doi:10.1016/j.jmb.2011.04.071 J. Mol. Biol. (2011) xx, xxx–xxx

Contents lists available at www.sciencedirect.com
Journal of Molecular Biology
journal homepage: http://ees.elsevier.com.jmb

The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*

Scott R. Kennedy[1], Cheng-Yao Chen[1], Michael W. Schmitt[1], Cole N. Bower[1] and Lawrence A. Loeb[1,2*]

[1]*Department of Pathology, University of Washington at Seattle, Seattle, WA 98195, USA*
[2]*Department of Biochemistry, University of Washington at Seattle, Seattle, WA 98195, USA*

Received 16 March 2011;
received in revised form
15 April 2011;
accepted 27 April 2011

Edited by J. Karn

Keywords:
replication;
apicoplexan;
plDNA;
Pom1

Plasmodium falciparum, the major causative agent of human malaria, contains three separate genomes. The apicoplast (an intracellular organelle) contains an ~35-kb circular DNA genome of unusually high A/T content (>86%) that is replicated by the nuclear-encoded replication complex Pfprex. Herein, we have expressed and purified the DNA polymerase domain of Pfprex [KPom1 (*K*lenow-like *p*olymerase *o*f *m*alaria *1*)] and measured its fidelity using a *LacZ*-based forward mutation assay. In addition, we analyzed the kinetic parameters for the incorporation of both complementary and noncomplementary nucleotides using Kpom1 lacking 3′→5′ exonucleolytic activity. KPom1 exhibits a strongly biased mutational spectrum in which T→C is the most frequent single-base substitution and differs significantly from the closely related *Escherichia coli* DNA polymerase I. Using *E. coli* harboring a temperature-sensitive polymerase I allele, we established that KPom1 can complement the growth-defective phenotype at an elevated temperature. We propose that the error bias of KPom1 may be exploited in the complementation assay to identify nucleoside analogs that mimic this base-mispairing and preferentially inhibit apicoplast DNA replication.

Introduction

Plasmodium falciparum, the infectious agent associated with most cases of malaria, is responsible for an estimated 5 million deaths annually throughout the world.[1] This infectious parasite contains a nuclear genome and a mitochondrial genome, as well as a third unique genome that is encapsulated in an organelle termed the apicoplast. The apicoplast, which is thought to be derived from the secondary endocytosis of photosynthetic algae, is involved in a variety of biosynthetic pathways and is required for parasite survival. The apicoplast contains its own plastid-derived ~35-kb closed circular double-stranded DNA genome (plDNA) that is replicated by a bidirectional θ mechanism and segregated into daughter cells.[2,3] The genome encodes several subunits of rRNA and the accompanying ribosomal proteins, 25 species of tRNA, an RNA polymerase, and several open reading frames coding for chaperones, as well as other proteins of unknown function.[4]

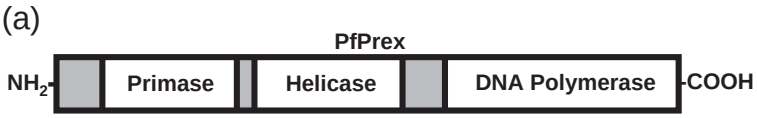
The biochemical and cellular processes involved in plDNA replication are poorly understood;

*Corresponding author. Department of Pathology, University of Washington at Seattle, HSB K-072, Box 357705, 1959 Pacific Street NE, Seattle, WA 98125-7705, USA. E-mail address: laloeb@u.washington.edu.

Abbreviations used: Pol I, polymerase I; EDTA, ethylenediaminetetraacetic acid.

0022-2836/$ - see front matter © 2011 Published by Elsevier Ltd.

Q1



(b)

	Motif 1	Motif 2
Pfprex	1679-INVNSQQQVLKAL	1757-NQLKTFSGRFSSEKPNLQ
EcPol I	578-FNLSSTKQLQTIL	660-HQAVTATGRLSSTDPNLQ
hPol ν	476-FLITSNNQLREIL	563-NQTGTVTGRLSAKHPNIQ

	Motif A	Motif B
Pfprex	1795-IIADFKQIELK	1847-RHIAKAINFGLIY
EcPol I	702-VSADYSQIELR	754-RRSAKAINFGLIY
hPol ν	620-LAADFSQIELR	674-REQTKKVVYAVVY

	Motif 6	Motif C
Pfprex	1930-ALNYPVQGTCADILK	1962-ILCVHDEIIIE
EcPol I	843-AINAPMQGTAADIIK	877-IMQVHDELVFE
hPol ν	763-AVNFVVQGSAADLCK	798-VAQIHDELLFE

Fig. 1. (a) The domain organization of Pfprex. (b) Sequence comparison of the evolutionarily conserved DNA polymerase motifs of the *P. falciparum* Pfprex polymerase domain, *E. coli* DNA Pol I, and human Pol *ν*.

however, several proteins that are involved in DNA metabolic processes are encoded by the nuclear genome, synthesized in the cytoplasm, and transported to the apicoplast. These include a bacteria-like gyrase, a DNA ligase, and several unclassified open reading frames that are homologous to DNA repair enzymes.[5,6] Of these, the apicoplast gyrase is a known target for inhibition by the drug ciprofloxin, a major therapeutic agent for the treatment of malaria, suggesting that other enzymes involved in plDNA replication and maintenance may be useful drug targets.[7–10]

Apicoplast DNA replication is catalyzed by the nuclear-encoded apicoplast-targeted polyprotein Pfprex, which is a large (2016 aa) multifunctional peptide that contains three distinct domains that exhibit DNA primase, DNA helicase, and DNA polymerase activities, respectively[11] (Fig. 1a). The helicase and primase segments are homologous to T7 bacteriophage helicase and primase proteins.[11] The third domain contains a DNA polymerase that is evolutionarily related to prokaryotic DNA polymerase I (Pol I), an A-family polymerase, based on sequence homology. Similar to DNA polymerase γ, which is localized to the mitochondria, Pfprex is believed to be the only DNA-synthesizing enzyme in the apicoplast and is thought to be involved in DNA replication, repair, and recombination.

The three genomes present in *P. falciparum* (i.e., nuclear, mitochondrial, and plastidial) are among the most A/T-rich yet sequenced genomes, with plDNA being the richest (86.9% A/T).[4] Given the highly biased sequence composition of the apicoplast genome, we asked if the plDNA polymerase preferentially incorporated dATP and/or dTTP and thus has a role in the maintenance of the A/T-rich genome. To this end, we have expressed and purified the DNA polymerase domain of Pfprex [henceforth referred to as KPom1 (*K*lenow-like *p*olymerase *o*f *m*alaria *1*)] and have determined the frequencies of misincorporations and the effects of neighboring nucleotides using the M13mp2 forward mutation assay for KPom1 with and without a 3′→5′ exonucleolytic activity. In addition, we also characterized the kinetics of incorporation of complementary and noncomplementary nucleotides. Interestingly, we find that KPom1 exhibits a strongly biased error spectrum, with the T→C single-base substitution being the most frequent, and thus does not account for the maintenance of the A/T-rich plastid genome. Even though the catalytic site motifs are highly conserved between *Escherichia coli* Pol I and KPom1, the spectrum of misincorporation by KPom1 is markedly different from that by *E. coli* Pol I. This finding suggests that residues outside of active-site motifs may influence the fidelity of these two enzymes. Despite these differences, we established that KPom1 is able to substitute for *E. coli* Pol I *in vivo*. We suggest that the error bias of KPom1 may be exploited in the complementation assay to identify nucleoside analogs that mimic base-mispairing and preferentially inhibit plDNA replication.

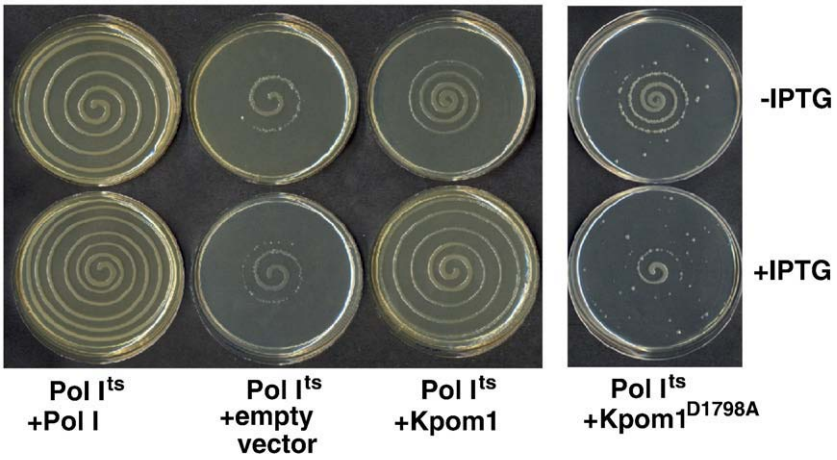
Results

KPom1 can substitute for DNA polymerase Pol I in *E. coli*

KPom1 is a member of the A-family of DNA polymerases. There is a high degree of conservation of amino acid sequences at the catalytic sites of DNA polymerases; sequence alignment of KPom1 with *E. coli* Pol I shows that all the required motifs for polymerase activity in Pol I are present in KPom1 and are highly conserved (Fig. 1b). We have previously shown that both mammalian Pol β[12]

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

(a)



(b)

Pol I^{ts}	+Pol I	+empty vector	+KPom1	+KPom1^{D1798A}
-IPTG (% relative growth)	100%	<0.001%	29%	<0.001%
+IPTG (% relative growth)	111%	<0.001%	60%	<0.001%

Fig. 2. Functional complementation of *E. coli* DNA Pol I by *P. falciparum* KPom1 *in vivo*. (a) *E. coli* JS200 cells were transformed with an empty pHSG576 vector or with a vector harboring Pol I, KPom1, or the KPom1^{D1798A} gene. (b) The complementation efficiency of Pol I^{ts} cells by exogenously expressed Pol I, KPom1, or KPom1^{D1798A} was quantified as described by Camps *et al.*[14] The complementation efficiency of Pol I^{ts} cells by KPom1 or KPom1^{D1798A} was normalized to the conditions with Pol I in the absence of IPTG, which was set as 100%. Assays were performed in triplicate.

and human immunodeficiency virus reverse transcriptase[13] can complement *E. coli* harboring a temperature-sensitive mutation in Pol I.[12,13] To determine whether KPom1 can substitute for *E. coli* Pol I, we inserted KPom1 into an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible vector and transformed it into *E. coli* that expresses a temperature-sensitive variant of Pol I, Pol I^{ts}. At 37 °C, Pol I^{ts} is inactivated, and the strain is dependent on an exogenously supplied polymerase for colony formation. Figure 2a and b demonstrates that KPom1 almost fully restores wild-type growth when induced with IPTG. In the absence of IPTG, the KPom1-containing plasmid fails to form colonies at the nonpermissive temperature. In addition, essentially no growth is observed in vector-only control or when cells are transformed with a copy of KPom1 harboring an active-site mutation (D1798A) that inactivates the polymerase (Fig. 2a and b).

KPom1 is error prone *in vivo*

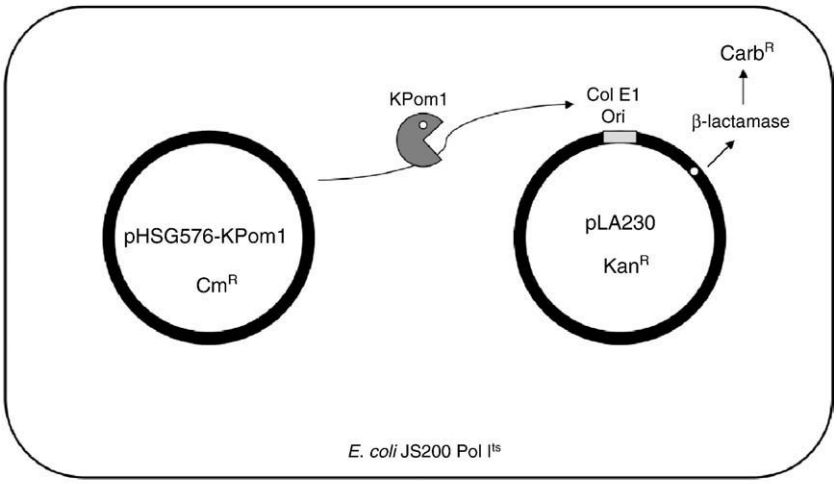
We took advantage of the ability of KPom1 to substitute for the endogenous Pol I activity of *E. coli* in order to determine its *in vivo* fidelity. We utilized a two-plasmid system detailed by Camps *et al.*[14] that uses a bacterial host, Pol I^{ts} (Fig. 3a).[12] In this system, one plasmid encodes the polymerase of interest (Klenow fragment of Pol I or KPom1 and its exonuclease-deficient derivatives), while the other contains the β-lactamase gene with an ochre (TAA) mutation. Any mutation that results in a reversion of the ochre codon will lead to active β-lactamase activity and carbenicillin resistance. Thus, the frequency of bacteria rendered carbenicillin resistant reflects the frequency of ochre codon mutagenesis in the target plasmid.

Figure 3b shows the results for the *in vivo* reversion of mutations in β-lactamase by Klenow, Klenow^{exo-}, KPom1, and KPom1^{exo-}. The reversion frequencies for Klenow and Klenow^{exo-} are consistent with previous reports on β-lactamase reversion with this system.[14] Specifically, the reversion frequency with Klenow^{exo-} is 4.5-fold greater than that observed with exonuclease-proficient Klenow and is consistent with previous reports.[14,15] In JS200 strains expressing the KPom1 and KPom1^{exo-} enzymes, the reversion frequency is much higher than that of the corresponding controls. The KPom1 enzyme exhibits a 9-fold-higher reversion frequency than the Klenow fragment of *E. coli* Pol I. Surprisingly, KPom1^{exo-} exhibits an increase in the *in vivo* reversion frequency 793-fold relative to *E. coli* Pol I^{wt} and 41-fold higher than KPom1 (Fig. 3b), and suggests that the polymerase domain of KPom1 is very error prone and that the majority of polymerase mistakes are corrected by the exonuclease domain of the enzyme in the context of this system. Interestingly, the extremely high reversion frequency by KPom1^{exo-} occurs in the context of the A/T-rich TAA ochre codon, and the plDNA genome that this enzyme replicates is also A/T rich.

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Q1

(a)



(b)

	β-lactamase reversion frequency	Fold difference (relative to E. coli Pol I)
***E. coli* Pol I**	$2.9 \pm 0.7 \times 10^{-7}$	1
***E. coli* Pol I (exo$^-$)**	$1.3 \pm 0.6 \times 10^{-6}$	4.5
KPom1	$5.5 \pm 2.9 \times 10^{-6}$	9
KPom1 (exo$^-$)	$2.3 \pm 1.0 \times 10^{-4}$	793

Fig. 3. *In vivo* fidelity of KPom1. (a) Schematic diagram of the two-plasmid β-lactamase reversion assay. JS200 (Pol I^{ts}) cells were transformed with two plasmids. Plasmid pHSG576 is a low-copy-number plasmid and carries the *polA* gene under the control of the *tac* promoter. The pLA230 plasmid carries the β-lactamase gene placed in close proximity downstream of a pUC19 (ColEl-type) origin of replication. The polymerase of interest is expressed from pHSG576 and initiates replication of pLA230. If the polymerase makes a misinsertion when copying the ochre codon, it will lead to a reversion of a functional β-lactamase enzyme. (b) Reversion frequencies for the β-lactamase reversion assay. The fold increase in the reversion frequency is relative to *E. coli* Pol I.

Biochemical characterization of KPom1

In order to determine if KPom1 exhibits a lower fidelity for nucleotide misinsertions opposite template dA or dT relative to template dC or dG, we used the purified enzyme to measure the *in vitro* fidelity and incorporation kinetics for all 16 possible base pairs. We were unable to express the apicoplast DNA polymerase in *E. coli* or yeast using different high-expression vectors. We reasoned that this lack of high expression could be due to the disparity in the codon usage bias between *P. falciparum* and *E. coli*. Therefore, we obtained a chemically synthesized gene, optimized for *E. coli* and coding for amino acids 1431–2016 of Pfprex (Supplementary Fig. 1), which corresponds to the Pol I-like DNA polymerase domain (KPom1). The expressed N-terminal maltose-binding domain fusion protein was purified to near homogeneity using a heparin column, followed by amylose resin (Supplementary Fig. 2). After cleavage and removal of the maltose-binding peptide, the purified DNA polymerase is highly active, with a specific activity of 77 pmol of dNMPs incorporated per minute per nanogram of protein for KPom1^{WT} and 53 pmol/min ng for KPom1^{exo-}. KPom1 exhibits maximal activity at ~10 mM Mg^{2+} (Supplementary Fig. 3a) and pH 9.0 (Supplementary Fig. 3b). Polymerase activity increases 2-fold from 18 °C to ~40 °C and then rapidly declines, consistent with the idea that enzyme activity has evolved for DNA replication in warm-blooded animals (Supplementary Fig. 3c).

Fidelity of DNA synthesis by KPom1 DNA polymerase

The fidelity of DNA synthesis by the wild-type and exonuclease-deficient KPom1 DNA polymerases was determined using the M13mp2 forward mutation assay.[16] The substrate is double-stranded M13mp2 with a 407-nucleotide single-stranded gap in the *LacZ*α gene. A purified DNA polymerase is used to fill in the single-stranded segment *in vitro*, which is then transformed into *E. coli* and plated on a lawn of α-complementation cells in the presence of 5-bromo-4-chloro-3-indolyl-β,D-galactopyranoside and IPTG. Accurate synthesis by the polymerase results in the faithful replication of the *LacZ*α gene and the formation of dark-blue plaques, while polymerase errors yield light-blue or colorless plaques. The fidelity of the polymerase is determined from the ratio of mutant to wild-type plaques; the error spectrum of the polymerase is determined by sequencing the resulting plaques. This assay allows for the monitoring of a broad range of mutations, including all 12 single-nucleotide misinsertion mutations and small insertion–deletion mutations, as well as short duplications, additions,

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Q1

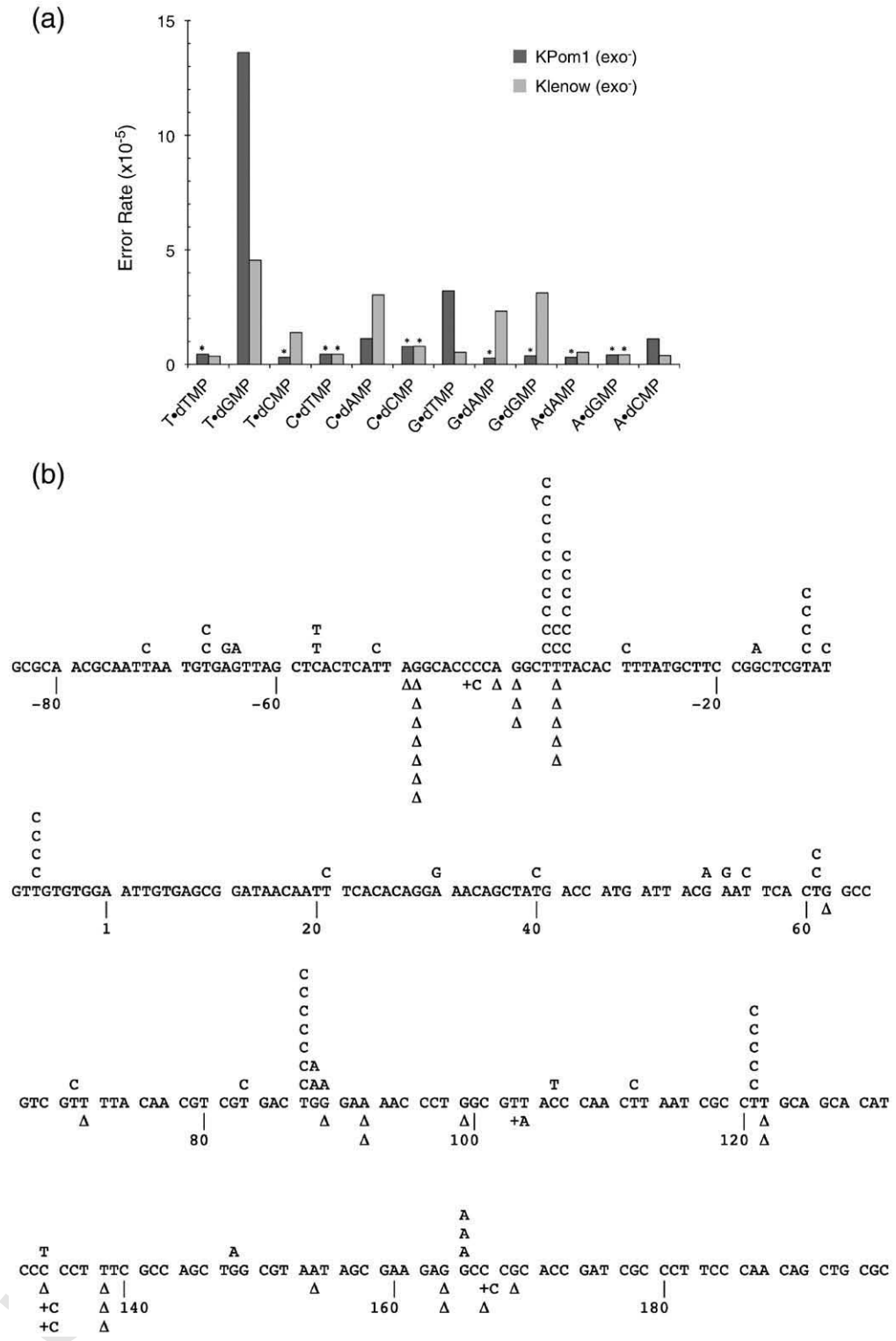


Fig. 4. Base substitution error rates of *E. coli* and *P. falciparum* KPom1 DNA polymerase in the *LacZ* forward mutation assay. (a) Error rates for all possible base–base mispairs are shown for KPom1^{exo-} (dark bars) and Klenow^{exo-} (light bars). An asterisk (*) indicates less-than or equal-to values due to the failure to identify any mutations of a particular class. Data for Klenow^{exo-} were adapted from Bebenek and Kunkel.[16] (b) Mutation spectrum of Kpom1^{exo-}. The target sequence of the *LacZ* fragment is shown, with detected single-base substitutions indicated above the target sequence and with single-nucleotide deletions (Δ) and insertions (+) indicated below the target.

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Table 1. Mutation rates in the *LacZ* forward mutation assay

Mispair	Number of detectable sites	Wild type: Number detected	Wild type: Error rate ($\times 10^{-5}$)	Exo$^-$: Number detected	Exo$^-$: Error rate ($\times 10^{-5}$)
T·dTMP	16	0	≤0.3	0	≤0.4
T·dGMP	27	1	0.17	52	14
T·dCMP	23	0	≤0.2	0	≤0.3
C·dTMP	16	0	≤0.3	0	≤0.4
C·dAMP	25	2	0.36	4	1.1
C·dCMP	9	0	≤0.5	0	≤0.8
G·dTMP	22	0	≤0.2	10	3.2
G·dAMP	25	0	≤0.2	0	≤0.3
G·dGMP	19	0	≤0.2	0	≤0.4
A·dAMP	23	0	≤0.2	0	≤0.3
A·dGMP	17	0	≤0.3	0	≤0.4
A·dCMP	29	1	0.24	3	1.1
Base substitutions	125	4	0.14	69	3.9
+Insertions	199	0	≤0.02	5	0.2
−1 deletions	199	6	0.14	34	1.2
Large deletion[a]	Not defined	4	25	7	37
LacZ mutation frequency			1.3×10^{-3}		6.3×10^{-3}

[a] For large deletions, mutation frequency is reported instead of the error rate.

and rearrangements.[16] With the 'wild-type' KPom1, only 14 phenotypic colonies (i.e., light-blue or colorless plaques) were observed in the 10,769 observed colonies, resulting in a mutation frequency of 0.13%—less than 2-fold greater than the background of the assay (0.07%).[15,16] In contrast, with KPom1^{exo-}, 114 mutant plaques were detected in a total of 18,254 plaques, resulting in a mutation frequency of 0.63%. These mutation frequencies for both the wild type and the exonuclease-deficient KPom1 are similar to those observed for other high-fidelity A-family polymerases, including *E. coli* Pol I,[15] *Taq* DNA polymerase,[17] and human DNA polymerase γ.[18] The presence of the exonuclease domain imparts at least a 6-fold increase in fidelity, thus showing that it is able to correct most misincorporation events that the polymerase active site makes.

Table 2. Comparison of mutation rates between KPom1^{exo-} and other polymerases

DNA polymerase	Error rate for single-base deletion ($\times10^{-5}$)	Error rate for single-base substitution ($\times10^{-5}$)
Kpom1 (exo$^-$)[a]	1.2	3.9
E. coli Klenow (exo$^-$)[b]	0.6	2.5
Taq Pol (exo$^-$)[c]	0.6	1.7
Human Pol ν[d]	17	350
Human Pol θ[e]	140	240
Human Pol γ (exo$^-$) (+p140 and p55)[f]	0.8	4.1
Human Pol α[g]	2.8	7.5
Human Pol δ (exo$^-$)[h]	2.0	4.4
Yeast Pol ε (exo$^-$)[i]	5.6	24
Human Pol β[j]	14	23
Human Pol λ[j]	450	90
Human Pol η[k]	240	3500
Human Pol κ[l]	180	580

[a] Error rates were derived from this study.
[b] Single-base deletions were adapted from Minnick *et al.* (1996), whereas single-base substitutions were adapted from Bebenek *et al.*[15]
[c] Error rates were adapted from Eckert and Kunkel.[17]
[d] Error rates were adapted from Arana *et al.*[19]
[e] Error rates were adapted from Arana *et al.*[20]
[f] Single-base deletions were adapted from Longley *et al.*,[18] whereas single-base substitutions were reported in Table 2 of Arana *et al.*[19]
[g] Error rates are unpublished data reported in Table 2 of Arana *et al.*[20]
[h] Error rates were adapted from Schmitt *et al.*[21]
[i] Error rates were adapted from Shcherbakova *et al.* (2003).
[j] Error rates were adapted from Bebenek *et al.* (2003).
[k] Error rates were adapted from Matsuda *et al.* (2000, 2001).
[l] Error rates were adapted from Ohashi *et al.* (2000).

The error spectrum of KPom1 was determined by sequencing the observed *lacZ* mutants and by tabulating the types and sequence contexts of the mutations. The results for both KPom1^{wt} and KPom1^{exo-} were used to calculate the error rates for each type of mutation (Fig. 4, Table 1), and the error spectrum of KPom1^{exo-} was compared to that of other characterized polymerases (Table 2). For KPom1^{exo-}, 39 of the 115 sequenced verified mutants were single-base insertion or deletions; the error rates were 1.2×10^{-5} and 0.2×10^{-5}, respectively. Of the 34 detected −1 frameshift mutations found, 28 were in nucleotide repeats of two or more identical bases, consistent with a mechanism involving a template slippage event.[22]

For KPom1^{exo-}, 69 of the 115 phenotypic mutants consisted of single-base substitutions—a base substitution error rate of 3.9×10^{-5}. Even though the M13mp2 forward mutation assay can detect all of the 12 single-base substitutions, only a subset was observed. Of the 69 sequenced single-base substitutions, the majority (52/69 or 75%) were the result of a T→C transition (i.e., T:dGMP). The

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Q1

frequency of this transition mutation was 14×10^{-5}, three times greater than the frequency generated by the closely related *E. coli* Pol I^{exo-} (Fig. 4, Tables 1 and 2).[15] The second most common observed base substitution (10/69 or 14%) was the G→A (i.e., G:dTMP) transition, which occurs at a rate of 3.2×10^{-5}—6-fold higher than what is observed for *E. coli* Pol I^{exo-}.[15] Both C→T (C:dAMP) and A→G (A:dCMP) mutations occur at a rate of 1.1×10^{-5} and, combined, only account for 10% of the observed single-base substitutions. No other base substitutions were observed. This error spectrum deviates significantly from *E. coli* Pol I^{exo-}, even though these two enzymes share a significant sequence identity in their active-site motifs (Fig. 1b) and have a similar overall error rate. Specifically, KPom1^{exo-} catalyzes primarily T→C transitions, while *E. coli* Pol I^{exo-} misincorporations result predominantly in C→T transitions, G→T transversions, and G→C transversions.[15] In addition, a majority of T→C mutations (~65%) with KPom1^{exo-} DNA polymerase occurred in a sequence context in which the 5′-template base is either a C or a G. The mutation spectrum for KPom1^{exo-} is most similar to that of the A-family lesion bypass polymerase Pol *v* (see Discussion for details).

Steady-state kinetics of KPom1

Examining the fidelity of KPom1 using steady-state kinetics gives an indication of which step in the nucleotide discrimination process is rate limiting for the incorporation of specific deoxynucleoside triphosphates. For example, a higher K_m would suggest that nucleotide discrimination is due to an increase in substrate dissociation (increase in k_{off}), while a lower V_{max} would indicate an unfavorable geometry of the bound nucleotide in the active site.[15,23] We used a steady-state gel-based assay to determine the apparent kinetic parameters (V_{max} and K_m) for the incorporation of the correct or incorrect nucleotide across all 16 possible base pairings. These values were then used to calculate the fidelity of KPom1^{exo-}.[23]

KPom1^{exo-} efficiently incorporated dAMP, dCMP, dTMP, or dGMP across from their complementary template bases and discriminated against the incorrect nucleotide for all possible nucleotide mismatches (Table 3). The calculated values for the fidelity of nucleotide misinsertion range from 4.6×10^{-3} (T:dGMP) to 2.6×10^{-5} (A:dCMP) (Table 3). Values for some mismatches could not be calculated due to a lack of observable incorporation. The most prevalent misinsertions observed in the kinetic assay are the same as the mispairs with the lowest observed fidelity in the M13mp2 forward mutation assay (Table 1). The apparent values for K_m ranged from a 325-fold increase to a 3000-fold increase for incorrect *versus* correct incorporation, but did not vary substantially for the correct incorporation reactions. The mispairs that were the least frequent in the M13mp2 forward mutation assay also exhibited a relatively high value for K_m (>1000 μM), whereas base mispairs with the lowest fidelity tended to have much lower values of K_m. Specifically, the values for K_m were highest when the identity of the template base was dG or dC, or when base-pairing involved an incoming nucleotide with the same identity as the template base. Conversely, when the template base was dA or dT, the values of K_m were on the order of 2-fold to 3-fold lower than the values observed for dG or dC. This observation suggests that KPom1 uses a mechanism whereby nucleotide discrimination for certain base–base mispairs (namely template dG, template dC, and

Table 3. Fidelity of single-nucleotide insertion by KPom1^{exo-}

Base pair	K_m (μM)	V_{max} (fmol/min)	V_{max}/K_m (fmol/μM min)	Fidelity[a]
T·dAMP	1.6±0.7	12.5±0.7	7.8	1
T·dTMP	ND[b]	ND[b]	—	—
T·dGMP	521±54	18.8±0.6	0.0361	4.61×10^{-3}
T·dCMP	648±350	0.91±0.13	0.0014	1.80×10^{-4}
C·dGMP	0.36±0.09	8.2±0.4	22.8	1
C·dTMP	1300±450	0.81±0.09	0.0006	2.74×10^{-5}
C·dAMP	ND[b]	ND[b]	—	—
C·dCMP	ND[b]	ND[b]	—	—
G·dCMP	0.72±0.09	10.4±0.3	14.4	1
G·dTMP	290±120	3.2±0.6	0.0110	7.64×10^{-4}
G·dAMP	1450±650	3.9±0.7	0.0027	1.86×10^{-4}
G·dGMP	2200±620	3.0±0.4	0.0014	9.44×10^{-5}
A·dTMP	0.70±0.15	11.4±0.6	16.30	1
A·dAMP	1500±260	1.6±0.2	0.0011	6.55×10^{-5}
A·dGMP	264±117	2.9±0.2	0.0110	6.74×10^{-4}
A·dCMP	885±228	0.92±0.15	0.0010	6.38×10^{-5}

[a] Fidelity is defined as $(V_{max}/K_m)_{wrong}/(V_{max}/K_m)_{right}$.
[b] The reaction was too slow to be determined.

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Q1

pyr:pyr/pur:pur) involves poor binding (i.e., increased off-rate).

In contrast to the values of K_m, which exhibit an increase in magnitude for all mismatches relative to the correct base pair, the values for the apparent V_{max} ranged from a 13.7-fold decrease to a 1.5-fold increase for incorrect *versus* correct incorporation. The apparent V_{max} values for the correct incorporation reactions were very similar to one another, regardless of the template base identity, while the V_{max} values for most mismatch reactions varied in a range of about 4-fold, regardless of the identity of the template base. A notable exception is the T:dGMP mispair, which surprisingly exhibits a V_{max} higher than that of the correct base pair.

Taken together, it appears that, depending on the template base identity, KPom1 uses two different strategies for nucleotide discrimination. When the identity of the template base is dG or dC, the polymerase primarily discriminates between the correct nucleotide and the incorrect nucleotide based on a reduced occupancy of the active site (i.e., the rate of dissociation is much faster than the rate of incorporation). This is most likely due to the inability of these mispairs to form stable hydrogen bonds between the template and the substrate. This is consistent with the idea that fidelity for template dG and dC is mostly governed by substrate dissociation. However, this observation is not the case for the G:dTMP mispair, which exhibits a reduced K_m, while the value of V_{max} for this mismatch is unchanged relative to the other mispairs. This observation suggests that, for this base mismatch, the fidelity is increasingly governed by the rate of product formation.

Interestingly, when the template base identity is dA or dT, the values of K_m tend to be much lower than those of dG or dC and only vary by less than 2-fold, while V_{max} exhibits a greater than 20-fold variation (Table 3). This observation is consistent with the fidelity of mismatches involving template dA or dT, being governed mostly by the base-pair geometry in the polymerase active site. As mentioned previously, the T:dGMP base pair has a higher V_{max} than the correct base pair (Table 3) and indicates that the V_{max} for this mismatch does not contribute favorably to the fidelity and that the increased fidelity is primarily due to substrate dissociation.

Discussion

Pfprex is a multifunctional fusion protein targeted to the apicoplast. Even though definitive proof is lacking, it is believed to be responsible for the replication of the apicoplast genome using a DNA Pol-I-like domain. In addition, it may also function in apicoplast DNA repair and other DNA synthetic functions. Thus, its presence in the apicoplast is somewhat reminiscent of mammalian DNA polymerase γ, which is also nuclearly encoded and has been shown to carry out a variety of DNA synthetic functions in mitochondria. Because the apicoplast is a unique and required component of the malaria parasite, it could serve as an important target for specifically preventing or treating infections by the malaria parasite. To this end, we have expressed and purified—as well as present fidelity and kinetic studies of—KPom1, the plastid-targeted replicative DNA polymerase domain from the malaria parasite *P. falciparum*. Our results show that the polymerase domain of the apicoplast genome replication enzyme replicates the genome with high fidelity and has an overall fidelity similar to other fidelity A-family polymerases. The high level of fidelity stems from an efficient discrimination of mismatch nucleotides at the active site, as well as the presence of a 3′→5′ exonuclease domain involved in proofreading, which is efficient at removing the vast majority of mutations caused by misinsertion at the active site.

Characterization of the exonuclease-deficient form of KPom1 reveals the 'spectrum' of errors introduced by the polymerase. The frequencies of insertions and deletions (indels) of only a few nucleotides in length are similar to the evolutionarily related *E. coli* DNA Pol I^{exo-}, as well as other high-fidelity A-family polymerases.[15,18] The finding that these indels occur at nucleotide repeats suggests that these mutations may be the result of a strand slippage event.[22] In addition to indels, single-base substitutions at specific template positions may also be caused by a strand slippage. The M13mp2 forward mutation assay shows that a majority of the observed T→C mutations occur after a template dG or a dC. Slippage events could lead to a transient template misalignment, with the template dG or dC remaining in the catalytic site, followed by incorporation of the 'correct' nucleotide (i.e., dCMP and dGMP, respectively). Such a mechanism has been suggested to occur for human Pol γ.[22]

One of the most intriguing aspects of the KPom1 polymerase is the error spectrum for single-base substitutions. The most frequent mutations observed with both Klenow^{exo-} and KPom1^{exo-} are T→C transitions; however, the error frequency for T→C transitions is more than 3-fold greater during copying with KPom1. The next most frequently scored mutation for KPom1 is G→A, while with *E. coli* DNA Pol I^{exo-}, C→T and G→C mutations are the next most frequent base substitutions.[15] The active-site motifs in the polymerase sites for KPom1 and *E. coli* DNA Pol I are nearly identical, as is the overall accuracy in DNA synthesis; therefore, in accounting for the differences in the error spectrum, it is likely that distant amino acids have a profound

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

effect on the types of misincorporations catalyzed at the polymerase site.

Of the DNA polymerases that have been analyzed using the M13mp2 forward mutation assay, the error spectrum of KPom1^{exo-} most closely resembles that of human Pol ν, an error-prone lesion bypass polymerase that lacks a proofreading exonuclease activity.[19] While the frequency of misincorporation by human Pol ν is greater than KPom1^{exo-} (Table 2), Pol ν makes primarily G→A (G:dTMP) mutations, followed by T→C (T:dGMP) substitutions.[19] In contrast, for KPom1, the primary mutations are T→C mutations, followed by G→A substitutions. When considering the flanking sequence immediately 5′ to a mutation, it is interesting to note that G→A mutations by Pol ν are primarily preceded by a template dA or dT, while the T→C mutation of KPom1 is primarily preceded by a template dG or dC. These observations suggest that the interaction of a polymerase with the template DNA must also influence the incorporation of specific incorrect nucleotides.

The apicoplast genome base-pair composition is >86% A/T,[4] yet the most frequent mutations produced by KPom1 are T→C transitions, both in the presence and in the absence of exonuclease activity. This finding would suggest that a genome with a higher G/C content would evolve over time, thus leading to an apicoplast genome with reduced A/T content. One possible explanation as to why a G/C bias is not observed may involve biased nucleoside pools within the parasite. *In vivo* studies have previously shown that the levels of adenosine and thiamine nucleosides are several times higher in the malaria parasite than are cytosine and guanine,[24] and studies have shown that, at a genomic level, there is a distinct trend in the mutational bias towards high A/T content in bacterial obligate endosymbiotes and parasites.[25,26] This enzyme may have evolved to lower the K_m value for dGTP and dCTP due to their reduced availability in the pool. This idea is consistent with our observations that mispairs involving dGTP and dCTP tend to have a lower K_m. Therefore, the skewed mutational bias exhibited by KPom1 may be the result of a selective advantage that keeps the A/T content of the genome from becoming too skewed and reducing organismal fitness. An alternative possibility as to why a G/C bias is not observed is that the misincorporation bias at the level of the DNA polymerase could be compensated for by a mismatch repair process that preferentially removes dGMP mismatches in the newly replicated DNA strand. However, we are unaware of evidence for a mismatch repair system that functions in apicoplast DNA.

The unique error signature of KPom1 is a product of its specific structural features that govern fidelity. This begs the question as to why KPom1 has an error spectrum similar to that of human Pol ν even though the amino acid sequence is more closely related to *E. coli* Pol I. A comparison of all six conserved polymerase motifs of KPom1, *E. coli* Pol I, and human Pol ν shows that there are no amino acids conserved between KPom1 and Pol ν that are not also conserved in *E. coli* Pol I (Fig. 1b). This suggests that these motifs are not the only ones that govern the substrate selection by DNA polymerases and that other structural features or amino acids in the enzyme are also likely to affect substrate selection. Indeed, several multi-amino-acid insertions flanking these conserved motifs in human Pol ν and Pol θ (another A-family bypass polymerase) have been implicated in its reduced fidelity, suggesting that the determinants of fidelity are not strictly confined to the conserved polymerase motifs.[20,27,28]

Finally, our finding that KPom1 complements the *in vivo* activity of Pol I in *E. coli* will allow us to exploit genetic complementary for both structure–function studies and drug screening. Genetic complementation has been shown to facilitate the screening of large libraries of DNA polymerase mutants with random nucleotides at designated positions. An analysis of KPom1 mutants will be important in identifying amino residues that govern nucleotide selection and account for the unique substrate specificity. These mutants can then be compared to other A-family polymerases, as described previously.[29,30]

The ability to substitute KPom1 for Pol I brings about the feasibility of using *E. coli* as a vehicle for evaluating the potency of DNA polymerase inhibitors in a high-throughput manner. Such a system has been used to sensitize *E. coli* to the antiviral compound AZT by complementing human immunodeficiency virus reverse transcriptase.[13] The uniquely biased error spectrum of the KPom1 polymerase suggests that nucleoside analogs that exploit this bias as a method of terminating apicoplast DNA synthesis could be developed. Since the T→C error rate of KPom1 is much higher than that of human Pol δ or Pol ε, in the replicative DNA polymerases in the nucleus, the design of specific nucleoside analogs that take this fact into account will more specifically target the parasitic polymerases instead of the replicative polymerases, thus minimizing incorporation into the host genome.

Materials and Methods

Construction of recombinant plasmids

The pHSH576 derivative plasmids pECpolI and pECpolI-3′exo$^-$ that carry the *E. coli* wild-type and 3′→5′ exonuclease-deficient Pol I gene, respectively, were constructed as described previously.[14] The synthetic codon-optimized KPom1 or exonuclease-deficient KPom1^{D1531A} (KPom1^{exo-}) genes (the amino acid numbering system is

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

based on full-length Pfprex; Integrated DNA Technologies, Coralville, IA) (Supplementary Fig. 1) were cloned into the pMAL-c2X expression plasmid using the BamHI and HindIII sites. The pMAL-c2X vector encodes a maltose-binding protein moiety, which was fused to the polymerase gene to expedite purification.

Protein expression and purification

E. coli Rosetta2 cells (Novagen, EMD Chemicals) containing a pMAL-c2X-KPom1 plasmid were grown at 37 °C in 3 L of LB medium containing 50 μg/mL carbenicillin and 30 μg/mL chloramphenicol. Protein expression was induced by the addition of 0.2 mM IPTG at an OD_{600} of ~0.6. Cells were grown for an additional 20 h at 21 °C and harvested by centrifugation. Cell pellets were resuspended in 30 mL of buffer A [20 mM Tris–HCl (pH 7.4), 1 mM DTT, 1 mM ethylenediaminetetraacetic acid (EDTA), 5% (wt/vol) glycerol, 2 mM benzamidine, 500 μg/mL lysozyme, and 1 mM PMSF] + 200 mM NaCl and incubated on ice for 1 h. The crude cell extract was sonicated and clarified by centrifugation at 15,000*g* for 20 min (Supplementary Fig. S2, lane I). The supernatant was diluted with buffer A + 200 mM NaCl to 50 mL and loaded onto an amylose column preequilibrated in buffer A. The column was then washed with 100 mL of buffer A + 200 mM NaCl. KPom1 was eluted with 20 mL of buffer A + 200 mM NaCl + 20 mM maltose (Supplementary Fig. 2, lane II). The fractions containing KPom1 were pooled and directly diluted with an equal volume of buffer A without NaCl and loaded onto a heparin column preequilibrated in buffer A + 100 mM NaCl. The column was washed with 40 mL of the same buffer. KPom1 was eluted with 30 mL of buffer A + 1 M NaCl. Column fractions (2 mL each) containing fusion maltose-binding protein–KPom1 protein were pooled and dialyzed at 4 °C overnight against factor Xa cleavage buffer [20 mM Tris–HCl (pH 8.0), 2 mM $CaCl_2$, and 100 mM NaCl] and then concentrated using an Amicon filter unit (molecular weight cutoff, 30,000). The maltose-binding domain was cleaved by addition of factor Xa according to the manufacturer's protocol (New England Biolaboratories). After cleavage, the sample was diluted with 4 vol of buffer A and loaded onto an amylose column equilibrated with buffer A. The column was then washed with 15 mL of buffer A. The column flow-through (Supplementary Fig. 2, lane III) was directly diluted with an equal volume of buffer A without NaCl and loaded on a heparin column preequilibrated in buffer A + 100 mM NaCl and washed with 20 mL of the same buffer. KPom1 was eluted with 15 mL of buffer A + 1 M NaCl. Column fractions (1 mL) containing KPom1 proteins were pooled and dialyzed at 4 °C overnight against enzyme storage buffer [20 mM Tris–HCl (pH 7.4), 1 mM DTT, 0.5 mM EDTA, 50 mM NaCl, 2 mM benzamidine, and 15% (wt/vol) glycerol] and concentrated using an Amicon filter unit (molecular weight cutoff, 30,000) (Supplementary Fig. 2, lane IV). The mutant $KPom1^{exo-}$ protein was purified with the same protocol as the wild-type KPom1 preparation.

Two-plasmid β-lactamase reversion assay

The assay was performed as described previously.[14] Briefly, the Pol I^{ts} *E. coli* strain JS200 (*SC-18 recA718 polA12 uvrA155 trpE65 lon-11 sulA1*) harboring the reporter plasmid pLA230 was transformed with plasmids encoding the gene for either wild-type Pol I, Pol I^{exo-}, KPom1, or $KPom1^{exo-}$. The recombinant strains were cultured at 30 °C for 18 h in LB containing 30 mg/mL kanamycin, 12.5 mg/mL tetracycline, and 30 mg/mL chloramphenicol. A 0.01 vol of the precultured broth was inoculated into fresh growth media, then cultured at 37 °C until an A_{600} of 1.0 had been attained. Cells were then plated onto prewarmed 2×YT agar plates supplemented with 30 μg/mL kanamycin, 12.5 μg/mL tetracycline, and 30 μg/mL chloramphenicol in the presence or in the absence of 100 μg/mL carbenicillin. After incubation at 37 °C for 24 h, colonies were counted, and reversion frequencies were calculated as the ratio of carbenicillin-resistant colonies to total colonies.

Genetic complementation assay

The spiral assay for testing the genetic complementation of exogenously expressed *E. coli* Pol I and KPom1 DNA polymerases was conducted as described previously using the temperature-sensitive *E. coli* strain JS200.[31] To quantitatively determine the complementation efficiency of Pol I^{ts} cells by wild-type and KPom1 DNA polymerases, we plated ~1000 cells of JS200 harboring either pHSG576, pECpolI, pHSG576-KPom1, or pHSG576-$KPom1^{D1798A}$ (a mutant lacking polymerase activity) on 2×YT agar plates containing tetracycline and chloramphenicol at 30 °C or 37 °C. The complementation efficiency of each construct was determined as the ratio of viable colonies at 37 °C on 2×YT agar plates to those at 30 °C on plates after the 24-h incubation.[14] The results shown in all figures represent the average of three experiments carried out independently.

Polymerase activity assays

The DNA polymerase activity of all purified proteins was quantified using activated calf thymus DNA, as previously described.[14] KPom1 and $KPom1^{exo-}$ protein were assayed for 3′→5′ exonuclease activity using a duplex 5′-[^{32}P]-labeled 27/36-mer DNA containing a 3′-terminal G:G mismatch. Reaction mixtures (10 μL) contained 10 nM 5′-[^{32}P]-labeled 27/36-mer DNA and 0.01 U of wild-type Klenow fragment (New England Biolaboratories), exonuclease-deficient Klenow fragment (New England Biolaboratories), 20 nM KPom1, or $KPom1^{exo-}$ protein in the appropriate reaction buffer. Reactions were incubated at 37 °C for 15 min, terminated by addition of 10 μL of 2× gel loading buffer [100% formamide, 0.03% bromophenol blue (wt/vol), and 0.03% xylene cyanol (wt/vol)], and boiled at 95 °C for 5 min. Ten microliters of each reaction mixture was analyzed by electrophoresis on an 8 M urea/18% polyacrylamide gel.

M13 forward mutation assay

The M13mp2 gap-filling forward mutation assay was performed as described previously.[16] Briefly, the gap-filling reaction (15 μL) was carried out in polymerase reaction buffer [25 mM Tris–HCl (pH 9.0), 1 mM DTT, 10 mM $MgCl_2$, and 50 mg/mL bovine serum albumin] containing

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

4 fmol of purified double-stranded bacteriophage M13mp2 DNA with a 407-nucleotide single-stranded gap within the *lacZ*α complementation target sequence, 5 pmol of purified KPom1 or KPom1^{exo-}, and 100 μM of each dNTP. Reactions were incubated at 37 °C for 10 min and terminated by the addition of EDTA to a concentration of 10 mM. The completion of gap-filling reactions was confirmed by agarose gel electrophoresis. Aliquots of gap-filling reactions were transformed into MC1061 cells and plated on agar plates containing 5-bromo-4-chloro-3-indolyl-β,D-galactopyranoside, IPTG, and a lawn of CSH50 *E. coli* host cells. After incubation of plates at 37 °C for 16 h, the numbers of wild-type (dark blue) and mutant (light blue and white) plaques were scored. Mutant plaques were isolated and individually grown in liquid culture, and the M13mp2 DNA was isolated and sequenced using the sequencing primer 5′-TCGGAACCACCATCAAAC-3′. Error rates were calculated as previously described.[32]

Single-nucleotide insertion kinetics

The primer extension reaction was performed to determine single-nucleotide insertion kinetics, as described previously.[21,23] Briefly, the 5′-[^{32}P]-labeled 16-mer primer 5′-CATGAACTACAAGGAC-3′ was annealed to a 1.5-fold molar excess of the template 36-mer 5′-GCATTCAGTXGTCCTTGTAGTTCATG-3′, where the 'X' in boldface was either A, C, T, or G. Primer extension reactions (40 μL) were carried out in polymerase reaction buffer [25 mM Tris–HCl (pH 9.0), 1 mM DTT, 10 mM $MgCl_2$, and 50 mg/mL bovine serum albumin] containing 10 nM ^{32}P-labeled primer/template DNA, 20 nM purified KPom1^{exo-}, and indicated concentrations of dNTPs. Reactions were incubated at 37 °C from 1 min to 5 min. The time of each reaction was chosen based on prior experiments that were performed to determine single completed hit conditions, with less than 20% of the total primer extended.[23,33] Reactions were terminated by adding 40 μL of 2× gel loading buffer [95% formamide, 15 mM EDTA, 0.05% (wt/vol) bromophenol blue, and 0.05% (wt/vol) xylene cyanol]. Samples were boiled and loaded onto an 8 M urea/15% polyacrylamide gel for analysis by electrophoresis.

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.04.071

Acknowledgements

This work was supported by National Institutes of Health grants R01CA115202, R01CA102029, and P01AG033061. We thank Eddie Fox and Sharath Balakrishna for their useful comments and discussion.

References

1. Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. & Hay, S. I. (2005). The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*, **434**, 214–217.
2. He, H. Y., Shaw, M. K., Pletcher, C. H., Striepen, B., Tilney, L. G. & Roos, D. S. (2001). A plastid segregation defect in the protozoan parasite *Toxoplasma gondii*. *EMBO J.* **20**, 330–339.
3. Stanway, R. R., Witt, T., Zobiak, B., Aepfelbacher, M. & Heussler, V. T. (2009). GFP-targeting allows visualization of the apicoplast throughout the life cycle of live malaria parasites. *Biol. Cell*, **101**, 415–430.
4. Wilson, R. J. M., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A. *et al.* (1996). Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* **261**, 155–172.
5. Dahl, E. L. & Rosenthal, P. J. (2008). Apicoplast translation, transcription and genome replication: targets for antimalarial antibiotics. *Trends Parasitol.* **24**, 243–284.
6. Kumar, A., Tanveer, A., Biswas, S., Ram, E. V. S. R., Gupta, A., Kumar, B. & Habib, S. (2010). Nuclear-encoded DnaJ homologue of *Plasmodium falciparum* interacts with replication *ori* of the apicoplast genome. *Mol. Microbiol.* **75**, 942–956.
7. Ram, E. V. S. R., Kumar, A., Biswas, S., Kumar, A., Chaubey, S., Siddiqi, M. I. & Habib, S. (2007). Nuclear *gyrB* encodes a functional subunit of the *Plasmodium falciparum* gyrase that is involved in apicoplast DNA replication. *Mol. Biochem. Parasitol.* **154**, 30–39.
8. Dahl, E. L. & Rosenthal, P. J. (2007). Multiple antibiotics exert delayed effects against the *Plasmodium falciparum* apicoplast. *Antimicrob. Agents Chemother.* **51**, 3485–3490.
9. Dahl, E. L., Shock, J. L., Shenai, B. R., Gut, J., DeRisi, J. L. & Rosenthal, P. J. (2006). Tetracyclines specifically target the apicoplast of the malaria parasite *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **50**, 3124–3131.
10. Goodman, C. D., Su, V. & McFadden, G. I. (2007). The effects of anti-bacterials on the malaria parasite *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **152**, 181–191.
11. Seow, F., Sato, S., Janssen, C. S., Riehle, M. O., Mukhopadhyay, A., Phillips, R. S. *et al.* (2005). The plastidic DNA replication enzyme complex of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **141**, 145–153.
12. Sweasy, J. B. & Loeb, L. A. (1992). Mammalian DNA polymerase β can substitute for DNA polymerase I during DNA replication in *Escherichia coli*. *J. Biol. Chem.* **267**, 1407–1410.
13. Kim, B. & Loeb, L. A. (1995). Human immunodeficiency virus reverse transcriptase substitutes for DNA polymerase I in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **92**, 684–688.
14. Camps, M., Naukkarinen, J., Johnson, B. P. & Loeb, L. A. (2003). Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc. Natl Acad. Sci. USA*, **100**, 9727–9732.
15. Bebenek, K., Joyce, C. M., Fitzgerald, M. P. & Kunkel, T. A. (1990). The fidelity of DNA synthesis catalyzed by derivatives of *Escherichia coli* DNA polymerase I. *J. Biol. Chem.* **265**, 13878–13887.
16. Bebenek, K. & Kunkel, T. (1995). Analyzing fidelity of DNA polymerases. *Methods Enzymol.* **262**, 217–232.
17. Eckert, K. A. & Kunkel, T. A. (1990). High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **18**, 3739–3752.

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071

Q1

18. Longley, M. J., Nguyen, D., Kunkel, T. A. & Copeland, W. C. (2001). The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit. *J. Biol. Chem.* **276**, 38555–38562.
19. Arana, M. E., Takata, K. I., Garcia-Diaz, M., Wood, R. D. & Kunkel, T. A. (2007). A unique error signature for human DNA polymerase ν. *DNA Repair*, **6**, 213–223.
20. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I. B. & Kunkel, T. A. (2008). Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleuc Acids Res.* **36**, 3847–3856.
21. Schmitt, M. W., Matsumoto, Y. & Loeb, L. A. (2009). High fidelity and lesion bypass capability of human DNA polymerase δ. *Biochimie*, **91**, 1163–1172.
22. Bebenek, K. & Kunkel, T. A. (2000). Streisinger revisited: DNA synthesis errors mediated by substrate misalignments. *Cold. Spring Harbor Symp. Quant. Biol.* **65**, 81–92.
23. Boosalis, M. S., Petruska, J. & Goodman, M. F. (1987). DNA polymerase insertion fidelity—gel assay for site-specific kinetics. *J. Biol. Chem.* **262**, 14689–14696.
24. Dyke, K. V., Trush, M. A., Wilson, M. E. & Stealey, P. K. (1977). Isolation and analysis of nucleotides from erythrocyte-free malarial parasites (*Plasmodium berghei*) and potential relevance to malaria chemotherapy. *Bull. World Health Organ.* **55**, 253–264.
25. Hershberg, R. & Petrov, D. M. (2010). Evidence that mutation is universally biased towards AT in bacteria. *PLOS Genet.* **6**, e1001115.
26. Mann, S. & Chen, Y. P. P. (2010). Bacterial genomic G+C composition-eliciting environmental adaptation. *Genomics*, **95**, 7–15.
27. Takata, K. I., Arana, M. E., Seki, M., Kunkel, T. A. & Wood, R. D. (2010). Evolutionary conservation of residues in vertebrate DNA polymerase N conferring low fidelity and bypass activity. *Nucleic Acids Res.* **38**, 3233–3244.
28. Hogg, M., Seki, M., Wood, R. D., Doublié, S. & Wallace, S. S. (2011). Lesion bypass activity of DNA polymerase θ (POLQ) is an intrinsic property of the pol domain and depends on unique sequence inserts. *J. Mol. Biol.* **405**, 642–652.
29. Loh, E., Choe, J. & Loeb, L. A. (2007). Highly tolerated amino acid substitutions increase the fidelity of *Escherichia coli* DNA polymerase I. *J. Biol. Chem.* **282**, 1201–12209.
30. Patel, P. H. & Loeb, L. A. (2000). DNA polymerase active site is highly mutable: evolutionary consequences. *Proc. Natl Acad. Sci. USA*, **97**, 5095–5100.
31. Shinkai, A. & Loeb, L. A. (2001). *In vivo* mutagenesis by *Escherichia coli* DNA polymerase I Ile^{709} in motif A functions in base selection. *J. Biol. Chem.* **276**, 46759–46764.
32. Glick, E., Anderson, J. P. & Loeb, L. A. (2002). *In vitro* production and screening of DNA polymerase η mutants for catalytic diversity. *BioTechniques*, **33**, 1136–1144.
33. Creighton, S., Bloom, L. B. & Goodman, M. F. (1995). Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. *Methods Enzymol.* **262**, 232–256.

Please cite this article as: Kennedy, S. R., *et al.*, The Biochemistry and Fidelity of Synthesis by the Apicoplast Genome Replication DNA Polymerase Pfprex from the Malaria Parasite *Plasmodium falciparum*, *J. Mol. Biol.* (2011), doi:10.1016/j.jmb.2011.04.071