Structure and kinetic characterization of plasmodium falciparum apicoplast DNA polymerase (apPOL)

Supreet Kaur
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Biochemistry Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/17481

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Structure and kinetic characterization of *Plasmodium falciparum* apicoplast DNA polymerase (apPOL)

by

Supreet Kaur

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 Nelson, Major Professor
Richard Honzatko
Vincenzo Venditti

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University
Ames, Iowa
2019

Copyright © Supreet Kaur, 2019. All rights reserved.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................................................iii  
ABSTRACT ..................................................................................................................................v  

CHAPTER 1. INTRODUCTION ........................................................................................................1  
  Literature Review ..................................................................................................................... 1  
  References ............................................................................................................................... 5  
  Figures ...................................................................................................................................... 9  

CHAPTER 2. STRUCTURE AND KINETIC CHARACTERIZATION OF PLASMODIUM  
FALCIPARUM APICOLAST DNA POLYMERASE (apPOL) ....................................................... 10  
  Abstract ................................................................................................................................. 10  
  Introduction .......................................................................................................................... 11  
  Materials and Methods ......................................................................................................... 13  
  Results and discussions ........................................................................................................ 18  
  References ............................................................................................................................ 22  
  Figures ................................................................................................................................... 24  

CHAPTER 3. CONCLUSIONS ...........................................................................................................42
Firstly, I would like to express my sincere gratitude to my advisor Prof. Scott Nelson for the continuous support of my Masters’ study and related research, for his patience, motivation, and immense knowledge. I appreciate all his contributions of time, ideas, and funding to make my research experience productive and stimulating. The enthusiasm he has for his research was motivational for me and his guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Richard Honzatko, and Dr. Vincenzo Venditti, for their insightful comments and encouragement. I would like to express my special appreciation and thanks to Prof. Honzatko for guiding me through Crystallography. Prof. Honzatko has provided me with every bit of guidance, assistance, and expertise that I needed during my experiments of Crystallography.

I thank my fellow labmates in for the stimulating discussions, their help in performing the experiments, and for all the fun we have had in the last three years. In particular, I am grateful to Tibebe for always helping and troubleshooting when I get stuck in conducting experiments. I would like to take this opportunity to acknowledge Shravanti’s contribution in my thesis. She has been a great help in proofreading and answering my silly questions.

Lastly, I would like to thank my family for all their love and encouragement. For my parents who raised me with a love of science and supported me in all my pursuits. I am so grateful to my mother-in law, father-in-law, my mother, and father for great help in taking care of my new born girl “Sarmaya”. To my beloved daughter Pinaaz Kaur, I would like to express
my thanks for being such a good girl and a caring daughter. And most of all for my loving, supportive and encouraging husband Prabhjot Singh, whose faithful support during my research is so appreciated. Thank you for proofreading all the chapters of this thesis and guiding me to write effectively. His unconditional help and support has led me to complete my thesis.
Malaria is caused by the parasite *Plasmodium falciparum* and is amongst the world’s most deadly infectious diseases for which there is no vaccine. Malaria is currently controlled through drug therapy, but the formation of resistance is common and new drugs that target novel parasite functions are urgently needed. *P. falciparum* contains non-photosynthetic organelle known as the apicoplast. The apicoplast harbors its own DNA and a single DNA polymerase (apPOL) is targeted to the apicoplast to carry-out genomic replication and repair. apPOL has no direct ortholog with mammalian polymerases, making it an appealing drug target for the treatment and/or prevention of malaria. The Medicines for Malaria Venture (MMV) selected 400 structurally diverse compounds that resulted from very large-scale screening campaigns performed by St. Jude Children's Research Hospital, Novartis, and GSK. These compounds are able to kill *P. falciparum* in cultured red blood cells and are non-toxic to cultured human cells, but their molecular targets are unknown. We previously screened the Malaria Box for apPOL inhibitors and identified a single compound (MMV666123) that completely inhibits apPOL at a concentration of 10 µM. This thesis describes the kinetic characterization of that compound, along with 36 additional structural derivatives that were generated by our collaborators in an effort to improve inhibitor potency. In addition, to guide the synthesis of additional derivatives, we determined the x-ray crystal structure of one of the derivatives bound to apPOL. The compound binds at a previously unidentified allosteric pocket that sits adjacent to the polymerase active site. We propose that the binding of the derivative to the allosteric site prevents the polymerase from adopting the closed, active conformation of the enzyme. The allosteric site is present in other A-family polymerases, but the identity of the residues lining the pocket differ, suggesting the site could be targeted with high specificity.
CHAPTER 1. INTRODUCTION

Literature Review

Malaria

For humans, one of the deadliest creatures on earth is the mosquito. Nearly 220 million people die every year because of diseases that are transmitted through mosquito bites (1). Out of these, 600,000 deaths occur due to malaria alone (2). Parasites from the genus *Plasmodium* are responsible for Malaria with five species infecting humans: *Plasmodium falciparum; Plasmodium vivax; Plasmodium ovale; Plasmodium malariae;* and *Plasmodium knowlesi*. Among these five species, the deadliest is *P. falciparum*, which especially poses a great threat to tropical and the subtropical regions that include Africa, Asia, and Latin America (3).

*P. falciparum* follows a complex lifecycle that is divided between the mosquito vector and the human host. The malaria parasite’s life cycle is very complex and involves many distinct stages. As the mosquito draws blood meal from human, malarial sporozoites enters into the human liver and develops into to a mature schizont. A schizont creates thousands of merozoites, which bursts out of the cell into the bloodstream, where it infects red blood cells. A few of the merozoites mature into gametocytes that may be ingested by another mosquito during subsequent feeding (4). Inside the mosquito’s gut, gametocytes are fertilized and then develop into a zygote that further transforms to oocysts, eventually releasing sporozoites. Sporozoites are finally transmitted to new human host during an infective bite (2). Nearly half of the world’s population is at risk of malaria and although efforts to control this disease are on the rise, the disease has not been eradicated yet (5). The greatest challenge to this problem lies in drug resistance. The
malaria parasite has developed resistance against drugs like chloroquine and sulfadoxinopyrimethamine (3). Additionally, research suggests that artemisinin resistance is rapidly spreading in southeast Asia. There is no drug available that does not have some level of resistance (4). Therefore, more research on identifying effective drugs is warranted to eradicate this disease.

**Apicoplast**

*Plasmodium* is a member of the phylum Apicomplexa that consists of an unusual non-photosynthetic plastid known as Apicoplast (Figure 1). The apicoplast has a 35 kb circular, supercoiled DNA and is the smallest of the plastid genomes. Evidence suggests that the apicoplast emerged from the chloroplast from a red-algae (9). The chloroplast is surrounded by four membranes, resulting from the secondary endosymbiotic event. It is important to note that the secondary endosymbiosis is a symbiotic relationship between two eukaryotic cells in which one cell contained a plastid (a modified photosynthetic bacterium residing inside a eukaryotic cell) and was engulfed by the other eukaryotic cell. As a consequence, *P. falciparum* has both eukaryotic and prokaryotic genomes. The apicoplast has evolved and lost its photosynthetic characteristics over time (6). Nevertheless, heme synthesis, fatty acid biosynthesis, isoprenoid precursor synthesis and iron-sulfur complex synthesis are four biosynthetic pathways that have been found to be associated with the apicoplast (7). Identification of apicoplast-targeting sequence confirmed that the proteins required for these biosynthetic pathways are encoded in the nucleus that are targeted to apicoplast (8). This suggests that the apicoplast could a promising drug target for the treatment and/or prevention of malaria.
The apicoplast genome consist only of 68 coding genes that include duplicated t-RNAs, large and small subunit rRNAs, RNA polymerase subunits, several protein chaperons and ribosomal proteins (6, 9). The apicoplast has also retained the minimal autonomy of gene order by undergoing several deletions and rearrangements of the genes which resulted in its reduced genome size from 150 kb to 35kb (6).

**Replication of the apicoplast genome**

The apicoplast does not encode for genes responsible for the replication of its own genome. The majority of the proteins residing in the apicoplast are nuclear encoded and must be imported to function properly (10). In the presence of signal peptide, nuclear encoded apicoplast proteins are targeted to the organelle through secretory pathway (11). Signal peptide facilitates the entry of nascent polyprotein into the endoplasmic reticulum (ER) and with the help of transit peptide, the proteins are targeted to apicoplast (12).

The replication machinery of the apicoplast consists of a primase, a DNA helicase, DNA polymerase, single-stranded binding (SSB) proteins, DNA gyrase (topoisomerase) and ligase (11). Because the apicoplast is essential to the parasite and the apicoplast cannot function without genomic replication, all of these proteins could be considered targets for anti-malarial drugs. DNA gyrase is inhibited by fluoroquinolones and ciproflaxin, both of which show anti-malarial activity (11).

Prevention of apicoplast replication does not result in immediate cell death; however, blocking the replication process does prevent the parasite from entering other host cells. This phenomenon is called delayed cell death, which highlights the importance of apicoplast in invasion or production of parasitophorous vacuole in the host cell (13). This further suggests that
blocking of replication seems plausible by either targeting DNA polymerase or any of the proteins mentioned above.

**The apicoplast DNA polymerase**

On the basis of its amino acid sequence, the *P. falciparum* apicoplast DNA polymerase is classified as an atypical A-family polymerase. It is also one of the most widely studied enzymes among Prex domain enzymes. Prex (*Plasmodium falciparum* Plastidic DNA Replication/repair Enzyme Complex) is a poly-protein that encodes DNA primase, helicase, and polymerase (14). Prokaryotic A-family polymerases generally contain three functional domains: 1) a 5’ to 3’ polymerase domain; 2) a 3’-5’ exonuclease domain for proof reading incorrect nucleotide incorporations; and 3) a 5’-3’ exonuclease domain which is responsible for removing RNA primers from the lagging strand during replication and nick translation in during DNA repair. However, apicoplast DNA polymerases in the phylum Apicomplexa are lacking the 5’-3’ exonuclease domain, instead they have an independent gene coding for an this activity (15).

The apicoplast DNA polymerase is considered to be a promising drug target because it targets to the apicoplast and its inhibition, thereby preventing all DNA replication within the organelle (13). A high-throughput inhibition assay would be an appropriate experimental methodology in exploring effective inhibitors for DNA polymerase (16). The suggested methodology would be helpful in studying the inhibitors through different mechanisms: 1) Inhibitor that competes for the active site 2) Inhibitor that binds to the active site and also blocks the DNA/substrate from binding to the enzyme/polymerase 3) Inhibitor that binds away from DNA/substrate binding site and 4) Inhibitor that binds with the substrate forming an inhibitor-substrate complex (17).
Prior research has widely used the Malaria Box, a library that contains about 400 compounds that are active against the red blood cell stages of *Plasmodium falciparum*. These compounds were selected from 2 million compounds that were screened by GlaxoSmithKline (GSK), Novartis and St. Jude Children’s Research Hospital using a phenotypic assay for parasite killing (18). Out of the 400 compounds, 200 are considered to be “drug-like” compounds (for oral drug discovery and development) and 200 were “probe-like” compounds that could be used as biological tools for malaria search. It is noted that these 400 compounds are unique, such that they are different from currently available anti-malarial compounds in the market. These compounds have been screened for cytotoxicity and it was found that they are within the range considered acceptable for an initial drug discovery program. The purpose of Malaria box was to identify novel drug target and compound optimization (19). We previously screened the Malaria Box for apPOL inhibitors and identified a single compound (MMV666123) that had inhibitory activity against apPOL. We then kinetically characterized the parent compound along with 36 structural derivatives to improve inhibitor potency. One of these derivatives was crystallized with apPOL and the structure of the complex was determined. The structure revealed a novel allosteric site that when liganded may prevent the polymerase from achieving the active, closed conformation of the enzyme.

References


Figure 1: Apicoplast Structure (adapted from Baum et al. 2006).
CHAPTER 2. STRUCTURE AND KINETIC CHARACTERIZATION OF PLASMODIUM FALCIPARUM APICOPLAST DNA POLYMERASE (apPOL)

Supreet Kaur and Scott W. Nelson

From the Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa 50011

Modified from a manuscript in preparation for Journal of biological chemistry

Abstract

Malaria is caused by parasitic protozoan Plasmodium falciparum, which contains an essential organelle referred to as the apicoplast. The apicoplast is a non-photosynthetic plastid harboring its own genetic material. A single DNA polymerase, apPOL, is targeted to the apicoplast, where it is responsible for replicating and repairing the genome. apPOL does not contain direct orthologs in mammals; therefore, it is considered to be a promising drug target for the treatment and/or prevention of malaria. The Malaria Box is a collection of 400 compounds, selected from a phenotypic screen of nearly four million compounds for activity against the blood stage of Plasmodium falciparum. We previously screened the Malaria Box, finding a single compound (MMV666123) as a selective inhibitor of apPOL. Steady-state kinetics indicate that the mechanism of inhibition is non-competitive against deoxynucleotides. A structure-activity relationship analysis of MMV666123 was performed using 36 derivatives. One derivative class of MMV666123 facilitated crystal structure determinations of inhibitor-apPOL complexes. These derivatives bind to a novel allosteric pocket approximately 20 Å from the active site. Derivative binding to the allosteric site could prevent the polymerase from adopting the closed conformation of the enzyme necessary for catalytic activity.
Introduction

Malaria is a deadly disease caused by protozoans *Plasmodium falciparum* and *Plasmodium vivax* among other species belonging to the genus *Plasmodium*. According to the WHO malaria report, about 435,000 deaths in 2017 were due to malaria (1). Children under 5 years of age are most affected by this disease in tropical and sub-tropical regions (2). Around half the world’s population is under the threat of malaria (1). With the rapid emergence of malaria parasites resistant to standard drugs (chloroquine, sulfadoxine–pyrimethamine, mefloquine, amodiaquine and artemisinin), there is an urgent need for drugs that target novel features of disease progression (3). The identification of potential drug targets and the development of potent inhibitors of those targets is one of several approaches, the combination of which will eliminate this disease.

*Plasmodium* organisms of the phylum Apicoplexa have an non-photosynthetic plastid known as apicoplast that is indispensable for the parasite’s existence (3). Other parasites of the phylum Apicoplexa are responsible for diseases of humans such as Toxoplasmosis, Cyclosporiasis, and Cystoisosporiasis, and diseases of livestock such as Babesiosis, Theileriosis, and Coccidiosis (4). Studies support that apicoplast has evolved from the chloroplast of red algae from secondary endoplasmic events (5). With the course of time during its evolution, the apicoplast lost its photosynthetic characteristics (13). However, it retained biosynthetic pathways for heme, fatty acid, isoprenoid precursor and iron-sulfur complex synthesis (6).

The apicoplast genome lacks genes for DNA replication. However, the nuclear genome has a gene named _Plastidic DNA Replication/repair Enzyme Complex_ (Prex). Prex encodes a DNA primase, helicase, and polymerase. Prex is co-translated as a polyprotein into the lumen of
the endoplasmic reticulum (ER), ultimately reaching the apicoplast with the help of apicoplast targeting sequence (3). DNA replication is the process of producing two identical copies of an organism’s genetic information. The central enzyme involved in replication is the DNA polymerase, which catalyzes the addition of deoxyribose mononucleotides to the 3'-end of the primer of a DNA primer/template complex.

On the basis of sequence, DNA polymerases fall into different families: A, B, C, D, X, Y, and reverse transcriptase (RT). The A-family DNA polymerases include some of the best structurally and biochemically characterized polymerases (7). All DNA polymerases resemble the shape of a human right-hand, with the thumb, a palm and fingers representing separate domains, all of which play a role in DNA replication. The palm activates the 3'-hydroxyl of the primer, catalyzing the nucleophilic displacement of pyrophosphate from deoxynucleoside triphosphates (dNTPs) bound primarily to the fingers domain. The thumb domain positions the duple moiety of the DNA primer/template, and plays a role in translocation and processivity (7). dNTPs that complement the template by correct Watson-Crick base pairing, drive the fingers domain from an open to a closed (catalytically productive) conformational state. “Enforced” Watson-Crick interactions at polymerase active site is a major factor in governing the fidelity of the polymerase (7). apPOL also undergoes opening and closing conformational changes with high rates of fidelity (4,7). X-ray crystal structures have demonstrated that all DNA polymerase families share a similar palm domain, but there is a lack of homology in the thumb and finger domains.

DNA replication is indispensable for the proliferation and the survival of all living organisms (9). Because of this, DNA replication has been a target of many therapeutic strategies for controlling disease. The *P. falciparum* apicoplast DNA polymerase (apPOL) is an A-family
polymerase of prokaryotic origin, therefore selective inhibition of its activity should be possible. Studies suggest that apPOL is indispensable for the apicoplast genome’s replication (10) (2). Known inhibitors of prokaryotic DNA replication, such as ciprofloxacin, inhibit apicoplast DNA replication and cause a “delayed death” phenotype in *Plasmodium* parasites. Delayed death is a phenomenon in which the drug does not directly kill the parasite; instead, parasite progenies are unable to complete their replicative cycle. Even if the drugs are discontinued before the completion of the parent’s life cycle, the progeny still die (10). This suggests that the apicoplast could be a promising drug target for malaria treatment and/or prevention and could even be used to boost natural immunity to the parasite (11).

In order to determine compounds that inhibit apPOL, we screened the 400 compound Malaria Box using an in vitro DNA synthesis assay and identified one compound (MMV66123) that showed selective inhibition of apPOL. Inhibition parameters and the mechanism of inhibition of the compound were determined by kinetics and determined that it is non-competitive against dNTPs with two binding sites. Similarly, the *IC*$_{50}$ of 36 derivatives of MMV666123 were determined, along with an x-ray crystal structure of one of the derivatives bound to apPOL. The structure is consistent with the results of kinetics and suggests modification to the inhibitor structure that may lead to tighter binding.

**Materials and Methods**

**Materials**

Nickel-agarose was sourced from Sigma-Aldrich Chemical Company and deoxyribonucleotides were bought from Thermo Fischer Scientific. Derivatives of the inhibitor MMV666123 were synthesized by Dr. Kerns Laboratory at the University of Iowa. DNA hairpin
substrate with Cy3 dye at the 5’-end and Black Hole Quencher-1 at the 3’-end was synthesized by Integrated DNA Technologies. pET-SUMO vector and SUMO protease for crystallization were prepared by Morgan Milton. PEG 1,000 was bought from Hampton Research. Spin-X centrifuge tube filters used for filtration of protein were purchased from Sigma-Aldrich.

**Protein purification**

The purification procedure was identical for the wild type and exonuclease deficient polymerases. *E. coli* BL21 (DE3) cells were transformed with pet28-apPOL\textsuperscript{exo−} C terminal polyhistidine tag vector. A single colony was obtained from a LB/kanamycin agar plate and transferred to flasks containing 100 ml LB/kanamycin medium. After shaking the flasks for 16h at 37°C, 10 ml of the LB starter culture were transferred to two flasks each containing a liter of LB. The flasks were shaken (225 rpm) at 37°C until A\textsubscript{600} reached 0.8 and then were cooled to 18°C by holding them at 4°C for one hour, after which apPOL expression was induced by adding 200 μM (final) isopropyl β-D-1-thiogalactopyranoside (IPTG). After 16 h with constant shaking at 16°C, the cells were collected by centrifugation at 4000 rpm for 10 minutes. After constant shaking, the pellet was obtained and was re-suspended in 20 mL nickel binding buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole, pH 8, 4°C) and passed through a homogenizer at 12,000 psi. After centrifugation for 50 mins at 17000 rpm, the supernatant was loaded on Ni-agarose resin. 200 ml of binding buffer and then 100 ml of high salt wash buffer (20 mM Tris, 1 M NaCl, 25 mM imidazole and 20% glycerol, pH 8.0, 4°C) were passed through the column which was then followed by 30 ml (20 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole and 20% glycerol, pH 8.0, 4°C). Then the elution buffer (20 mM Tris-HCl, 200 mM NaCl, 150 mM imidazole and 20% glycerol, pH 8.0, 4°C) was used to obtain the protein. The protein was then concentrated to 7 mg/ml as determined spectroscopically with the help of extinction coefficient of ε\textsubscript{280} =56750
M⁻¹ cm⁻¹. Aliquots of purified apPOL<sub>exo</sub> were flash-frozen in liquid nitrogen and stored at 193 K. The freezing/thawing process does not affect the polymerase activity.

For crystallization of apPOL with inhibitors, the following steps were performed. First, a N-terminal polyhistidyl-SUMO (Small Ubiquitin-like Modifier) tag was inserted before the N-terminal residue of apPOL protein through molecular cloning. The SUMO tagged apPOL was then expressed and purified with nickel-NTA agarose column chromatography. Second, after eluting the protein from the nickel-NTA agarose column, it was further purified through TOYOPEARL HW-50F (150 mL) gravity column (TOSOH), which was previously equilibrated with 20 mM Tris-HCl, 400 mM NaCl, pH 8.0 at 4°C. Fractions were analyzed using Bradford assays and SDS-PAGE and those containing apPOL were pooled. The purified protein was incubated for 16 hours at 4°C with a 1:50 protein to SUMO protease ratio. SUMO protease specifically cuts the amide linkage after the C-terminal residue of the SUMO domain. Thereafter, the protein solution was passed through the nickel-NTA-agarose column, retaining the unwanted polyhistidyl-SUMO fragment and passing the tag-free apPOL. The tag-less protein was confirmed by SDS-PAGE using Coomassie Blue stain, and concentrated to 12 mg/ml. Freshly prepared apPOL was passed through a cellulose acetate Spin-X centrifuge tube filter (0.22 μm pore size), and then used in crystallization experiments.

**Kinetic experiments**

Assays of apPOL<sub>exo</sub> employed a DNA hairpin substrate in 20 mM Tris-HCl, 10 mM magnesium acetate, 50 mM potassium acetate, pH 8.0 at 25 °C. Reactions used glass cuvettes and a reaction volume of 280 mL. Progress curves of were monitored with a Cary Eclipse Fluorescence Spectrophotometer. An excitation and emission wavelength of 545 nm and 570 nm respectively were used for the experiments. Data were collected using the Cary Kinetics
software. Inhibition assays used a 96-well plate and plate reader at the same wavelengths. The assays were carried out for 20 minutes at an enzyme and dNTP concentration of 20 nM and 16 mM, respectively. A time course at these concentrations confirmed the linearity of the reaction.

MMV666123 (referred to here as PC-01-171D) was identified as an inhibitor of apPOL<sub>exo</sub>- through high-throughput screening. The \( IC_{50} \) of PC-01-171D for apPOL was determined using 20 nM apPOL<sub>exo</sub>-, 16 \( \mu \)M dNTPs and 100 nM DNA substrate with varying concentration of PC-01-171D from 0-40 \( \mu \)M. PC-01-171D was incubated with apPOL<sub>exo</sub>- for 10 minutes in 96 well plate. With the addition of DNA substrate, the reaction was started with the help of dNTPs for 20 minutes. Finally, the reaction was quenched using 0.05 M ethylenediaminetetraacetic acid (EDTA). \( IC_{50} \) values were determined through Sigma plot software. Similarly, \( IC_{50} \) for all the derivatives of PC-01-171D were determined.

The inhibition mechanism of PC-01-171D versus nucleotides was also performed in 96 well plates by varying inhibitor concentration from 0 to 6 \( \mu \)M and dNTPs concentrations from 0 to 128 \( \mu \)M. DNA substrate and apPOL<sub>exo</sub>- concentrations were kept constant at 100 nM and 20 nM, respectively. We used Dynafit software to fit the resulting data against inhibition models containing competitive, noncompetitive, mixed, and uncompetitive inhibition.

**Crystallization and data collection**

In general, the inhibitors exhibited limited solubility in aqueous buffers (as determined visually). As the urea-based derivatives (such as PC-01-194) seemed most soluble in aqueous buffers, these were tried first. Solid inhibitor was added incrementally to a concentrated solution of apPOL (12 mg/mL) to the point of saturation and incubated at room temperature for 10
minutes. Excess (undissolved) inhibitor was removed by centrifugation. Each inhibitor required a separate preparation of an inhibitor-saturated apPOL solution.

Crystallization experiments four commercial screens (PEGRx, SaltRx, Morpheus, and Structure) in 96-well sitting-drop plates obtained from Hampton Research. Screening employed a Mosquito Crystallization Robot in the X-Ray facility at Iowa State University. Out of 400 conditions, one condition produced crystals. The results were successfully replicated and then optimized manually by the method hanging-drop vapor diffusion. In so doing, the apPOL-saturated inhibitor solution in 20 mM Tris-HCl, 400 mM NaCl, pH 8.0 was mixed with the well solution in equal ratio, producing an initial droplet. Identical crystallization experiments of the protein inhibitor complexes were kept at room temperature and at 4°C. A third set of experiments had 400 mM NaCl in the well solution before combining 1:1 with the protein solution. The latter experiment was pursued to force the migration of water from the droplet to the well. After crystal formation, the crystals were looped and immediately plunged into liquid nitrogen for storage and subsequent data collection.

The data of diffraction was collected on Advanced Photon Source beamline 23-ID-D at 100 K. 600 frames were collected using an oscillation range of 0.2° with a 20 μm beam. HKL-3000 was used for data processing. Molecular replacement was performed using the PHENIX software suite, followed by model building and refinement using Phenix.refine and Coot.

Results and Discussion

Protein purification

The yield of apPOLexo− was approximately 25 mg in a liter of Luria broth. The SUMO tag was at least 95% cleaved, as determined using Coomassie Blue stained SDS–PAGE (Figure 1).
A spin concentrator was then used to concentrate the tag-less protein to 12 mg/mL. As with N- and C-terminal polyhistidyl constructs of apPOL, the presence of 400 mM NaCl prevents precipitation of the protein at concentrations of 12 mg/mL.

**Kinetic experiments**

A stem-loop oligonucleotide with a primer annealed to the loop and a Cy3 fluorophore at 5’ end and Black Hole Quencher at 3’ end of the stem was used as a DNA substrate for the DNA synthesis assays (Figure 2). apPOL extends the primer at the 3’-end and performs strand displacement synthesis when it reaches the stem section. This opens the DNA duplex and separates the Cy3 fluorophore from the quencher, causing increased fluorescence. We performed fluorescence assays varying concentrations of apPOL<sup>exo−</sup> and found the reaction to be approximately linear (Figure 3A). Additionally, we observed that apPOL<sup>exo−</sup> was stable for several hours and multiple preparations of apPOL<sup>exo−</sup> had similar activity. Continuous and stopped-time assays were conducted and we observed linearity for approximately 20 minutes (Figure 3B & 3C).

We used apPOL<sup>exo−</sup> activity as an output to measure the $IC_{50}$ of the parent compound, PC-01-171-D at varying concentrations (Figure 4A & 4B). For this compound, we determined an $IC_{50}$ of 5.0 μM. Using Dynafit software, the concentrations of dNTPs and PC-01-171-D were varied on Km and $IC_{50}$. We noted that the best-fit model was a true noncompetitive with two independent binding sites of different affinities (i.e., two molecules of inhibitor can bind per enzyme). The $K_i$ of high-affinity binding site was 0.6 μM whereas the $K_i$ of low-affinity binding site was 12 μM (Figure 4C). The parent compound has a moderate $IC_{50}$ and displays 100% inhibition activity at 10 μM; however, it does not meet the guidelines suggested by the Medicines for Malaria Venture (MMV). The guidelines for an ideal inhibitor according to MMV
are 1) IC$_{50}$ is less than 10 nM against target; 2) The compound is soluble; and 3) The compound is stable for 24 months. In order to improve the properties of PC-01-171-D, several derivatives were synthesized by Pratik Chheda in the Kerns laboratory at the University of Iowa. These derivatives included sulfonamide and urea-based analogs (Figure 5). IC$_{50}$/s were measured for the 36 derivatives to define structure-activity relationships and determine whether any derivative showed improved inhibitory properties (Figure 6 A and 6B). We noted that the structural differences between the derivatives resulted in different values of IC$_{50}$/s. At least three derivatives: PC-01-171-C, PC-01-186, and PC-01-194 retained the inhibitory properties of the parent compound (100% inhibition activity at 10 μM), but were more soluble based on a visual analysis; however, none of the 36 compounds were significantly better inhibitors.

**Crystallization**

Structural work pursued the urea-based derivatives of PC-01-171-D, because of their higher solubility and comparable IC$_{50}$ values. Crystals were obtained by the hanging-drop method in droplets containing a 1:1 ratio of inhibitor-saturated protein solution and well solution, and grew from 0.1 M sodium citrate tribasic dihydrate pH 5.5, 30% w/v polyethylene glycol 1000 (PEGRx 1 condition No.17). However, after four days at the room temperature, crystals began to dissolve, and disappeared after a week. Crystals were stable (for at least a month) and bigger in size if crystal plates were kept at 4 °C and/or if 400 mM NaCl was added to the well solution, before combination with the protein solution. This indicates that the optimum temperature and condition to obtain stable crystals is 4 °C and excess NaCl. Next, we found that the crystals resembled (see Figure 8) the leaf of a flower with an approximate dimension of 200 × 50 × 50 μm (Figure 7). We note that the crystals obtained were larger in size and much more reproducible than crystals formed with the N- or C-terminally tagged constructs.
Structure determination

A 20 μm microbeam provided diffraction from crystals of the apPOL-PC-01-194 complex to a resolution of 2.7 Å with an exposure time of 0.1 seconds per frame. Previous crystals formed from protein with the N- or C-terminal hexahistidine tags diffracted to 2.9 Å, suggesting that either the presence of the compound or the lack of tags results in a slight improvement in overall resolution. Data reduction indicated a space group P6$_5$ with unit-cell parameters a = b = 145.4 Å, c = 164.9 Å, α = β = 90°, γ = 120°. The number of observations used in refinement were 54,200 with R-work (last shell) = 0.19 (0.29) and R-free (last shell) = 0.24 (0.33). The position of the inhibitor bound to apPOL is distal to the polymerase active site, which is consistent with the non-competitive kinetic mechanism that was determined (Figure 9).

The asymmetric unit of the crystal contains two polypeptides chains of apPOL (Chain A and Chain B). Each chain has at least two bound chloride ions. We have found apPOL to be somewhat unstable at NaCl concentrations less than 400 mM, therefore the protein is stored and crystallized in the presence of 400 mM NaCl. The observed chloride near the polymerase active site (interacting with Lys$^{412}$, Gln$^{413}$, Lys$^{417}$, Glu$^{615}$) has four hydrogen bonds with proton donors (Figure 10). Similarly, there is a second chloride present near the PC-01-194 binding site (interacting with Asn$^{234}$, Asn$^{433}$, Asn$^{529}$, Lys$^{531}$) (Figure 11). Asn$^{433}$ of the inhibitor binding site comes from the other polypeptide chain of asymmetric unit. This chloride ion appears to prevent the inhibitor from binding in its “optimal” conformation (the optimal binding mode is in Figure 12), therefore we are currently pursuing co-crystallization conditions with reduced NaCl concentrations.
Two molecules of PC-01-194 are present among the apPOL chains of the asymmetric unit where inhibitor #1 binds at Chain A while inhibitor #2 binds at Chain B. Hydrogen bond is present between N11 of inhibitor #2, Ser 525 (chain B) and Thr 526 (chain B) (Figure 13).

At the tri-methyl end of the PC-01-194 binding pocket lies a highly conserved aromatic residue (Trp$^{512}$ in apPOL). All A-family polymerases undergo repeated cycles of opening and closing their finger and thumbs domains when correct incoming nucleotide H-bonds with the templating base at the polymerase active site. This closing of the polymerase around the Watson-Crick base pairs is thought to be the structural basis for polymerase fidelity (8). A comparison of the open and closed states of the A-family polymerase from Thermus aquaticus (Klentaq), indicates that this conserved aromatic residue (Trp$^{706}$) moves 4 Å into the compound’s binding pocket in the closed state of the enzyme. Assuming apPOL undergoes a similar conformational change in the closed conformation, the trimethyl group PC-01-194 would clash with the indol nitrogen of Trp$^{512}$ (Figure 14). This strongly suggests that mechanism of inhibition of PC-01-194 is to prevent the closed, active conformation of the polymerase by sterically blocking the movement of Trp$^{512}$.

It is important to note that the binding pocket for PC-01-194 is found in all A-family DNA polymerases, although the identities of the residues lining the pocket are only partially conserved. The lack of conservation with the three human A-family polymerases (theta, nu, and gamma), Figure 15 suggest that PC-01-194 would have low affinity for these enzymes, consistent with the observation that the parent compound is non-toxic to a variety of different human cells lines (12). The ubiquitous presence of this allosteric pocket in A-family polymerases suggests that it could be targeted to control the activity of other medically relevant A-family polymerases, particularly for the development of new antibacterial drugs.
References


8. The Closing Mechanism of DNA Polymerase I at Atomic Resolution | Elsevier Enhanced Reader [Internet]. [cited 2019 Apr 14]. Available from:


Figure 1. SDS-PAGE gel analysis following purification of tagless apPOL.
Figure 2. A schematic of DNA substrate and the steps leading to an increase in fluorescence.

Additional details can be found in the text.
Figure 3. Polymerase activity. (A) Optimal concentration of polymerase was identified to be 20 nM for inhibition assays. (B) Time course showing the polymerase-dependent increase in fluorescence using the substrate. (C) Time points to optimize reaction time. The optimum reaction time for inhibition assays was 20 mins.
A

B
Figure 4. Mechanism of the parent compound of MMV666123 (named as PC-01-171D). (A) Structure of PC-01-171D. (B) The IC\textsubscript{50} of PC-01-171D was determined by the percentage of polymerase activity at variable concentrations of PC-01-171D in \( \mu \text{M} \). Data were fit to the standard IC\textsubscript{50} equation and an IC\textsubscript{50} of 5.0 \( \mu \text{M} \) was determined. (C) Plots between initial velocity and dNTPs concentration of PC-01-171D (0 \( \mu \text{M} \), 2 \( \mu \text{M} \), 4 \( \mu \text{M} \) and 6 \( \mu \text{M} \)). The solid lines represent non-competitive mechanism with two binding sites of different affinities.
Figure 5. All the derivatives of the parent compound.
Figure 6 (A & B). Structures and $IC_{50}$ of the derivatives of PC-01-171D.
Figure 7. apPOL- inhibitor complex crystal with dimensions of approximately $200 \times 50 \times 50$ μm.
Figure 8. Leaf like crystals of apPOL-inhibitor complex.
Figure 9. Crystal structure of apPOL-inhibitor complex. (A). Structure and IC$_{50}$ of PC-01-194. (B) Crystal structure showing the binding site of the inhibitor that is located in borderline druggable pocket.
Figure 10. Polymerase active site showing chloride ion (purple) near bound to it.
Figure 11. Inhibitor binding site showing chloride ion (blue) bound near to it.
Figure 12. Possible conformation of the inhibitor when chloride ion is not present.
Figure 13. Hydrogen bond of N11 of inhibitor with Ser$^{525}$ and Thr$^{526}$
Figure 14. Possible mechanism of PC-01-194. Green color shows the open confirmation of apPOL and cyan shows the closed confirmation of klentag. Closed confirmation is blocked due to steric clash of tryptophan residue with the inhibitor.
Figure 15. Residues lining the pocket showing lack of conservation with the three human A-family polymerases (theta, nu, and gamma).
CHAPTER 3. CONCLUSION

In this thesis, I determined the IC_{50} and preliminary kinetic mechanism of the MMV666123, an apPOL inhibitor that was selected from the 400 compounds of Malaria Box. It was found that the compound did not meet the suggested guidelines presented by Medicine for Malaria Venture (MMV). In an attempt to develop a better compound, several derivatives were synthesized with a few modifications in the structure by our collaborators. However, the unguided modification of the derivatives did not lead us to improved properties. Therefore, it was decided that structural characterization of the compounds would be necessary for meaningful improvements in inhibitor potency.

Co-crystallization was performed to explore the apPOL-inhibitor complex. The crystal structure of apPOL with the inhibitor PC-01-194 was consistent with the kinetic mechanism experiments, which showed two binding sites with different affinities. However, preliminary data of another crystal structure of apPOL with the inhibitor PC-01-188 exhibited the same type of binding in two sites that had similar affinities. These intriguing results reveal the necessity for exploring other derivatives for possible kinetic mechanisms that can explain the polymerase-inhibitor interaction.

Future work will include examining the impacts of point mutations. Moreover, the substrate can be bound to the crystal structure of apPOL and inhibitor so that we can have a better understanding of the inhibition mechanism. In so doing, this would guide our collaborators to build more potent and specific variations in the compounds that inhibit apPOL.

Additionally, future work can address the challenges posed by the presence of chloride ions in the apPOL inhibitor complex. Putatively, chloride ions prevent the inhibitor from
obtaining its optimal conformation. To overcome this problem, co-crystallization can be performed by reducing the amount of NaCl in the well solution. Another way to address this problem would be to get rid of the excess of NaCl by letting the crystal grow at 4 °C.