3-2015

Crystallization and preliminary X-ray analysis of the Plasmodium falciparum apicoplast DNA polymerase

Morgan Eilise Milton
Iowa State University, eilise.miller@gmail.com

Jun-yong Choe
Rosalind Franklin University of Medicine and Science

Richard B. Honzatko
Iowa State University, honzatko@iastate.edu

Scott W. Nelson
Iowa State University, swn@iastate.edu

Crystallization and preliminary X-ray analysis of the Plasmodium falciparum apicoplast DNA polymerase

Abstract
Infection by the parasite *Plasmodium falciparum* is the leading cause of malaria in humans. The parasite has a unique and essential plastid-like organelle called the apicoplast. The apicoplast contains a genome that undergoes replication and repair through the action of a replicative polymerase (apPOL). apPOL has no direct orthologs in mammalian polymerases and is therefore an attractive antimalarial drug target. No structural information exists for apPOL, and the Klenow fragment of *Escherichia coli* DNA polymerase I, which is its closest structural homolog, shares only 28% sequence identity. Here, conditions for the crystallization of and preliminary X-ray diffraction data from crystals of *P. falciparum* apPOL are reported. Data complete to 3.5 Å resolution were collected from a single crystal (2 × 2 × 5 µm) using a 5 µm beam. The space group *P*6522 (unit-cell parameters $a = b = 141.8$, $c = 149.7$ Å, $\alpha = \beta = 90$, $\gamma = 120°$) was confirmed by molecular replacement. Refinement is in progress.

Keywords
apicoplast, DNA polymerase, Plasmodium falciparum

Disciplines
Biochemistry | Molecular Biology

Comments
This article is from *Acta Crystallographica Section F:Structural Biology Communications* 71 (2015): 333, doi: 10.1107/S2053230X15002423. Posted with permission.

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystallization and preliminary X-ray analysis of the *Plasmodium falciparum* apicoplast DNA polymerase

Morgan E. Milton,[a] Jun-yong Choe,[b] Richard B. Honzatko[a] and Scott W. Nelson[a]*

[a]Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, IA 50011, USA, and [b]Department of Biochemistry and Molecular Biology, Rosalind Franklin University of Medicine and Science, IL 60064, USA. *Correspondence e-mail: swn@iastate.edu

Received 22 December 2014
Accepted 4 February 2015

Keywords: DNA polymerase; apicoplast; *Plasmodium falciparum*.

Supporting information: this article has supporting information at journals.iucr.org/f

Infection by the parasite *Plasmodium falciparum* is the leading cause of malaria in humans. The parasite has a unique and essential plastid-like organelle called the apicoplast. The apicoplast contains a genome that undergoes replication and repair through the action of a replicative polymerase (apPOL). apPOL has no direct orthologs in mammalian polymerases and is therefore an attractive antimalarial drug target. No structural information exists for apPOL, and the Klenow fragment of *Escherichia coli* DNA polymerase I, which is its closest structural homolog, shares only 28% sequence identity. Here, conditions for the crystallization of and preliminary X-ray diffraction data from crystals of *P. falciparum* apPOL are reported. Data complete to 3.5 Å resolution were collected from a single crystal (2 × 2 × 5 µm) using a 5 µm beam. The space group $P6_522$ (unit-cell parameters $a = b = 141.8$, $c = 149.7$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$) was confirmed by molecular replacement. Refinement is in progress.

1. Introduction

Almost half of the world's population is at risk of malaria. In 2012, there were approximately 207 000 000 cases worldwide, resulting in 627 000 deaths (World Health Organization, 2013). The cause of malaria is a group of parasites from the genus *Plasmodium*, with the most common being *P. falciparum* and *P. vivax*. Drug resistance has been a constant problem and the spread of drug-resistant parasites in central Africa, India, southeast Asia and northern South America has left common antimalarial drugs such as chloroquine, artemisinin and sulfadoxin–pyrimethamine ineffective (Petersen *et al.*, 2011).

Plasmodium spp. are members of the phylum Apicomplexa, which is composed exclusively of unicellular protozoan parasites and contains several other important human and animal pathogens such as *Toxoplasma gondii* (toxoplasmosis), *Babesia bovis* (babesiosis), *Cyclospora cayetanensis* (cyclosporiasis) and *Eimeria falciformis* (coccidiosis). All of these organisms contain a unique organelle called the apicoplast that has been shown to be essential for the survival of the parasites within their host. The apicoplast arose through a secondary endosymbiosis event with red algae and can be viewed as an ancient chloroplast (Kalanon & McFadden, 2010). While it has not maintained the photosynthetic abilities of its parent organelle, the apicoplast is responsible for the execution of several biochemical processes such as the synthesis of isoprenoids, fatty acids and hemes, along with iron–sulfur cluster maturation (Seeber & Soldati-Favre, 2010)

Like the mitochondria and chloroplast, the apicoplast houses its own genome that must undergo replication and

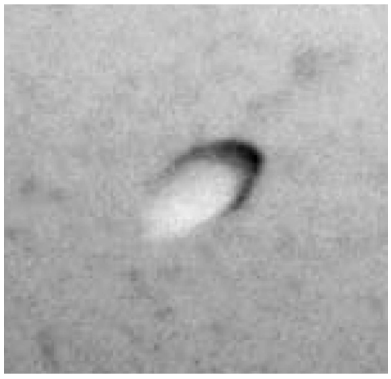


Table 1
Macromolecule-production information.

Source organism	*P. falciparum*
DNA source	GenScript
Forward primer†	5′-GGTGGTCATATGGATGAAATCACCAAAAAAT-AAATCACCAAAAAATACATCAAAG-3′
Reverse primer‡	5′-GGTGGTGGATCCTTAATCTTTGCTACCCCAG-TCTTTGCTACCCCAG-3′
Exonuclease mutation primer	5′-GATATTAAATATTGCGGCCTGAATATCCAAA-ACCAGGGGTCTGGAAGTG-3′
Expression vector	pET-28b
Expression host	*E. coli* BL21(DE3)
Complete amino-acid sequence of the construct produced	MGSSHHHHHHSSGLVPRGSHDEITKKYIKDNIIN-VDDNIIKKKDIFKLKNENNEITECAFEYFESK-KKFDDDIESRFFIINDNNYNENINLIYKDIKY-CGLDIETTGLEVFDENIRLIQIAVENYPVIIY-DMFNINKKDILDGLRKVLENKNIIKIIQNGKF-DAKFLLHNNFKIENIFDTYIASKLLDKNKNMY-GFKLNNIVEKYLNVILDKQQQNSVWNNSLLNN-NQLFYAARDSSCLLKLYKKLKEEIKKENLHIV-NDIENKCILPICDMELNGIKVDLENLQKSTNE-ILNELNIEKDNLKKKLKDENINVNSQQQVLKA-LQKNNVRDISNKLIENTSDSNLKNFLNHEEII-SLRNYRRLYKLYSAFYLKLPLHINTKTNKIHT-TFNQLKTFSGRFSSEKPNLQQIPRQKNIREIF-IPNDNNIFIIADFKQIELKIAAEITNDEIMLK-AYNNNIDLHTLTASIITKKNIPDINKEDRHIA-KAINFGLIYGMNYVNLKNYANTYYGLNMSLDQ-CLYFYNSFFEHYKGIYKWHNQVKQKRALQYST-LSNRKVIFPYFSFTKALNYPVQGTCADILKLA-LVDLYDNLKDINGKIILCVHDEIIIEVNKKFQ-EEALKILVQSMENSASYFLKKVKCEVSVKIAE-NWGSKD

† The NdeI site is underlined. ‡ The BamHI site is underlined.

repair, but its genome lacks genes coding for enzymes involved in DNA replication (Wilson *et al.*, 1996). In 2004, the *prex* (*p*lastidic DNA *r*eplication/repair *e*nzyme comple*x*) gene was identified. *Prex* encodes the putative apicoplast DNA primase, helicase and polymerase in the form of a polyprotein that has an N-terminal sequence targeting it to the apicoplast. Following import, Prex is cleaved by an unknown protease into three separate proteins (Seow *et al.*, 2005). The *P. falciparum* apicoplast DNA polymerase (apPOL) is the only known DNA polymerase targeted to the apicoplast, and we have previously demonstrated that the activity and fidelity of apPOL is consistent with that of a replicative DNA polymerase (Wingert *et al.*, 2013).

Owing to the essential nature of the apicoplast in malaria parasites, apPOL is an especially attractive target for the development of antimalarial drugs. Bioinformatics studies indicate that apPOL is of prokaryotic origin and is a member of the A-type family of DNA polymerases (Schoenfeld *et al.*, 2013). A-family DNA polymerases fall into five main groups: typical bacterial PolAs, thermophilic viruses, Aquificaceae/Hydrogenothermaceae, Apicomplexa, and other unrelated bacteria. While many members of the typical A-family DNA polymerases (for example, *Escherichia coli* PolI, *Taq* polymerase and mitochondrial DNA polymerase γ) have been extensively studied at a biochemical and structural level, very little is known regarding the polymerases from the other four distinct A-family clades. *P. falciparum* apPOL shares only 28% sequence identity with the Klenow fragment of *E. coli* PolI (PDB entry 2kfn; Brautigam *et al.*, 1999), which is its closest structurally characterized homolog (Beese *et al.*, 1993). On the other hand, the identity between different apPOLs is relatively high (84% identity between *P. falciparum* and *P. vivax*, the two primary agents of human malaria). This suggests that drugs designed to target apPOL from *P. falciparum* would be effective in treating malaria caused by other *Plasmodium* species (Supplementary Fig. S1).

Here, we report the expression, purification, crystallization and preliminary X-ray analysis of apPOL. The apicoplast DNA polymerase sequence is highly conserved among members of the *Plasmodium* genus and the structural information gathered through this study will aid in the discovery and development of antimalarial drugs.

2. Materials and methods

2.1. Protein preparation

The pET-28b vector (Novagen) containing apPOL (*P. falciparum* Prex residues 1389–2016) was used to generate a 3′-to-5′ exonuclease-deficient mutant apPOL^{exo-} (D1470N and E1470Q) as described previously (Wingert *et al.*, 2013; Table 1). The resulting vector was transformed into *E. coli* BL21(DE3) cells (Novagen), which were grown in LB medium at 310 K to an optical density of 0.7 at 600 nm. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a concentration of 0.2 m*M* to induce translation. The induced cultures were grown overnight at 291 K. The cells were harvested by centrifugation for 20 min at 3000*g* and 277 K, suspended in a minimal volume of 20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 5 m*M* imidazole (buffer 1) and stored at 193 K. The polymerase was purified at 277 K as described by Miller *et al.* (2014). The lysate from homogenized cells (EmulsiFlex-C5) was centrifuged for 1 h at 30 000*g* and 277 K. The supernatant was loaded onto a 5 ml Ni–agarose column and washed with buffer 1 and then with 20 m*M* Tris–HCl pH 8.0, 1 *M* NaCl, 25 m*M* imidazole. 1 *M* NaCl removes nucleic acids from the immobilized polymerase. 20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 150 m*M* imidazole was used to elute the polymerase. The eluted apPOL^{exo-} was concentrated to 10 mg ml^{-1} as determined spectrophotometrically using an extinction coefficient ε_{280} of 56 750 M^{-1} cm^{-1}. The sample was then subjected to size-exclusion chromatography using a 320 ml HiLoad 26/200 Superdex 200 column (GE Healthcare Biosciences) equilibrated with 20 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl. Fractions containing apPOL^{exo-} were pooled and concentrated to approximately 15 mg ml^{-1} using 50 000 NMWL Amicon Ultra-15 Centrifugal Filter Devices (Merck Millipore). High salt (400 m*M*) was used throughout the purification and crystallization process to ensure the stability of apPOL, which precipitates after prolonged periods (approximately 2 d) at lower salt concentrations. Protein purity was evaluated at each step by SDS–PAGE using Coomassie Blue stain. Control reactions indicated that our preparations were free of nuclease and phosphatase activity which would possibly interfere with polymerization assays. Polymerase activity was confirmed using methods described

previously (Miller *et al.*, 2014). Aliquots of purified apPOL^{exo-} were flash-frozen in liquid nitrogen and stored at 193 K. The freezing process does not impact on the polymerase activity or crystallization. Cells for the preparation of SeMet-apPOL^{exo-} were grown in M9 minimal medium supplemented with 50 mg SeMet 15 min prior to induction with IPTG. Labeled polymerase was purified in an identical manner to unlabeled polymerase and resulted in a slightly depressed protein yield, with unaltered catalytic activity.

Protein for crystallization was evaluated by multi-angle light scattering (MALS). The polymerase sample was passed through a 0.22 µm cellulose acetate Spin-X centrifuge tube filter (Electron Microscopy Sciences) prior to data collection. MALS data were collected by passing the eluant from a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with 20 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl into a DAWN HELEOS II light-scattering system and Optilab T-rEX refractive-index detector (Wyatt Technology). The data were analysed with the *ASTRA* software package.

2.2. Crystallization and X-ray data collection

Initial attempts employed Crystal Screen, Crystal Screen 2 and Index (Hampton Research) and vapor diffusion in 96-well sitting-drop plates. Promising outcomes were replicated and optimized by hanging-drop vapor-diffusion experiments. Polymerase (4–18 mg ml^{-1}) in 20 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl was combined with the well solutions in a 1:1 ratio to give a final droplet size of 4 µl (Table 2). Variation of the ratio of solutions and the drop size did not appear to affect the quality and the size of the crystals produced. The trays were equilibrated at room temperature. Prior to data collection, crystals were looped and immediately flash-cooled in liquid nitrogen.

Diffraction data were collected at 100 K using a MAR300 CCD detector on Advanced Photon Source beamline 23-ID-B (GM/CA-CAT). At a crystal-to-detector distance of 350 mm, 50 frames were collected using an oscillation range of 1° and a 5 µm mini-beam. Data were processed using *HKL*-3000 (Minor *et al.*, 2006). Additional analysis and molecular replacement was performed using *PHENIX* (Adams *et al.*, 2010).

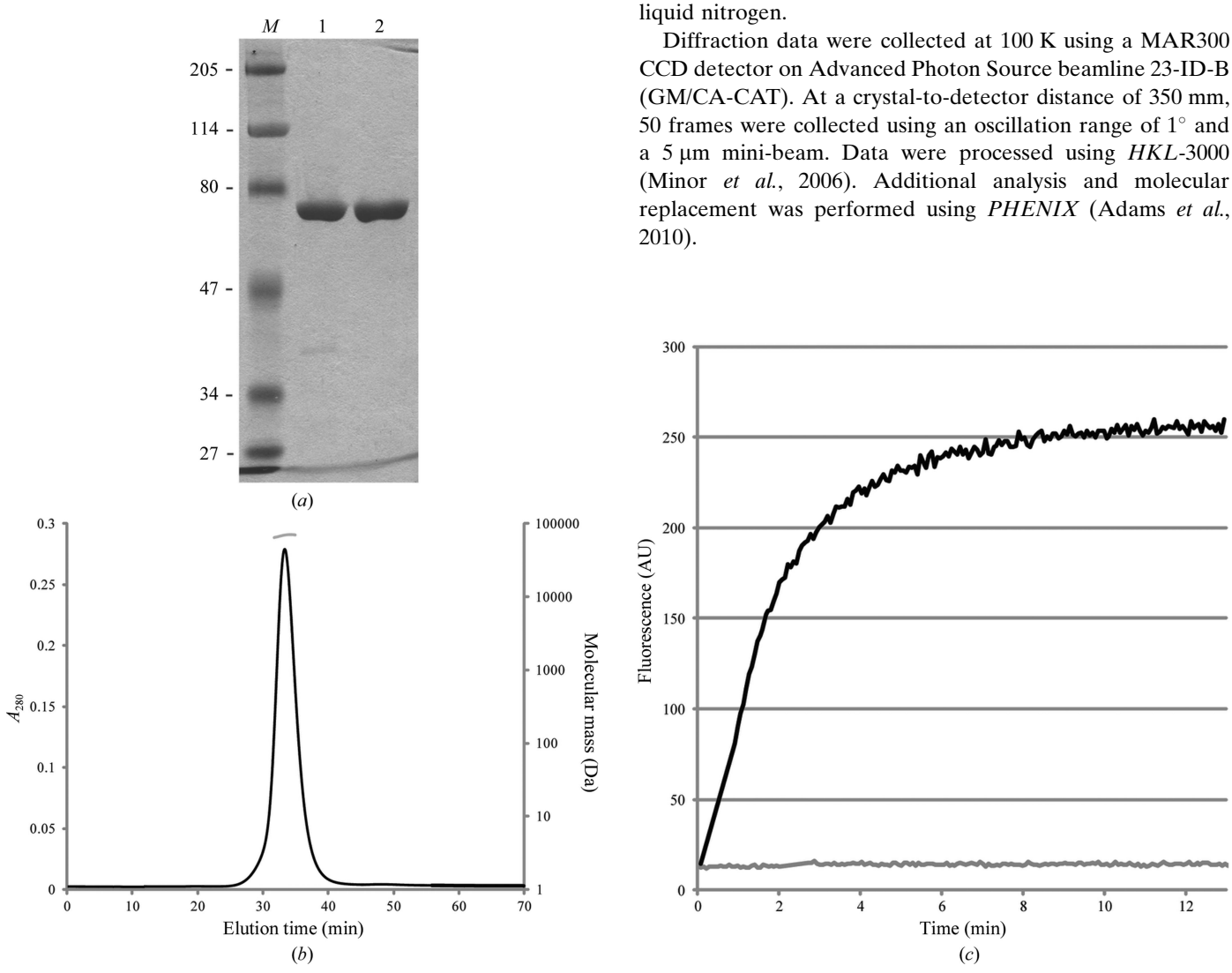


Figure 1
(*a*) SDS–PAGE analysis following purification of apPOL^{exo-}. Lane 1, pooled fractions after size-exclusion chromatography. Lane 2, peak fraction from MALS. Lane *M*, molecular-mass marker (labelled in kDa). (*b*) SEC–MALS trace for apPOL^{exo-} as monitored by absorbance at 280 nm (black). The calculated molecular mass was determined to be 68.9 kDa (grey). (*c*) Polymerase activity assay using the molecular beacon hairpin DNA substrate as previously described in Miller *et al.* (2014) without dNTPs (grey) and with dNTPs (black). The assay was performed with 50 n*M* DNA and 50 n*M* apPOL^{exo-}.

Table 2
Crystallization.

Method	Vapor diffusion
Plate type	Hanging drop
Temperature (K)	298
Protein concentration (mg ml^{-1})	4–18
Buffer composition of protein solution	20 m*M* Tris pH 8.0, 400 m*M* NaCl
Composition of reservoir solution	0.2 *M* ammonium sulfate, 0.1 *M* MES monohydrate pH 6.5, 30%(*w*/*v*) PEG monomethyl ether 5000
Volume and ratio of drop	4 µl, 1:1 protein:well solution
Volume of reservoir (µl)	500

3. Results and discussion

The yield of apPOL^{exo-}, which was 98% pure according to *ImageQ* (GE Healthcare) analysis of a Coomassie Blue-stained SDS–PAGE, was approximately 50 mg per litre of LB (Fig. 1*a*). MALS data revealed a single peak corresponding to a molecular mass of 68.9 kDa (Fig. 1*b*), which is in agreement with the gene-calculated mass of 73.9 kDa. Microcrystals grew within two weeks in droplets consisting of a 1:1 ratio of protein solution and well solution consisting of 0.2 *M* ammonium sulfate, 0.1 *M* MES monohydrate pH 6.5, 30% PEG monomethyl ether 5000 (Crystal Screen 2 condition No. 26). Crystals produced under this condition resembled short-grain rice, with approximate dimensions of 2 × 2 × 5 µm (Fig. 2). Although crystallization was highly reproducible using the premade solution from Hampton Research (regardless of the production lot), in-house solutions failed to consistently generate crystals and SeMet-labeled protein did not produce crystals under this, seeding or any other conditions.

A beam with a 20 µm cross-section provided diffraction from microcrystals to a resolution of only 8 Å; however, a beam with a 5 µm cross-section decreased the background radiation without loss of signal, revealing intensities to a resolution of 2.8 Å in a single image resulting from an exposure time of 10 s (Fischetti *et al.*, 2009). The crystals were radiation-sensitive, so the exposure time was limited to 2 s in order to collect complete data from a single crystal (at least 30 images). The crystal received an average radiation dose of 134.61 MGy as calculated by *RADDOSE-3D* (Zeldin *et al.*, 2013). This well exceeds the maximum dose of 30 MGy recommended for macromolecular crystallography (Owens *et al.*, 2006). Radiation damage was visually apparent over the 50 frames of data collected as the resolution decreased from just beyond 3.5 Å to approximately 4.5 Å. Also influencing the strategy for data collection was a frequency of one in 60 flash-cooled crystals that exhibited high-resolution ordered diffraction. The aforementioned circumstances resulted in complete data, but only to a resolution of 3.5 Å largely owing to the short exposure time of 2 s and radiation damage (Fig. 3*a*, Table 3). Data reduction initially assumed a trigonal lattice, but further analysis indicated space group $P6_122$ (or its enantiomorph), with unit-cell parameters $a = b = 141.8$, $c = 149.7$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. The Matthews coefficient of 2.94 Å^3 Da^{-1} is consistent with one polymerase molecule per asymmetric unit and a solvent content of 58.2%. Data were originally scaled with a resolution cutoff at 3.5 Å (Table 3). The inclusion of all available diffraction data resulted in qualitative improvement of the electron-density map, but the $I/\sigma(I)$ values for each shell suggest that the nominal resolution of the data is 3.5 Å.

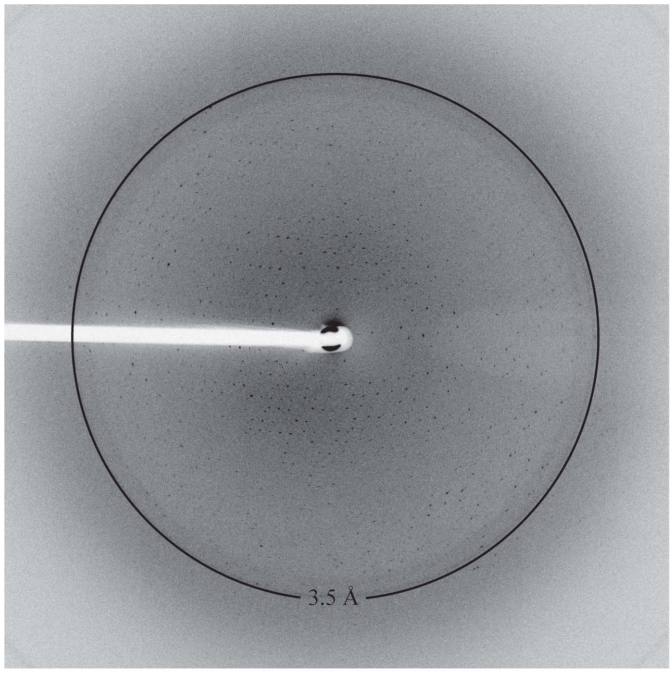
Figure 2
A typical apo apPOL^{exo-} microcrystal with dimensions of approximately 2 × 2 × 5 µm.

Table 3
Data collection and processing.

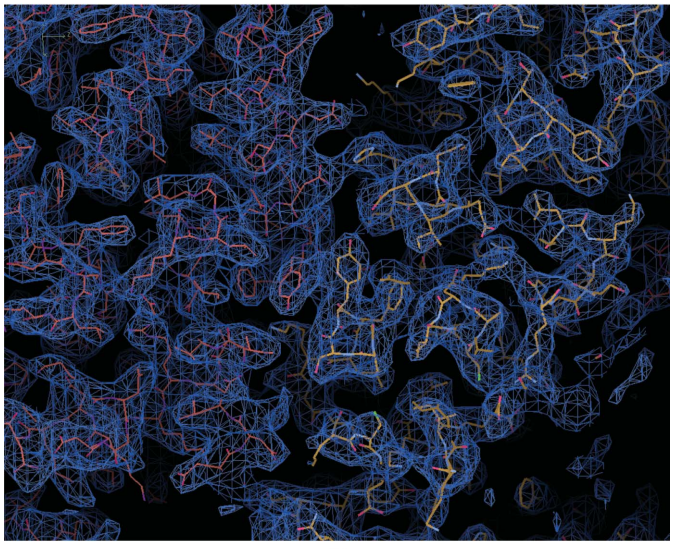
Values in parentheses are for the outer shell.

Wavelength (Å)	1.0332
Exposure time per image (s)	2.0
Space group	$P6_522$
Unit-cell parameters (Å, °)	$a = b = 141.8$, $c = 149.7$, $\alpha = \beta = 90$, $\gamma = 120$
Resolution range (Å)	50.0–3.5 (3.63–3.50)
Total No. of reflections	11190 (1081)
Completeness (%)	95.4 (95.1)
Multiplicity	3.6 (3.5)
$\langle I/\sigma(I)\rangle$	4.9 (1.75)
R_{merge}	0.151 (0.390)

Initial molecular replacement was performed using *Phaser* (McCoy *et al.*, 2007) and resulted in a translation-function *Z*-score (TFZ) of 4.4, a rotation-function *Z*-score (RFZ) of 3.1 and a log-likelihood gain (LLG) of 49.555 when an unmodified Klenow fragment (PDB entry 2kfn; Brautigam *et al.*, 1999) was used as the molecular-replacement model. Use of only the polymerase domain of the Klenow fragment (residues 519–928), which has 31% sequence identity to apPOL^{exo-} but alone only accounts for 63% of the total apPOL^{exo-} molecule, resulted in a molecular-replacement solution (electron density revealing right-handed α-helices as shown in Fig. 3*b*) only in space group $P6_522$, with a TFZ of 6.2, an RFZ of 3.9 and an LLG of 63.248 (Adams *et al.*, 2010). Subsequent use of *MR-Rosetta* (DiMaio *et al.*, 2011) using the unmodified PDB entry 2kfn (complete with DNA and ions) as the model independently indicated the same space group and a partial model with approximately 450 of the 628 expected residues placed in electron density. *MR-Rosetta* failed to model 67 residues at the N-terminus which make up part of the apPOL^{exo-} exonuclease domain. Over half of these residues are found



(*a*)



(*b*)

Figure 3
(*a*) X-ray diffraction pattern of an apPOL^{exo-} microcrystal produced from a 2 s exposure. The first image collected from a data set of 50 frames is shown. Weak diffraction extends to 2.8 Å resolution, with stronger diffraction extending to approximately 3.5 Å resolution as shown by the black circle. (*b*) Density map and model from early refinement after molecular replacement. The density shows clear right-handed helices and crystal lattice contacts. Coordinates are shown in yellow and symmetry atoms are shown in pink.

exclusively in members of the *Plasmodium* genus and appear to be important in producing an active polymerase (Wingert *et al.*, 2013). Density is clearly present for these residues and a model is currently being manually built.

Several circumstances were unfavorable for structure determination: radiation-sensitive microcrystals, most of which failed to diffract in an ordered manner, low sequence identity to known structures and failure to grow SeMet-derivative crystals. Nonetheless, the advanced capabilities of synchrotron facilities resulted in complete data and successful structure determination. Further development of the exonuclease domain will require manual building in regions of low or absent sequence identity. Preliminary work, however, clearly supports the thumb–palm–finger structure observed in other A-type polymerases.

Acknowledgements

We thank A. J. Milton for assistance with X-ray data collection and thank the beamline staff at GM/CA-CAT, Argonne National Laboratories. This work was funded by grants from Iowa State University (Signature Research Initiative, 2014) and The Roy J. Carver Charitable Trust (grant No. 10-3603).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Beese, L. S., Derbyshire, V. & Steitz, T. A. (1993). *Science*, **260**, 352–355.
Brautigam, C. A., Sun, S., Piccirilli, J. A. & Steitz, T. A. (1999). *Biochemistry*, **38**, 696–704.
DiMaio, F., Terwilliger, T. C., Read, R. J., Wlodawer, A., Oberdorfer, G., Wagner, U., Valkov, E., Alon, A., Fass, D., Axelrod, H. L., Das, D., Vorobiev, S. M., Iwaï, H., Pokkuluri, P. R. & Baker, D. (2011). *Nature (London)*, **473**, 540–543.
Fischetti, R. F., Xu, S., Yoder, D. W., Becker, M., Nagarajan, V., Sanishvili, R., Hilgart, M. C., Stepanov, S., Makarov, O. & Smith, J. L. (2009). *J. Synchrotron Rad.* **16**, 217–225.
Kalanon, M. & McFadden, G. I. (2010). *Biochem. Soc. Trans.* **38**, 775–782.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Miller, M. E., Parrott, E. E., Singh, R. & Nelson, S. W. (2014). *J. Biomol. Screen.* **19**, 966–972.
Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.
Owens, R. L., Rubiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
Petersen, I., Eastman, R. & Lanzer, M. (2011). *FEBS Lett.* **585**, 1551–1562.
Seeber, F. & Soldati-Favre, D. (2010). *Int. Rev. Cell Mol. Biol.* **281**, 161–228.
Seow, F., Sato, S., Janssen, C. S., Riehle, M. O., Mukhopadhyay, A., Phillips, R. S., Wilson, R. J. & Barrett, M. P. (2005). *Mol. Biochem. Parasitol.* **141**, 145–153.
Wilson, R. J., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., Whyte, A., Strath, M., Moore, D. J., Moore, P. W. & Williamson, D. H. (1996). *J. Mol. Biol.* **261**, 115–172.
Wingert, B. M., Parrott, E. E. & Nelson, S. W. (2013). *Biochemistry*, **52**, 7723–7730.
World Health Organization (2013). *World Malaria Report 2013*. Geneva: World Health Organization. http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/.
Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Appl. Cryst.* **46**, 1225–1230.