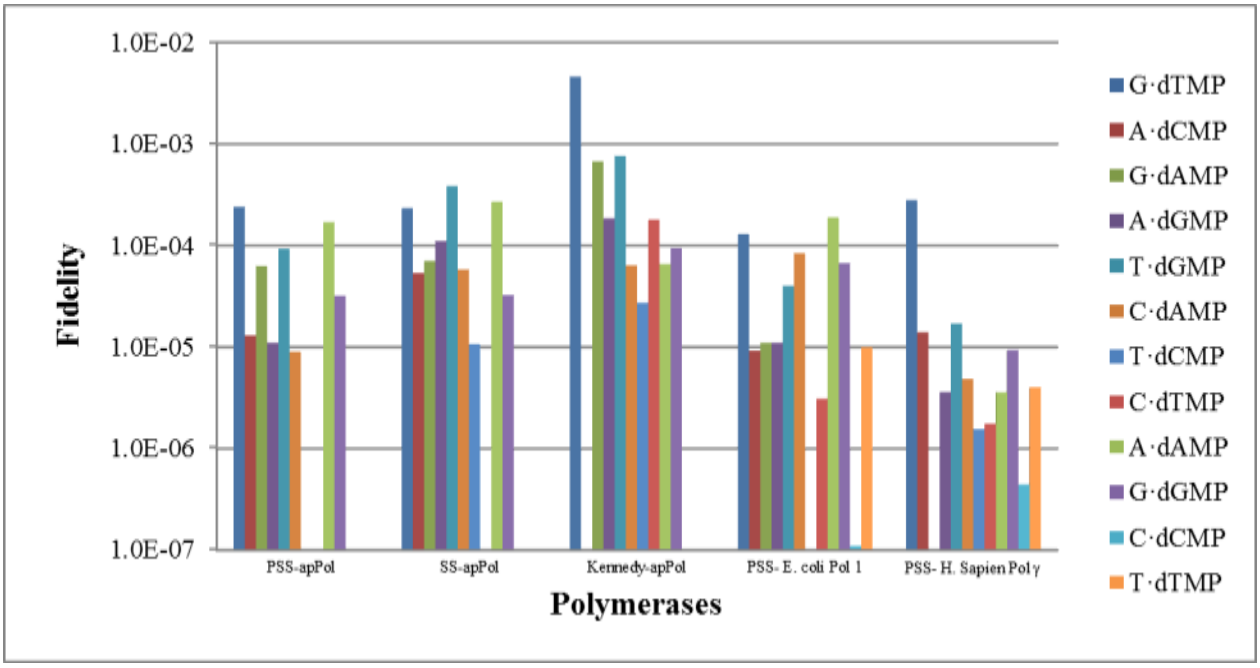

Figure 5. Comparisons of fidelity among selection of family A polymerases. (From the Left) Pre-steady State apPOL^{exo-} fidelity reported here. Steady State apPOL^{exo-} fidelity reported **(Wingert, et al. 2013)**, Steady state apPOL^{exo-} fidelity **(Kennedy, et al. 2011)**, Pre-steady state fidelity *E. Coli* Pol I **(Bertram, et al. 2010)**, Pre-steady State *H. sapiens* Pol γ **(Lee and Johnson 2006)**.

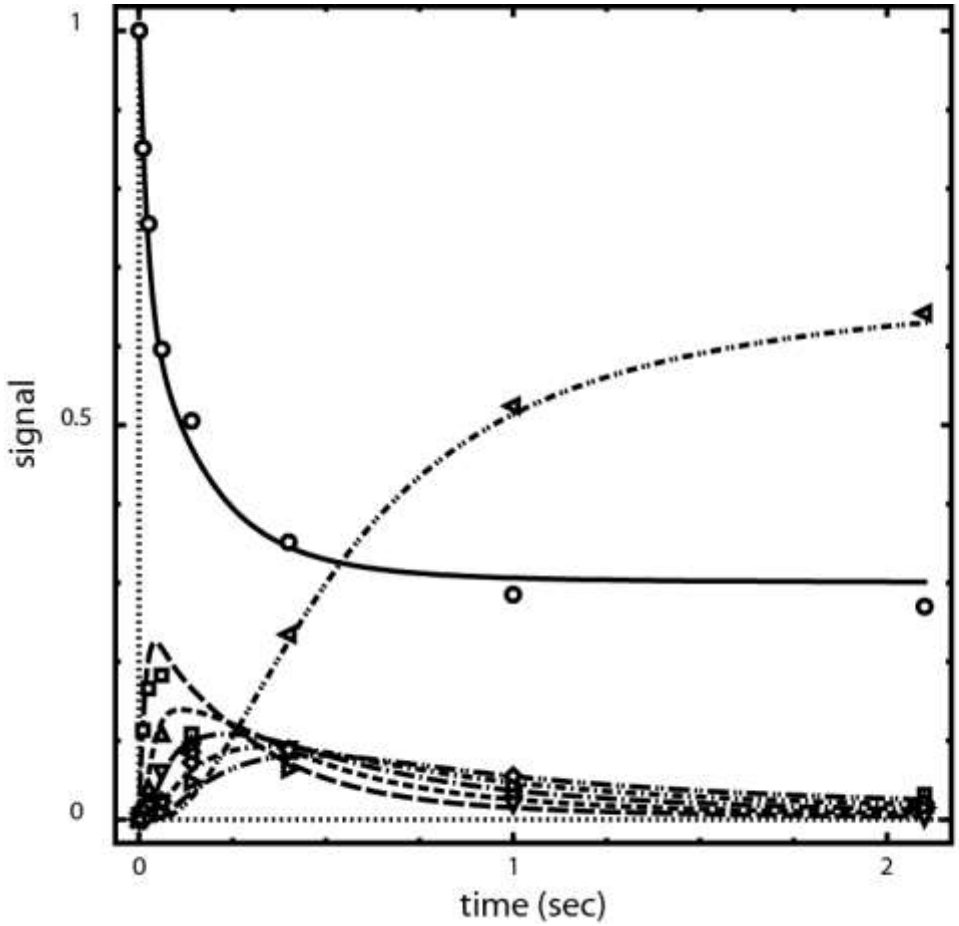

Figure 6. Multiple Nucleotide Incorporations. (o) % Primer-template, (□) % n+1 product, (up triangle) % n+2 product, (down triangle) % n+3 product, (diamond) % n+4 product, (right triangle) % n+5 product, (left triangle) % n+6 product.

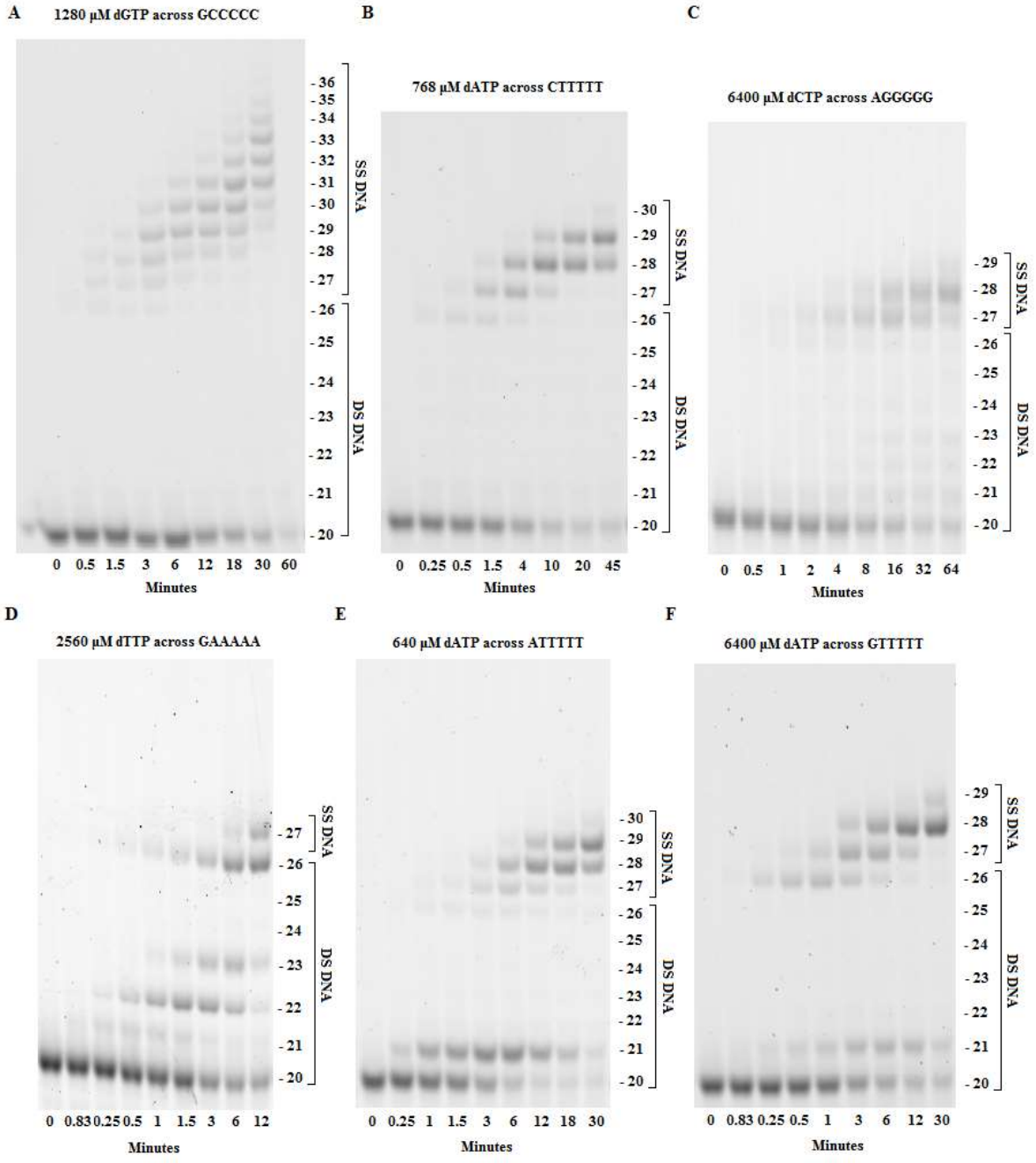

Figure 7. Mismatch extension stalling and non-template single strand extension. (A) dGTP non-template single strand extension. (B) dATP non-template single strand extension (C) dCTP non-template single strand extension (D) dTTP non-template single strand extension, n-2,n+3 stalling for T:dGMP mismatch (E) n+1 stalling of A:dAMP mismatch (F) n+1 stalling of A:dGMP mismatch

Supplemental Material

Dynafit Script—

```
[task]

task = fit
data = progress
model = chem ?

[mechanism]

A --> A : k10
ED + NTP <==> EDNTP  : k1 k2
EDNTP --> EDD + PP : kpol
ED <==> E + D : koff1 kon1
EDNTP <==> E + NTP + D : koff1 kon1

EDD + NTP <==> EDDNTP  : k1 k2
EDDNTP --> EDDD + PP : kpol
EDD <==> E + DD : koff1 kon1
EDDNTP <==> E + NTP + DD : koff1 kon1

EDDD + NTP <==> EDDDNTP  : k1 k2
EDDDNTP --> EDDDD + PP : kpol
EDDD <==> E + DDD : koff1 kon1
EDDDNTP <==> E + NTP + DDD : koff1 kon1

EDDDD + NTP <==> EDDDDNTP  : k1 k2
EDDDDNTP --> EDDDDD + PP : kpol
EDDDD <==> E + DDDD : koff1 kon1
EDDDDNTP <==> E + NTP + DDDD : koff1 kon1

EDDDDD + NTP <==> EDDDDDNTP  : k1 k2
EDDDDDNTP --> EDDDDDD + PP : kpol
EDDDDD <==> E + DDDDD : koff1 kon1
EDDDDDNTP <==> E + NTP + DDDDD : koff1 kon1

EDDDDDD + NTP <==> EDDDDDDNTP  : k1 k2
EDDDDDDNTP --> EDDDDDDD + PP : kpol
EDDDDDD <==> E + DDDDDD : koff1 kon1
EDDDDDDNTP <==> E + NTP + DDDDDD : koff1 kon1

[constants]
```

```
k1 = 50
k2 = 1
kpol = 10?
k10 = 1
koff1 = 6?
kon1 = .1?

[concentrations]
ED = .15?
A = .3
NTP = 200
E = 0.25

[progress]

directory ./fitting/polymerase/data
extension txt

file n
response ED = 5,EDNTP = 5, D = 5, A = 1

file n1
response EDDNTP = 5, EDD = 5,  DD = 5
file n2
response EDDDNTP = 5, EDDD = 5,  DDD = 5
file n3
response EDDDDNTP = 5, EDDDD = 5,  DDDD = 5
file n4
response EDDDDDNTP = 5, EDDDDD = 5,  DDDDD = 5
file n5
response EDDDDDDNTP = 5, EDDDDDD = 5,  DDDDDD = 5
file n6
response  EDDDDDDD = 5,

[output]
directory ./fitting/polymerase/output

[task]
task = fit
data = progress
model = trans ?

[mechanism]

A --> A : k10
ED + NTP <==> EDNTP  : k1 k2
```

EDNTP --> EDD* + PP : kpol
EDD* ---> EDD : kt
ED <==> E + D : koff1 kon1
EDNTP <==> E + NTP + D : koff1 kon1

EDD + NTP <==> EDDNTP : k1 k2
EDDNTP --> EDDD* + PP : kpol
EDDD* ---> EDDD : kt
EDD <==> E + DD : koff1 kon1
EDDNTP <==> E + NTP + DD : koff1 kon1

EDDD + NTP <==> EDDDNTP : k1 k2
EDDDNTP --> EDDDD* + PP : kpol
EDDDD* ---> EDDDD : kt
EDDD <==> E + DDD : koff1 kon1
EDDDNTP <==> E + NTP + DDD : koff1 kon1

EDDDD + NTP <==> EDDDDNTP : k1 k2
EDDDDNTP --> EDDDDD* + PP : kpol
EDDDDD* ---> EDDDDD : kt
EDDDD <==> E + DDDD : koff1 kon1
EDDDDNTP <==> E + NTP + DDDD : koff1 kon1

EDDDDD + NTP <==> EDDDDDNTP : k1 k2
EDDDDDNTP --> EDDDDDD* + PP : kpol
EDDDDDD* ---> EDDDDDD : kt
EDDDDD <==> E + DDDDD : koff1 kon1
EDDDDDNTP <==> E + NTP + DDDDD : koff1 kon1

EDDDDDD + NTP <==> EDDDDDDNTP : k1 k2
EDDDDDDNTP --> EDDDDDDD* + PP : kpol
EDDDDDDD* ---> EDDDDDDD : kt
EDDDDDD <==> E + DDDDDD : koff1 kon1
EDDDDDDNTP <==> E + NTP + DDDDDD : koff1 kon1

[constants]

k1 = 50
k2 = 1
kpol = 10?
kt = 10?
koff1 = 6?
kon1 = .1?
k10 = 1

[concentrations]

```
ED = .15?
A = .3
NTP = 200
E = 0.25

[progress]
directory ./fitting/polymerase/data
extension txt

file n
response ED = 5,EDNTP = 5, D = 5, A = 1
file n1
response EDDNTP = 5, EDD = 5,  DD = 5, EDD* = 5
file n2
response EDDDNTP = 5, EDDD = 5,  DDD = 5, EDDD* = 5
file n3
response EDDDDNTP = 5, EDDDD = 5,  DDDD = 5, EDDDD* = 5
file n4
response EDDDDDNTP = 5, EDDDDD = 5,  DDDDD = 5, EDDDDD* = 5
file n5
response EDDDDDDNTP = 5, EDDDDDD = 5,  DDDDDD = 5, EDDDDDD* = 5
file n6
response  EDDDDDDD = 5, EDDDDDDD* = 5

[output]
directory ./fitting/polymerase/output

[end]
```

CHAPTER III: CONCLUSIONS

Summary

In the course of this investigation it was our goal to accurately chart the substitution error map for the apPOL using pre-steady state kinetics. Our group has previously performed steady state kinetics for the apPOL; however, in terms of DNA polymerase fidelity, conclusions based on steady state kinetics have been questioned on theoretical grounds, primarily because the the slow dissociation of the DNA-polymerase complex may not reflect differences in nucleotide selectivity. Therefore, one of our goals was to confirm and extend our earlier steady state analysis (Wingert et al., 2013) with pre-steady state fidelity measurements. We were able to demonstrate the presence of a biphasic time course, confirming our pre-steady state kinetics represented single turnover conditions. We found that, for the most part, each study has indicated that apPOL has error rates typical of high fidelity A family polymerases (Kunkel, 2004). Additionally, it was our goal to expand and correct the work done previously by Kennedy et al. with a smaller apPOL fragment (Kennedy et al., 2011). We found that there was between a 500- and 8000-fold decrease in k_{pol} for incorrect incorporation compared to correct. We similarly found that $K_{d,app}$ values are inflated, but not to the same degree with a range of 1.1 to 46 fold increase for incorrect versus correct incorporation. The apPOL exhibits error in the 10^4 to the 10^6 range. The lowest calculated fidelity misincorporations were G:dTMP and A:dAMP and the highest fidelity that were able to observe were for A:dGMP and C:dAMP. Even greater estimated fidelities were found for the pyrimidine-pyrimidine mismatches (C:dCMP, T:dTMP, C:dTMP, and T:dCMP), which appear to be strongly disfavored as we had difficulty accurately quantitating their extremely slow rates. Other than slow pyrimidine-pyrimidine

misincorporations, we did not observe any obvious trends in fidelities (e.g. purine-purine mismatches were not favored or disfavored as compared to purine-pyrimidine).

We also determined that the apPOL has an average processivity of approximately 2, i.e., for every correct incorporation there is a 75% chance that the next nucleotide will be incorporated. Additionally, it appears that the rate of catalysis or a conformational change preceding is the rate limiting step for sequential polymerization as inclusion of a translocation step was not necessary to explain the product time courses.

The analysis of our misincorporation experiments revealed a high degree of stalling in the primer extension reaction after the initial turnover for T:dGMP, A:dAMP, and A:dGMP. This stalling may indicate that severe distortions of the active site following misincorporation significantly slows mismatch extension, which could increase the excision of the mismatch by the 3'-5' exonuclease activity of the wild-type apPOL. It was also surprising to observe non-template nucleotide additions to the primer strand by the apPOL^{exo-} beyond the end of the template with a preference for dGTP over dATP. More work will need to be done to see if this has any physiological relevance. The apPOL represents a potential target for drug discovery, and this report provides the essential data needed for further work on the polymerase. The impact of small molecules or nucleotide analogs on polymerase fidelity for drug screening will depend on the accurate fidelity data presented here for comparison.

Future Directions

We plan to continue expand the scope of the experiments reported here and explore some of the questions that arose during the course of these experiments. We are interested in

understanding the micro rate constants governing each step of nucleotide incorporation. The rate-limiting step for many polymerases is either the catalytic or a conformational change preceding catalysis (Johnson, 2010; Tsai and Johnson, 2006). To do this we plan to explore the dTTP(αS) elemental effect on the k_{pol} rate of incorporation. It has been shown in other polymerases, that the making or breaking of a phosphate bond will decrease the rate of catalysis (phosphothiote elemental effect). If the rate limiting step is the rate of catalysis then this effect will be observed, if not it suggests that the conformational change preceding catalysis is the rate limiting step (Fiala and Suo, 2004; Patel et al., 1991). We plan to determine the K_d-DNA through active site titration using various concentrations of DNA through a similar to a process described for the burst kinetics reported above. Additionally, this would provide us with a better calculation of the percentage active enzyme. We also plan to use a pulse-quench experiment to calculate the k_{off} of the DNA-enzyme complex (Fiala and Suo, 2004).

In the course of our pre-steady state experiments we noticed non-template nucleotide additions to the primer strand by the apPOL^{exo-} beyond the end of the template. These additions were done under "running-start" conditions where the polymerase rapidly incorporated nucleotides across the template before beginning non-template addition. We are interested to see whether or not the polymerase will catalyze addition to blunt-end DNA and single-stranded DNA, as it appears the apPOL extended dGTP well past the point where it would have been still bound to double-stranded portion of the DNA substrate. Non-template terminal transferase activity is unique in A family polymerases, but has been found to have a role in non-homologous end joining for X-family polymerases (Andrade et al., 2009). The physiological relevance, if any, of terminal transferase activity in apPOL is unclear.

References

Andrade, P., Martín, M.J., Juárez, R., López de Saro, F., and Blanco, L. (2009). Limited terminal transferase in human DNA polymerase mu defines the required balance between accuracy and efficiency in NHEJ. Proc. Natl. Acad. Sci. U. S. A. *106*, 16203–16208.

Fiala, K.A., and Suo, Z. (2004). Mechanism of DNA Polymerization Catalyzed by *Sulfolobus solfataricus* P2 DNA Polymerase IV †. Biochemistry (Mosc.) *43*, 2116–2125.

Johnson, K.A. (2010). The kinetic and chemical mechanism of high-fidelity DNA polymerases. Biochim. Biophys. Acta Bba - Proteins Proteomics *1804*, 1041–1048.

Kennedy, S.R., Chen, C.-Y., Schmitt, M.W., Bower, C.N., and Loeb, L.A. (2011). The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite Plasmodium falciparum. J. Mol. Biol. *410*, 27–38.

Kunkel, T.A. (2004). DNA Replication Fidelity. J. Biol. Chem. *279*, 16895–16898.

Patel, S.S., Wong, I., and Johnson, K.A. (1991). Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. Biochemistry (Mosc.) *30*, 511–525.

Tsai, Y.-C., and Johnson, K.A. (2006). A New Paradigm for DNA Polymerase Specificity †. Biochemistry (Mosc.) *45*, 9675–9687.

Wingert, B., Parrott, E., and Nelson, S.W. (2013). Misincorporation, mismatch extension, and proofreading activity of the Plasmodium falciparum apicoplast DNA polymerase. Biochemistry (Mosc.).

ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Scott Nelson, for allowing me to be a part of his laboratory for the last four and half years. I have appreciated his constant support, encouragement, patience, and long-suffering in during those years and during this research and writing process. He has been one of my favorite professors and his ability to clearly convey an idea or concept is second to none. I would also like to thank Dr. Timothy Herdendorf for all the assistance he has given to me regarding experimental procedures and techniques. I am deeply indebted to his advice on the scientific process and the frustrations that come with it. In addition, I would like to thank Bentley Wingert for his advice and help as we explored the kinetics of the apPOL. I would also like to thank the other current and former members of the lab; Dustin Albrecht, Tasida Barfoot, Morgan Miller, Haley Streff, and Bryanna Behning for assisting in the day-to-day operations and conversation that allow the lab to operate easily and enjoyably.

Last but not least; I would like to thank my parents, Evan and Kim Parrott, and my sister, Leah Parrott, for their listening ear, even when they had no idea what I was talking about; all my friends who dealt with my late nights and odd hours, especially Colin Judd, Drew Brown, Scott Henry, and Andrew Pringnitz; and finally my fiancé, Chelsea Anderson, who has listened to, encouraged, prodded, and stuck with me through everything. I am deeply indebted to all of them.