



# Structure of *Plasmodium falciparum* orotate phosphoribosyltransferase with autologous inhibitory protein–protein interactions

Shiva Kumar, Kalyanaraman Krishnamoorthy, Devaraja G. Mudeppa and Pradipsinh K. Rathod\*

Received 15 January 2015

Accepted 17 March 2015

Edited by W. N. Hunter, University of Dundee, Scotland

**Keywords:** pyrimidine; *de novo*; salvage; PRPP; OMP; prodrugs; peptide inhibition.

**PDB reference:** orotate phosphoribosyltransferase, 4fym

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

Department of Chemistry, University of Washington, Seattle, WA 98195, USA. \*Correspondence e-mail: rathod@chem.washington.edu

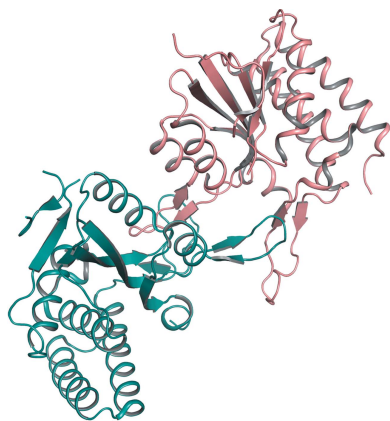
The most severe form of malaria is caused by the obligate parasite *Plasmodium falciparum*. Orotate phosphoribosyltransferase (OPRTase) is the fifth enzyme in the *de novo* pyrimidine-synthesis pathway in the parasite, which lacks salvage pathways. Among all of the malaria *de novo* pyrimidine-biosynthesis enzymes, the structure of *P. falciparum* OPRTase (*PfOPRTase*) was the only one unavailable until now. *PfOPRTase* that could be crystallized was obtained after some low-complexity sequences were removed. Four catalytic dimers were seen in the asymmetric unit (a total of eight polypeptides). In addition to revealing unique amino acids in the *PfOPRTase* active sites, asymmetric dimers in the larger structure pointed to novel parasite-specific protein–protein interactions that occlude the catalytic active sites. The latter could potentially modulate *PfOPRTase* activity in parasites and possibly provide new insights for blocking *PfOPRTase* functions.

## 1. Introduction

Historically, malaria has caused sustained morbidity and mortality in the human population. Many believe that the disease has imposed the most potent selective pressure in the recent evolution of the human genome (Kwiatkowski, 2005). The problem continues. In the last 20 years, malaria-related deaths increased from <1 million to 1.8 million in 2004 before decreasing to the current estimate of about 600 000 per year (Murray *et al.*, 2012). As encouraging as the recent numbers are, sustained management of malaria morbidity and mortality will require new treatments, including new drug targets (Dhanawat *et al.*, 2009; Gamo *et al.*, 2010).

The most severe form of malaria is caused by the obligate parasite *Plasmodium falciparum*, particularly in Africa. These intracellular parasites can be selectively attacked by inhibitors of essential enzymes. Potent and selective antimalarials have been developed to target *P. falciparum* dihydroorotate dehydrogenase, the fourth enzyme in the *de novo* pyrimidine-biosynthesis pathway (Baldwin *et al.*, 2005; Phillips *et al.*, 2008; Coteron *et al.*, 2011; Deng *et al.*, 2014; Gujjar *et al.*, 2011). Recently, one triazolopyrimidine derivative from this initiative (DSM265) has advanced to early human testing (<http://www.mmv.org/research-development/rd-portfolio>).

The pipeline for antimalarials could benefit from additional good targets and strategies that revolve around *de novo* pyrimidine metabolism in the parasite. Specifically, *P. falciparum* orotate phosphoribosyltransferase (*PfOPRTase*), the fifth enzyme in the indispensable *de novo* pyrimidine-synthesis pathway of the parasite (Reyes *et al.*, 1982), is of special interest. OPRTase catalyses the conversion of orotate and



5-phospho-D-ribose 1-pyrophosphate (PRPP) into orotidine 5'-monophosphate (OMP) (Fig. 1). While OPRase and OMP decarboxylase, the next enzyme in the *de novo* pyrimidine-synthesis pathway, are fused into one bifunctional enzyme in humans, they exist as two different gene products in the malarial parasite (Rathod & Reyes, 1983). OPRase is a dimeric enzyme with two active-site pockets, one on each protomer. Once the substrate has docked, two different loops, one from each protomer, cover the bound substrate to shield it from solvent (Scapin *et al.*, 1994; Henriksen *et al.*, 1996). The OPRase protein sequence is 100% identical between the two strains of *P. falciparum*, 3D7 and IT, as seen in PlasmoDB (Aurrecochea *et al.*, 2009). Additionally, among the 203 single-nucleotide polymorphism (SNP) data sets in PlasmoDB, *PfOPRTase* contains 30 SNPs. Two of these are stop-codon SNPs, 25 are nonsynonymous and three are synonymous. Previously, *PfOPRTase* has been used as a conduit for potent and selective prodrugs directed at downstream enzymes in pyrimidine biosynthesis (Rathod *et al.*, 1989; Gómez & Rathod, 1990). Potent direct inhibitors of *PfOPRTase* remain to be discovered.

The availability of an atomic resolution structure of *PfOPRTase* could help to reveal the precise layout of the substrate and prodrug-binding pocket, which in turn could help with the design of selective inhibitors. Unfortunately, like many other important malarial proteins, *PfOPRTase* has eluded structure determination. Molecular features unique to *Plasmodium* genes and gene products, such as AT-rich tracks, polyasparagine repeats and the binding of protein by cognate mRNA (Zhang & Rathod, 2002), can interfere with the heterologous expression of malaria proteins (Schneider *et al.*, 2005) and can also promote protein aggregation (Singh *et al.*, 2004). This has resulted in an underrepresentation of structures from *Plasmodium* proteins in the Protein Data Bank (wwPDB) that is even worse than for that of membrane proteins (Vedadi *et al.*, 2007; Mehlin *et al.*, 2006).

In 2012, a preliminary publication reported the crystallization of *PfOPRTase* (Takashima *et al.*, 2012). However, to date there is no public information on the underlying quality of the data from this crystal or an OPRase structure. Here, we report the optimum expression and crystallization of

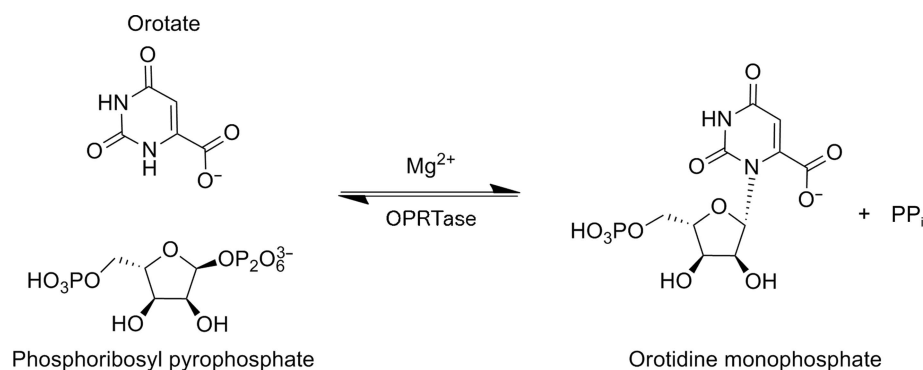
*PfOPRTase*. Our structure reveals novel parasite-specific protein-protein interactions that are of possible importance in malaria biology. In addition, we can map the positions of active-site amino-acid residues that differ between the host and the parasite OPRases and this should facilitate future new structure-based drug-development strategies.

## 2. Materials and methods

### 2.1. Protein production

Conventional methods for heterologous protein over-expression were not useful for producing *PfOPRTase* in *Escherichia coli* in the quantities required for screening crystallization conditions. Codon optimization to compensate for the unusual AT-rich *Plasmodium* genes failed to give desirable yields in *E. coli*. The expression of functional *PfOPRTase* was aided by obligatory protein production in auxotrophic bacteria, as previously performed to obtain *P. falciparum* serine hydroxymethyltransferase (Alfadhli & Rathod, 2000). Bacterial strain JW3617(DE3) from the *E. coli* stock centre (Yale University) was grown in M9 minimal medium (Sigma, St Louis, Missouri, USA) with *PfOPRTase* function provided on a pET-28a vector (Merck KGaA, Darmstadt, Germany). Since the JW3617(DE3) cells lacked a native OPRase gene, their survival depended on retaining a functional *PfOPRTase*. This system allowed us to explore different constructs of *PfOPRTase* for improved yield while retaining function. We hypothesized that diffraction-quality crystals may be best obtained by reducing the packing disorder caused by surface entropy. Various lengths of the polyasparagine repeats, which are unique to *P. falciparum* and do not appear to contribute to catalytic function, were individually deleted. In addition, a series of N-terminal truncations were also tested. The ability to rescue OPRase function in JW3617 (DE3) cells in M9 minimal medium agar was used to identify functionally active constructs. Using this criterion, in addition to full-length *PfOPRTase*-6×His, *PfOPRTase* 1–218(Δ37–58)-H6 with a C-terminal six-His tag appeared to be promising (Fig. 2a).

*PfOPRTase* protein was expressed in JW3617(DE3) cells grown in M9 medium supplemented with 40 μg ml<sup>-1</sup>



**Figure 1**

The orotate phosphoribosyltransferase (OPRTase) reaction. OPRases catalyze the conversion of orotate and 5-phospho-D-ribose 1-pyrophosphate (PRPP) to orotidine 5'-monophosphate (OMP).

kanamycin. The cells were grown at 310 K until the OD<sub>600</sub> reached 0.7–0.8. Subsequently, the cells were transferred to 291 K and, after 30 min, *PfOPRTase* production was induced with 100 μM IPTG. The cells were harvested 16 h after induction for protein extraction and purification. *PfOPRTase* 1–218(Δ37–58)-H6 (or full-length *PfOPRTase*-6×His) was released by lysing the host *E. coli* cells in buffer A (50 mM Tris–HCl pH 7.3, 150 mM NaCl, 25 mM imidazole, 5% glycerol, 2 mM β-mercaptoethanol) using a sonicator. The lysate was clarified by centrifugation (18 000g, 277 K, 20 min). The protein supernatant was passed through a 5 ml HisTrap FF column (GE Healthcare Biosciences, Pittsburgh, Pennsylvania, USA), after which the column was washed with ten volumes of buffer A. Adsorbed protein was then eluted off the column by passing five column volumes of elution buffer

(buffer A with 250 mM imidazole–HCl). An Amicon concentrator (Merck KGaA, Darmstadt, Germany) was used to equilibrate the protein with low-salt buffer B (50 mM Tris–HCl pH 7.3, 50 mM NaCl, 5% glycerol, 2 mM β-mercaptoethanol) before anion-exchange chromatography. The protein in buffer B was then loaded onto a 5 ml HiTrap Q-Sepharose FF column (GE Healthcare) and purified using a gradient from buffer B to buffer C (50 mM Tris–HCl pH 7.3, 1000 mM NaCl, 5% glycerol, 2 mM β-mercaptoethanol). The eluted protein was then injected onto an S200 XK16/60 size-exclusion chromatography column (GE Healthcare Biosciences) for final purification. The column was run on an ÄKTA FPLC system (GE Healthcare Biosciences) using buffer D (20 mM Tris–HCl pH 7.3, 100 mM NaCl, 0.5 mM TCEP). The final protein was concentrated to 30 mg ml<sup>-1</sup> before crystallization.

2.2. *In vitro* catalytic activity of purified protein

The final purified protein was tested for catalytic activity before use in crystallization experiments. Purified *PfOPRTase* 1–218(Δ37–58)-H6 (50 nM) was mixed with the substrates phosphoribosyl pyrophosphate (2 mM) and orotic acid (100 μM) in a buffer consisting of 50 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM TCEP. The decrease in absorbance was measured at 296 nm (Mudeppa & Rathod, 2013). Orotic acid utilization was observed with purified *PfOPRTase*, but not with BSA or with no protein added (Fig. 2b). The catalytic activity confirmed that the heterologous *PfOPRTase* 1–218(Δ37–58)-H6 protein was purified in an active form.

2.3. Protein crystallization

The protein was concentrated to 30 mg ml<sup>-1</sup> before crystallization. Full-length *PfOPRTase*-6×His did not crystallize under the conditions tested, but *PfOPRTase* 1–218(Δ37–58)-H6 did. Diffraction-quality crystals were grown in hanging-drop plates (Hampton Research) in 100 mM ammonium sulfate, 30% PEG 3350 at 277 K. Crystals were cooled in liquid nitrogen using the reservoir solution supplemented with 20% ethylene glycol. Diffraction data were collected on beamline 23-ID-B at the Advanced Photon Source (APS), Argonne National Laboratories, Argonne, Illinois, USA.

2.4. Structure determination

*PfOPRTase* 1–218(Δ37–58)-H6 crystallized in the primitive orthorhombic space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>, with unit-cell parameters *a* = 114.769, *b* = 152.487, *c* = 167.755 Å (Table 1). Estimating the precise space group was not trivial at the beginning of structure determination. The unit-cell volume was very large for a 31.37 kDa protein. Based on the Matthews coefficient, we expected eight to ten molecules in the asymmetric unit (Matthews, 1968; Kantardjieff & Rupp, 2003). The potentially large number of expected protein molecules in the asymmetric unit (*n*<sub>mol</sub>) made it difficult to conclusively predict their exact number *a priori*. We arrived at *n*<sub>mol</sub> = 8 by successively placing the molecules correctly in the asymmetric unit. A high number of molecules in an asymmetric unit can sometimes lead to a

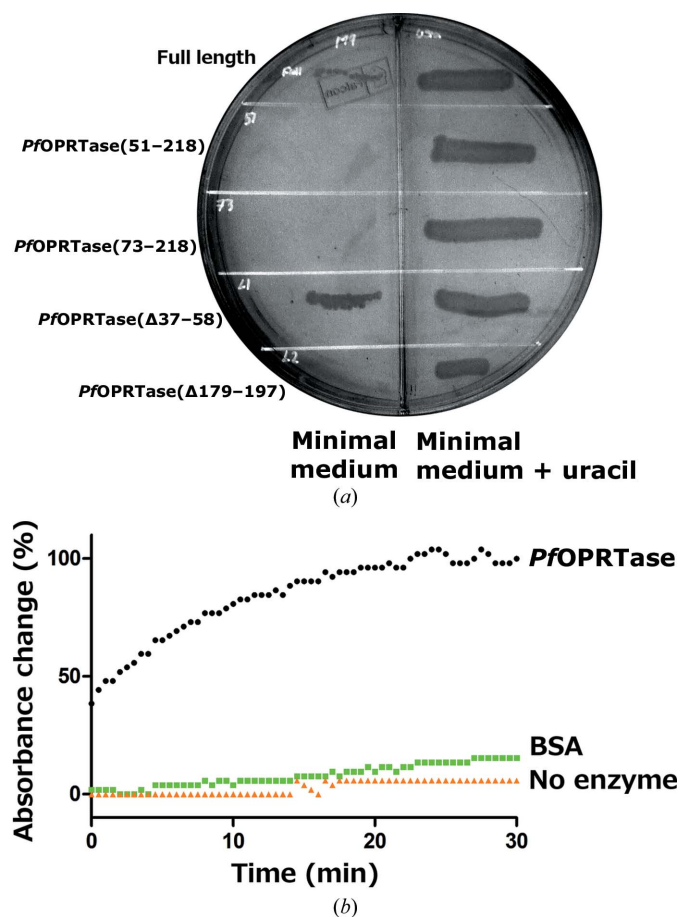


Figure 2 Activity assays validating functional *PfOPRTase* constructs. (a) Select *PfOPRTase* constructs can support the growth of *OPRTase*-deficient *E. coli*. Several variants of *PfOPRTase*-coding sequences [full length, *PfOPRTase*(51–218), *PfOPRTase*(73–218), *PfOPRTase*(Δ37–58) and *PfOPRTase*(Δ179–197)] in plasmid pET-28a were transformed into *E. coli* JW3617 cells. The cells were grown in M9 minimal medium (left half of the plate) or M9 medium supplemented with uracil (right half of the plate). Full-length *PfOPRTase* and partially deleted *PfOPRTase* (Δ37–58) successfully rescued *E. coli* JW3617 cells from *OPRTase* deficiency, but not the others. (b) Time-dependent reduction in absorbance as orotic acid is converted to orotidylate (OMP) by functional *PfOPRTase*. The change in the absorbance by orotic acid (at 296 nm) is shown in the presence of *PfOPRTase* (black), bovine serum albumin (green) or no added enzyme (orange).

**Table 1**  
Data-collection and refinement statistics.

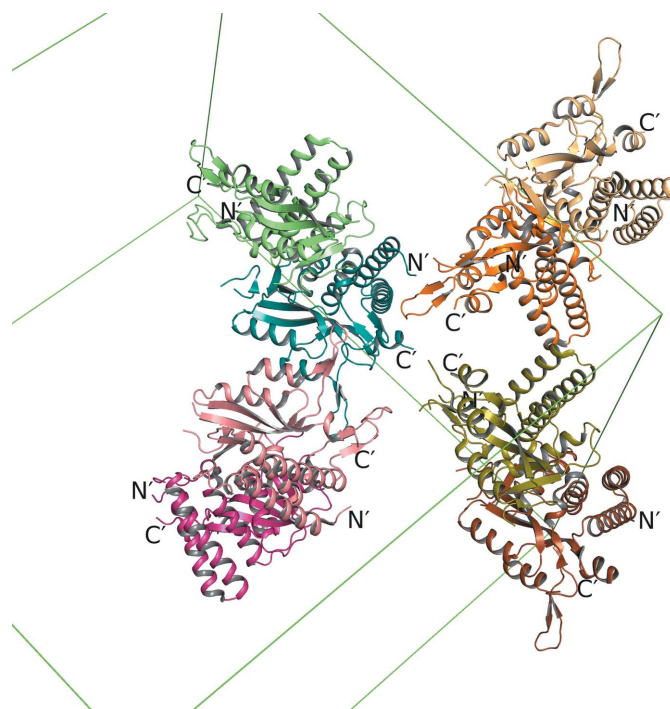
Data collection	
Diffraction source	Beamline 23-ID-B, APS
Wavelength (Å)	1.09
Temperature (K)	100
Detector	MAR Mosaic 300 mm CCD
Crystal-to-detector distance (mm)	320
Rotation range per image (°)	0.5
Space group	$P2_12_12_1$
$a, b, c$ (Å)	114.77, 152.49, 167.76
$\alpha, \beta, \gamma$ (°)	90, 90, 90
Resolution range (Å)	50.0–2.6 (2.63–2.60)
Total No. of reflections	1819537
No. of unique reflections	90605 (2832)
Completeness (%)	98.9 (93.9)
Multiplicity	5.6 (4.6)
$\langle I/\sigma(I) \rangle$	22.6 (1.97)
$R_{\text{r.i.m.}}^\dagger$	0.081 (0.711)
$R_{\text{merge}}$	0.073 (0.629)
Overall $B$ factor from Wilson plot (Å <sup>2</sup> )	51.16
Refinement statistics	
Resolution range (Å)	40.47–2.6 (2.66–2.59)
Completeness (%)	98.5
No. of reflections, working set	85965 (5765)
No. of reflections, test set	4588 (311)
Final $R_{\text{cryst}}$	0.197 (0.312)
Final $R_{\text{free}}$	0.218 (0.345)
No. of non-H atoms	
Protein	14453
Ion	165
Water	264
Total	14882
R.m.s. deviations	
Bonds (Å)	0.016
Angles (°)	1.839
Average $B$ factors (Å <sup>2</sup> )	
Protein	62.7
Ion	74.1
Water	47
Clashscore	3.81
Ramachandran plot	
Favored regions (%)	98.8
Additionally allowed (%)	1.1
Outliers (%)	0.1

$^\dagger R_{\text{r.i.m.}} = R_{\text{merge}}[N/(N-1)]^{1/2}$ , where  $N$  is the data multiplicity.

noncrystallographic symmetry axis within the asymmetric unit that is parallel to a crystallographic symmetric axis. The resulting possibility of pseudo-translation was tested by drawing a native Patterson map, and this gave a single off-origin peak of height 46% at fractional coordinates 0, 0.5, 0.5. Given that the vector, at a distance of 113 Å from the origin, was 46% of the origin, it was concluded that it was not a crystallographic  $A$ -centering vector but a noncrystallographic symmetry vector. Systematic absences of a  $2_1$  screw operator on either  $b, k = 2n$ , or on  $c, l = 2n$ , can satisfy the condition for systematic absences for  $A$ -centering, with  $k$  and  $l$  both being odd or even. This made it hard to discern whether a  $2_1$  screw was indeed present in the crystal at the  $b$  axis, at the  $c$  axis or both. In addition, molecular-replacement searches were expected to give a very low signal-to-noise ratio owing to the small scattering volume of the search model in relation to the asymmetric unit. Furthermore, the best molecular-replacement (MR) search model, *Saccharomyces cerevisiae* OPRt (PDB entry 2pry; González-Segura *et al.*, 2007), had 33% sequence identity, which by itself could reduce the signal-to-

noise ratio for the MR searches. A mammalian OPRtase domain from PDB entry 2wns (Structural Genomics Consortium, unpublished work) had a sequence identity of 24% using a *BLAST* alignment, which made the *S. cerevisiae* structure best suited as an MR search model.

Nevertheless, MR searches were carried out in all of the possible space groups of the point group  $P222$ . Using data processed and scaled in *HKL-2000* (Otwinowski & Minor, 1997), initial MR searches in *MOLREP* (Vagin & Teplyakov, 2010), *AMoRe* (Navaza, 1994) and *Phaser* (McCoy *et al.*, 2007) as part of the *CCP4* suite (v.6.3.0) failed to provide any solution with good signal when compared with a random position in the asymmetric unit. Next, all of the solutions provided by *Phaser* (v.2.5.2) while searching for the first ensemble were visually inspected in *PyMOL* (v.1.5.0.4; Schrödinger). Two solutions which looked like possible catalytic dimers were selected and fixed. The search was then repeated to find further solutions. The signal-to-noise ratio in the *Phaser* output improved as more molecules were chosen manually by visual inspection of pairs of polypeptides that seemingly formed a catalytic dimer. They were fixed in the asymmetric unit. Further solutions were automatically found by *Phaser*. After finding all eight molecules, *phenix.autobuild* (Adams *et al.*, 2010) was used to construct the initial structure. Further building was performed manually using *Coot* (Emsley



**Figure 3**  
*P. falciparum* OPRtase crystal packing. *PfOPRtase* packs with eight molecules in the asymmetric unit. All of the chains are shown in different colors: chains *A* and *B* in light orange and dark orange, chains *C* and *D* in light green and blue-green, chains *E* and *H* in light pink and dark pink and chains *F* and *H* in olive and brown, respectively. Catalytic dimers are colored in different shades of the same color. The unit cell is shown with green lines. The N- and C-termini of all chains are labeled. The rotation axis of the catalytic dimer is nearly parallel to the crystallographic axis, which caused the observed pseudo-translation.



*et al.*, 2010), while iterative refinement was performed using *REFMAC* (Murshudov *et al.*, 2011).

### 3. Results

#### 3.1. Crystal packing

In our crystal, *PfOPRTase* 1–218( $\Delta$ 37–58)-H6 packs as a set of four catalytic dimers, leading to eight molecules in the asymmetric unit (Fig. 3). None of the protomers in the four dimers are related by a crystallographic twofold. Instead, they are related by a near-perfect noncrystallographic twofold rotation (denoted here as  $\kappa$ ) with  $\kappa$  ranging from 179.3 to 179.8°.

#### 3.2. Two sites in an obligatory dimer

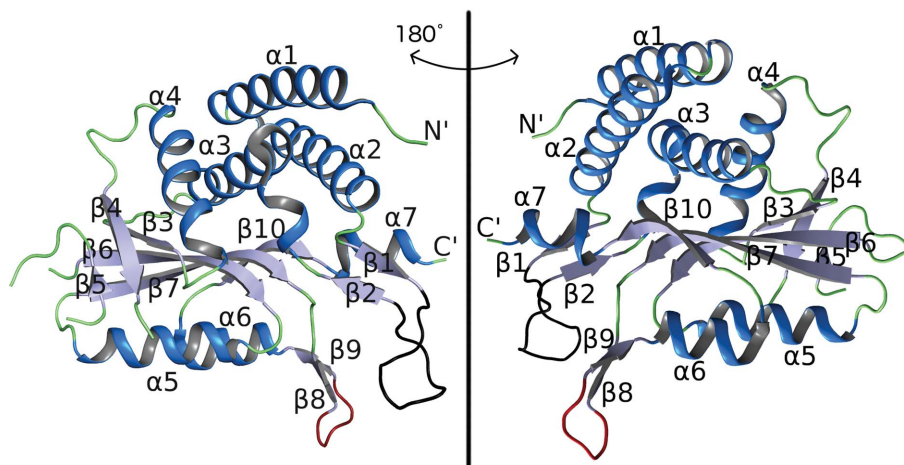
In the present crystal, the individual *PfOPRTase* molecules possess the conserved  $\alpha/\beta$  topology also seen in other OPRTases (Scapin *et al.*, 1994; Henriksen *et al.*, 1996; Liu *et al.*, 2010). There are seven  $\alpha$ -helices and ten  $\beta$ -strands in the structure, in addition to one *cis*-peptide (Fig. 4). The *PfOPRTase* structure conforms with the conserved PRTase fold, having a 321456 arrangement of  $\beta$ -sheets. The antiparallel strand 3 of this sheet is not tightly ordered into a distinctive  $\beta$ -strand in the solved structure.

At the core, the *PfOPRTase* 1–218( $\Delta$ 37–58)-H6 molecules pack as dimers. This is consistent with OPRTase structures from other organisms, which have shown that a dimer is essential to form catalytic active sites (Scapin *et al.*, 1994; Henriksen *et al.*, 1996; Liu *et al.*, 2010). The catalytic dimers have an average buried surface area of 1500 Å<sup>2</sup> (Fig. 5a). Among the database of OPRTase structures from other source organisms in the wwPDB, *PfOPRTase* is the most closely related to that from *S. cerevisiae* (*ScOPRTase*) in sequence as well as in structure. *PfOPRTase* differs from *ScOPRTase* with an r.m.s.d. of 2.1 Å (calculated by SSM Superpose in *Coot*; Emsley *et al.*, 2010), whereas with the human enzyme (PDB

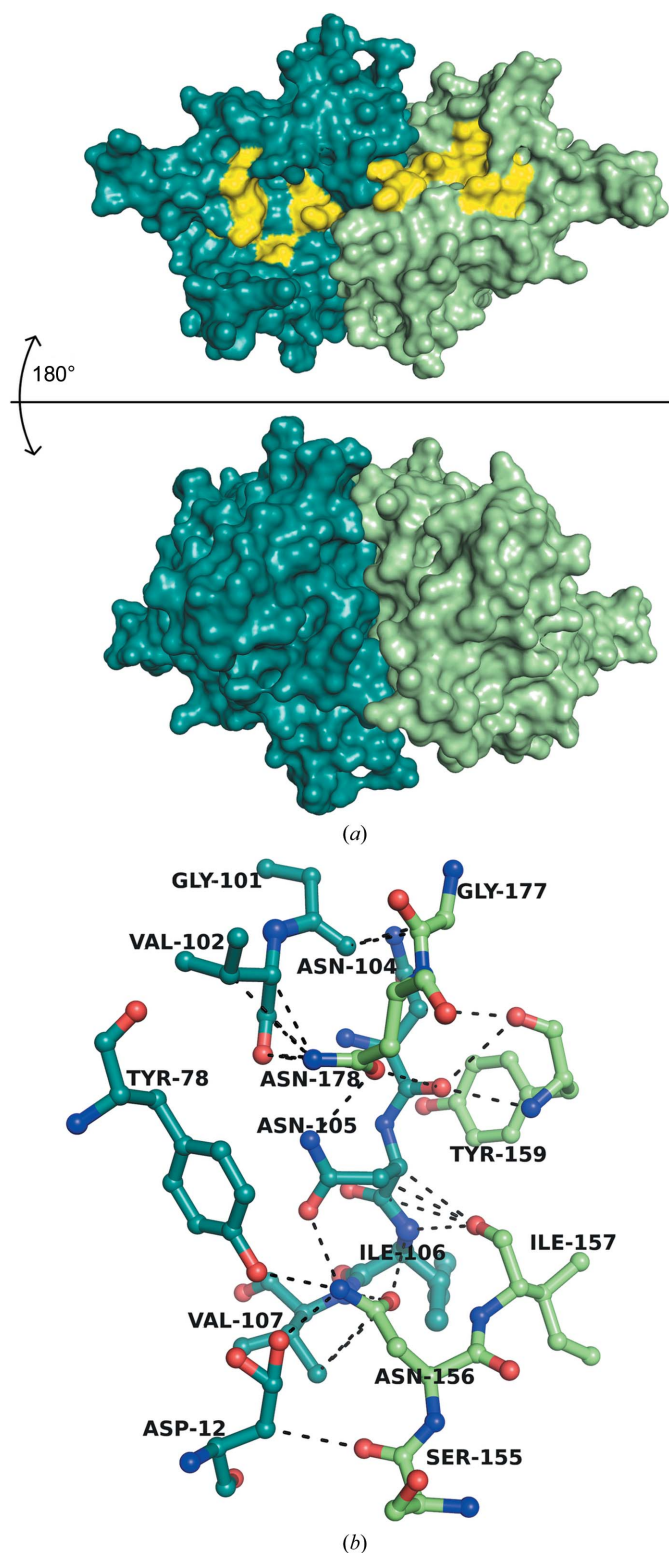
entry 2wns) the value is 2.6 Å. The interfaces between all of the four catalytic dimers in the *PfOPRTase* asymmetric unit are identical. The interface between chains *C* and *D* (PDB entry 4fym) is used in the present description since it has the best electron density of the four dimers. The catalytic dimer in our *PfOPRTase* structure is held together exclusively by hydrogen bonds (Fig. 5b). In particular,  $\alpha$ 4 and the loop residues leading up to it make dominant interactions that hold the catalytic homodimer together. Almost all of the surface residues of  $\alpha$ 4, which is comprised of the residues <sup>136</sup>KGIPMVSLTSHFLFE<sup>150</sup>, are buried at this dimer interface. The  $\alpha$ 4 helix then leads to  $\beta$ 4, which also makes predominant contacts at the interface.

#### 3.3. Substrate-binding sites

Each of the two active sites of *PfOPRTase* receives amino-acid contributions from the two polypeptides of the dimer. Beyond the dimer-interacting amino acids discussed above, there are eight disordered amino acids in the *PfOPRTase* structure. These eight amino acids have previously been observed to be disordered in the apo form of the yeast enzyme (PDB entry 2ps1; González-Segura *et al.*, 2007). These eight disordered amino acids (<sup>165</sup>EKKEYGDK<sup>172</sup> from each subunit of the dimer in *PfOPRTase*) are expected to wrap around PRPP based on available structures with substrate. PRPP is also bound to the dimeric partner, and together the two polypeptides contribute to the catalytic reaction (Henriksen *et al.*, 1996). This stretch of eight amino acids leads into  $\beta$ 5, which forms the only antiparallel strand of the conserved OPRTase  $\beta$ -sheet. Gly84 and Gly101 play an important role as hinge residues to provide flexibility to the hood domain. The hood domain moves on top of the substrate to sequester it from free solvent. The hood domain contains Phe97, which is known to form a  $\pi$ -stacking interaction with the substrate orotate (González-Segura *et al.*, 2007). Gly101 and Val102 at one hinge of the hood are part of a  $3_{10}$ -helix and they interact with  $\beta$ 5 at the dimeric interface. This interaction holds the hood



**Figure 4** Structure of an individual subunit of *PfOPRTase*. The two views depict a 180° horizontal rotation. All of the helices and strands are numbered according to their order of occurrence from the N-terminus of the protein subunit. Helices are shown in dark blue,  $\beta$ -strands in light blue and most loops in green. The loop in red between  $\beta$ 8 and  $\beta$ 9 (<sup>247</sup>NENNQ<sup>251</sup>) makes interdimer contacts (see §3 and Fig. 7). The loop in black between  $\beta$ 1 and  $\beta$ 2 (<sup>84</sup>GEFILKSKRKSN<sup>95</sup>) forms a ‘hood’ that is expected to wrap around the substrates (see §3 and Fig. 6)



**Figure 5**

Description of the OPRTase obligate dimer. (a) Two subunits contribute to each of the two active sites of the *Pf*OPRTase dimer. The two views, 180° apart, represent the surfaces of the two protomers (colored blue and green) of the *Pf*OPRTase dimer with a large buried surface area of 1500 Å<sup>2</sup>. The substrate-binding site on one face of the dimer is colored yellow. (b) The catalytic dimer of *Pf*OPRTase is held together primarily by interprotomer hydrogen bonds, with contributions from one protomer shown in dark green and those from the other in light green.

domain in the correct orientation to make the  $\pi$ -stacking possible.

The sulfates in the *Pf*OPRTase structure provide excellent clues to the detailed positioning of the natural phosphate-containing substrates and products of *Pf*OPRTase. There are 2–6 sulfate ions originating from the crystallization conditions associated with each polypeptide in the core dimer. They occupy the same positions in each of the eight chains in the asymmetric unit. The exact number of sulfates varies between the eight *Pf*OPRTase chains owing to differences in the quality of the electron density in the individual polypeptides. The two most interesting sulfate ions occupy the positions where the phosphate and pyrophosphate groups of the PRPP substrate are expected to bind. From the homologous structures of yeast OPRTase bound to PRPP (PDB entries 2ps1 and 1lh0; González-Segura *et al.*, 2007; Grubmeyer *et al.*, 2012), we can predict where the PRPP phosphates would make contact with the *Pf*OPRTase polypeptide. The 5'-phosphate should make contacts with the pocket formed by <sup>212</sup>TCGTA<sup>216</sup> (Supplementary Fig. S1), while the terminal phosphate of the 1'-pyrophosphate is expected to contact the scaffold formed by <sup>135</sup>YK<sup>136</sup>. In addition to the active-site sulfates, a sulfate close to the N-terminus mediates some of the contacts between  $\alpha$ 1 of one protomer and the loop between  $\alpha$ 4 and  $\beta$ 4 of the other protomer. Three water molecules that coordinate this sulfate ion also help to mediate interactions between the two protomers. Two other sulfates are present at the interface between  $\alpha$ 4 of the two protomers of each dimer. Residues from the two protomers coordinate each of these two sulfates.

A comparison of host–parasite differences in the OPRTase active-site residues within 4 Å of OMP in the mammalian structure (PDB entry 2wns) points to several differences that could contribute to the development of specific antimalarial inhibitors. Fig. 6 shows that Lys136 in the parasite enzyme, which interacts with an active-site sulfate ion that is thought to interact with the 1'-phosphate of PRPP, is a Thr in humans. In addition, Cys213 and Ala216 in the parasite enzyme replace the two Ser side chains in the human enzyme that normally interact with the 5'-phosphate of OMP. Phe97 in the malarial enzyme replaces the Tyr residue of the human enzyme known to form a stacking interaction and a hydrogen bond with orotate. A neighboring Phe98 in *Pf*OPRTase replaces an Ile which forms main-chain interactions with OMP in the human enzyme. Together, these mutations provide hope that it should be possible to find low-molecular-weight inhibitors that bind the parasite enzyme with selectivity.

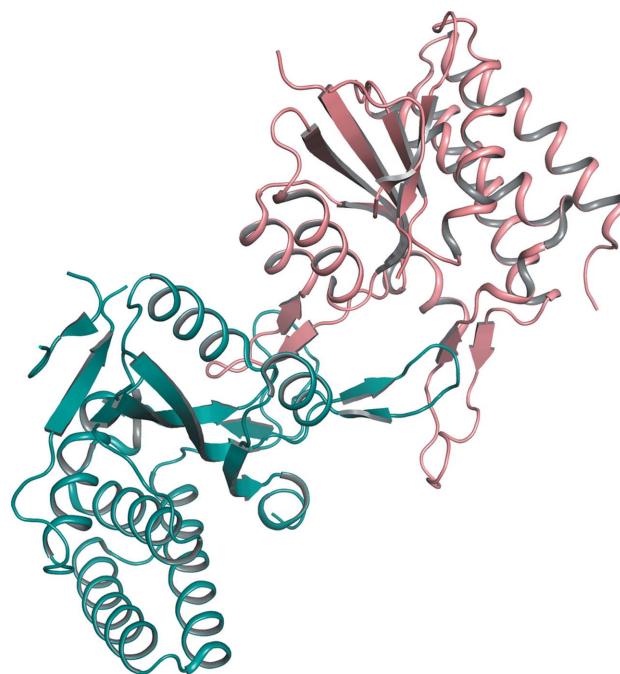
#### 3.4. Parasite-specific inhibitory protein–protein interactions

The individual repeat dimers within the *Pf*OPRTase crystal form unexpected malaria-specific quaternary arrangements that have not been reported for other OPRTases. Among the four catalytic dimers in the asymmetric unit, two are related by an additional noncrystallographic symmetry: a dimer of catalytic dimers form an almost perfect twofold symmetry with  $\kappa = 179.97^\circ$ .

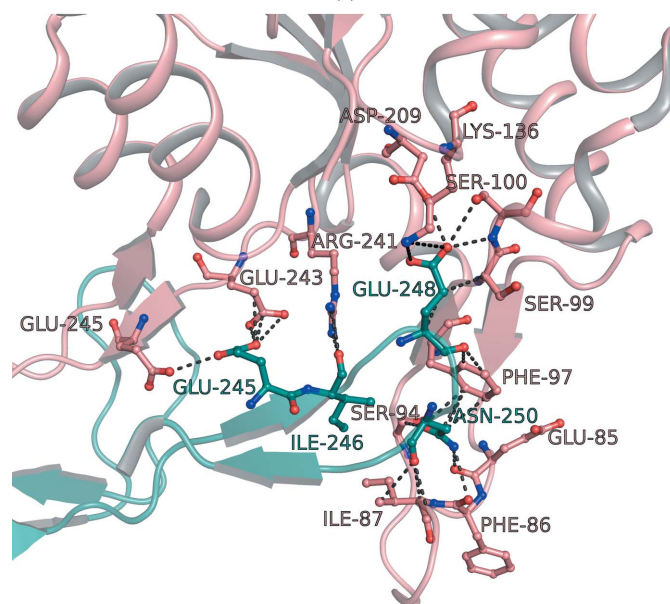
In this larger tetrameric structure, a protomer from one catalytic dimer forms an inhibitory interaction with the active site of another protomer but from a different neighboring catalytic dimer (Fig. 7). A  $\beta$ -hairpin formed by  $\beta 8$  and  $\beta 9$  (<sup>243</sup>EYEINENNQKIY<sup>254</sup>; Supplementary Fig. S2) inserts itself exactly at the substrate-binding pocket of the recipient active site. This completely occludes the closure of the hood domain and would prevent crucial substrate interactions, including  $\pi$ -stacking with Phe97, in the hood domain. The buried surface area at this inhibitory interface is  $\sim 1170 \text{ \AA}^2$ . Twofold symmetry at this inhibitory interface leads to completely isologous interactions. The same residues are involved in the same type of interactions on each of the monomer pairs, not within a dimer but between dimers. The interface hotspot is at Glu248, which makes a salt bridge with Lys136 on the other protomer. Glu248, which is part of the  $\beta$ -hairpin, is buried at the inhibitory interface and forms additional hydrogen bonds to Ser100 and Asp209. Another residue of the  $\beta$ -hairpin, Asn250, also makes crucial interactions (Fig. 7). It acts as a fulcrum around which the hood domain of its dimeric partner wraps itself, making extensive hydrogen bonds *via* the main-chain atoms of Glu85, Phe86 and Leu88 and the side-chain atoms of Ser94. Presumably, these intersubunit interactions blocking the active site of *Pf*OPRTase occurred at high crystallization concentrations, but they hint at protein–protein interactions leading to auto-logous catalytic inhibition of *Pf*OPRTase that may regulate flux through this enzyme. The crystal structure shows that active-site residues in this arrangement cannot interact with the substrate to catalyse the reaction. Inhibition through this

interaction has not been previously observed or studied in OPRTase and therefore these predictions are under active biochemical investigation in our laboratory.

The N-terminus of OPRTase is longer in the *Plasmodium* genus than in other species and this extension is thought to interact with OMP decarboxylase, the next enzyme in the pyrimidine-biosynthesis pathway (Krugkrai *et al.*, 2004). The parasite-specific residues in the longer N-terminus form an



(a)



(b)

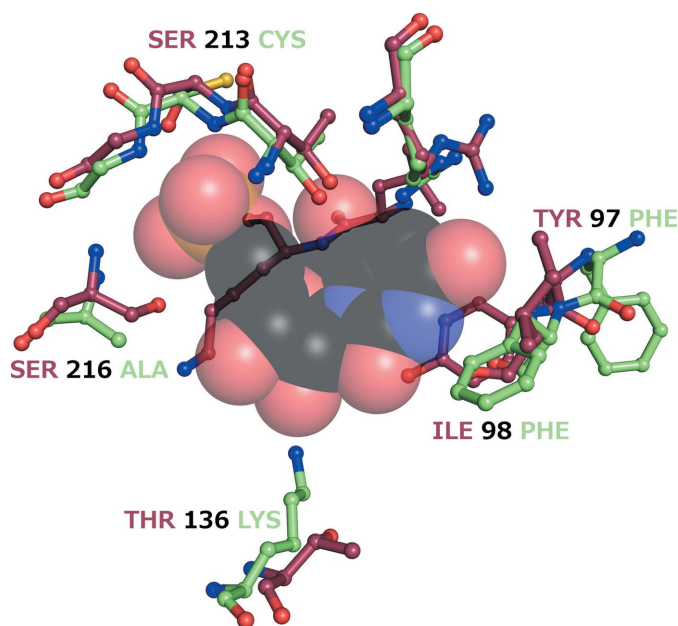


Figure 6

Host–parasite differences in the active site of OPRTase. Differences in key amino acids are color-coded, with the human enzyme structure in green and the malaria parasite structure in red. The space-filling model depicts the position of the product orotidylate (OMP) from the human OPRTase structure. The amino acids in green are those found within 4 Å of OMP (see §3.3).

Figure 7

Intersubunit inhibitory interactions. (a) Unique, parasite-specific, interdimer contacts involving the potentially inhibitory insertion of a loop from one dimer (green) and the active site of an adjacent dimer (pink). (b) Details of the interdimer contacts showing hydrogen-bond interactions that stabilize the incoming loop (green) and that help to form a defined structure on the open ‘hood’ (Glu85, Phe86 and Ile87).



amphipathic helix  $\alpha 1$ , the hydrophobic surface of which is sandwiched between  $\alpha 2$  and  $\alpha 3$  on one side of the *PfOPRTase* core, while the other predominantly polar surface is solvent-exposed. LCR1, a parasite-specific low-complexity sequence that was deleted in the construct that was crystallized, would have juxtaposed exactly between  $\alpha 1$  and  $\alpha 2$ , which now have a short loop separating them. From this position between  $\alpha 1$  and  $\alpha 2$ , LCR1 would not be expected to interact at the active site nor at the dimeric interface. There is another disordered region of 18 amino acids in length between the antiparallel  $\beta 5$  and  $\beta 6$ . This parasite-specific stretch is comprised of an Asp- and Lys-rich low-entropy insert (<sup>182</sup>DDKDILNLKKKTKN-NQDE<sup>199</sup>). While partially conserved in the *Plasmodium* genus, it is absent in other organisms, including humans.

#### 4. Discussion

Nucleotide biosynthesis is an important target in proliferating malarial parasites inside human erythrocytes (Rathod, 2000, 2001). In principle, the complete dependence of *Plasmodium* parasites on pyrimidine biosynthesis, combined with the ability of the host, but not the parasite, to salvage preformed pyrimidines, presents many potential opportunities to attack this important human pathogen with selectivity. In one approach, substrate analogues are selectively converted by the pathogen into potent toxins: 5-fluoroorotate (5FO) kills the parasite with low-nanomolar potency and with more than 1000-fold selectivity (Rathod *et al.*, 1989). A key step in this action is the conversion of 5FO to 5FOMP and ultimately other toxic fluorinated nucleotides by *PfOPRTase* (Rathod & Reyes, 1983; Rathod *et al.*, 1992). Exploiting the structural information on the *PfOPRTase* enzyme and its differences from the human enzyme may help in the design of even more specific and potent substrate analogues directed at malaria parasites. In a very different approach, high-throughput screening against *P. falciparum* dihydroorotate dehydrogenase and lead optimization of an active triazolopyrimidine has led to potent selective inhibitors of malaria parasites that are in human testing (Baldwin *et al.*, 2005; Phillips *et al.*, 2008; Coteron *et al.*, 2011; Deng *et al.*, 2014; Gujjar *et al.*, 2011; <http://www.mmv.org/research-development/rd-portfolio>).

Among all the early enzymes of the pyrimidine-biosynthesis pathway in *Plasmodium*, the structure of *PfOPRTase* is the only one that has been unavailable until now. Our first accomplishment here was to obtain crystallizable *PfOPRTase*. The *Plasmodium* proteome has an abundance of proteins with interspersed low-complexity regions (LCRs), which are often inserted within the coding regions of enzymes (Gardner *et al.*, 2002; Aravind *et al.*, 2003). Some AT-rich micro-satellites in the genome manifest themselves in the proteome in the form of continuous stretches of asparagines. These asparagine-repeat stretches are 37 residues in length on average (Zilversmit *et al.*, 2010). While the evolutionary and functional significance of these LCRs to the parasite is under active debate (Guler *et al.*, 2013; Muralidharan *et al.*, 2012; Aravind *et al.*, 2003), there is agreement that they can hinder over-expression in heterologous systems.

High protein expression was achieved by the removal of a 22-amino-acid asparagine-rich LCR that is close to the N-terminus of *PfOPRTase* without compromising the catalytic activity of the enzyme. This altered construct was able to complement wild-type *OPRTase* in the *E. coli* knockout strain JW3617(DE3) and permit its growth in M9 minimal medium. For crystallography, removing dispensable LCRs appeared to reduce conformational heterogeneity since it was not possible to identify a crystallization condition for wild-type *PfOPRTase*. However, the trimmed *PfOPRTase* 1–218( $\Delta 37$ –58) could be crystallized under a number of conditions. The successful removal of this particular LCR region without affecting the catalytic function of the protein can now be rationalized based on the available new structure, since the deleted segment (37–58) would normally be present between  $\alpha 1$  and  $\alpha 2$ , away from the core of the protein.

In the future, it may be possible to find selective active-site binding of inhibitors to this enzyme because the active site of *PfOPRTase* has a number of amino acids that differ in the parasite enzyme compared with the human enzyme. Remarkably, of the 11 active-site residues at the core of the human enzyme, five are different in the malarial enzyme. Perhaps even more significantly, all five of the unique amino acids are on the back wall of the active site where OMP and PRPP bind, while the fold that wraps around the substrate is well conserved. Additional structures of *PfOPRTase* 1–218( $\Delta 37$ –58)-H6 bound to substrates, products and inhibitors should be helpful in further exploiting this.

Beyond the active sites of individual dimers, there are higher order superstructures of *PfOPRTase* with potential implications for malaria biology and drug design. *PfOPRTase* 1–218( $\Delta 37$ –58)-H6, in addition to forming well known dimers at low concentration, seems to form inhibitory multimers at very high concentrations. Inside the unit cell of our *PfOPRTase* 1–218( $\Delta 37$ –58)-H6 crystal (PDB entry 4fym) there are 32 polypeptides. Equating this with the unit-cell volume, the enzyme concentration in the crystal is about 18 mM, similar to that used to set up the crystals. If these concentrations are approachable in cells, the interactions that are described could be of physiological importance in malaria biology. These protein-level auto-regulatory feedback loops have not previously been recognized. They may work hand in hand with the ability of malaria enzymes to form autologous complexes with nucleic acids (Zhang & Rathod, 2002).

The short peptide structures involved in interdimer interactions in *PfOPRTase* reveal novel protein–protein interactions that could serve as a template for other more rigid inhibitors directed at the *PfOPRTase* active site. The loop peptide entering the *PfOPRTase* active site not only precludes the normal *PfOPRTases* active-site hood from closing, but the hood domain becomes structured when the  $\beta$ -hairpin of the inhibitory dimer inserts itself over the active site. Since  $\beta$ -hairpins form stable secondary structures in isolation (Blanco *et al.*, 1994), their analogues may be used to specifically inhibit the enzyme activity of *PfOPRTase*.



5. Summary

*PfOPRTase*, an important enzyme in *de novo* pyrimidine biosynthesis in malaria parasites, has been overexpressed and crystallized. This is an important advancement in our quest to develop many different and increasingly better inhibitors to block malaria pyrimidine biosynthesis with selectivity. The present structure will help in the development of nucleotide precursors as prodrugs against parasites and also inhibitors that directly block *OPRTase* activity in malaria parasites. As an additional bonus, the new structure hints at a unique mechanism for autoregulation of *PfOPRTase* activity in the parasite based on interdimer protein–protein interactions. The nature of this interdimer structure points to a natural lead peptide inhibitor that could also guide future structure-inspired antimalarial drug-development strategies.

Acknowledgements

This work was supported by US NIH grants (AI099280 and AI093380). The authors thank the staff at GM/CA CAT and NE-CAT at Argonne National Laboratory for their kind assistance with diffraction data collection. We thank Laura Chery for critical reading of the manuscript and we are most grateful to the reviewers and the Editor for important suggestions to further improve the presentation of our data.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.  
 Alfadhli, S. & Rathod, P. K. (2000). *Mol. Biochem. Parasitol.* **110**, 283–291.  
 Aravind, L., Iyer, L. M., Wellems, T. E. & Miller, L. H. (2003). *Cell*, **115**, 771–785.  
 Aurrecochea, C. *et al.* (2009). *Nucleic Acids Res.* **37**, D539–D543.  
 Baldwin, J., Michnoff, C. H., Malmquist, N. A., White, J., Roth, M. G., Rathod, P. K. & Phillips, M. A. (2005). *J. Biol. Chem.* **280**, 21847–21853.  
 Blanco, F. J., Rivas, G. & Serrano, L. (1994). *Nature Struct. Mol. Biol.* **1**, 584–590.  
 Coteron, J. M. *et al.* (2011). *J. Med. Chem.* **54**, 5540–5561.  
 Deng, X., Kokkonda, S., El Mazouni, F., White, J., Burrows, J. N., Kaminsky, W., Charman, S. A., Matthews, D., Rathod, P. K. & Phillips, M. A. (2014). *J. Med. Chem.* **57**, 5381–5394.  
 Dhanawat, M., Das, N., Nagarwal, R. C. & Shrivastava, S. K. (2009). *Mini Rev. Med. Chem.* **9**, 1447–1469.  
 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.  
 Gamo, F. J., Sanz, L. M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J. L., Vanderwall, D. E., Green, D. V., Kumar, V., Hasan, S., Brown, J. R., Peishoff, C. E., Cardon, L. R. & Garcia-Bustos, J. F. (2010). *Nature (London)*, **465**, 305–310.  
 Gardner, M. J. *et al.* (2002). *Nature (London)*, **419**, 498–511.  
 Gómez, Z. M. & Rathod, P. K. (1990). *Antimicrob. Agents Chemother.* **34**, 1371–1375.  
 González-Segura, L., Witte, J. F., McClard, R. W. & Hurley, T. D. (2007). *Biochemistry*, **46**, 14075–14086.

Grubmeyer, C., Hansen, M. R., Fedorov, A. A. & Almo, S. C. (2012). *Biochemistry*, **51**, 4397–4405.  
 Gujjar, R. *et al.* (2011). *J. Med. Chem.* **54**, 3935–3949.  
 Guler, J. L., Freeman, D. L., Ah Yong, V., Patrapuvich, R., White, J., Gujjar, R., Phillips, M. A., DeRisi, J. & Rathod, P. K. (2013). *PLoS Pathog.* **9**, e1003375.  
 Henriksen, A., Aghajari, N., Jensen, K. F. & Gajhede, M. (1996). *Biochemistry*, **35**, 3803–3809.  
 Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.  
 Krungkrai, S. R., Prapunwattana, P., Horii, T. & Krungkrai, J. (2004). *Biochem. Biophys. Res. Commun.* **318**, 1012–1018.  
 Kwiatkowski, D. P. (2005). *Am. J. Hum. Genet.* **77**, 171–192.  
 Liu, C.-P., Xu, R., Gao, Z.-Q., Xu, J.-H., Hou, H.-F., Li, L.-Q., She, Z., Li, L.-F., Su, X.-D., Liu, P. & Dong, Y.-H. (2010). *Acta Cryst.* **F66**, 498–502.  
 Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.  
 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.  
 Mehlin, C. *et al.* (2006). *Mol. Biochem. Parasitol.* **148**, 144–160.  
 Mudeppa, D. G. & Rathod, P. K. (2013). *Eukaryot. Cell*, **12**, 1653–1663.  
 Muralidharan, V., Oksman, A., Pal, P., Lindquist, S. & Goldberg, D. E. (2012). *Nature Commun.* **3**, 1310.  
 Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., Fullman, N., Naghavi, M., Lozano, R. & Lopez, A. D. (2012). *Lancet*, **379**, 413–431.  
 Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.  
 Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.  
 Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.  
 Phillips, M. A., Gujjar, R., Malmquist, N. A., White, J., El Mazouni, F., Baldwin, J. & Rathod, P. K. (2008). *J. Med. Chem.* **51**, 3649–3653.  
 Rathod, P. K. (2000). *Biology of Parasitism*, edited by C. Tschudi & E. Pearce, pp. 73–93. Boston, Dordrecht, London: Kluwer.  
 Rathod, P. K. (2001). *Antimalarial Chemotherapy*, edited by P. Rosenthal, pp. 287–306. Totowa: Humana Press.  
 Rathod, P. K., Khatri, A., Hubbert, T. & Milhous, W. K. (1989). *Antimicrob. Agents Chemother.* **33**, 1090–1094.  
 Rathod, P. K., Leffers, N. P. & Young, R. D. (1992). *Antimicrob. Agents Chemother.* **36**, 704–711.  
 Rathod, P. K. & Reyes, P. (1983). *J. Biol. Chem.* **258**, 2852–2855.  
 Reyes, P., Rathod, P. K., Sanchez, D. J., Mrema, J. E., Rieckmann, K. H. & Heidrich, H. G. (1982). *Mol. Biochem. Parasitol.* **5**, 275–290.  
 Scapin, G., Grubmeyer, C. & Sacchettini, J. C. (1994). *Biochemistry*, **33**, 1287–1294.  
 Schneider, E. L., King, D. S. & Marletta, M. A. (2005). *Biochemistry*, **44**, 987–995.  
 Singh, G. P., Chandra, B. R., Bhattacharya, A., Akhouri, R. R., Singh, S. K. & Sharma, A. (2004). *Mol. Biochem. Parasitol.* **137**, 307–319.  
 Takashima, Y., Mizohata, E., Tokuoka, K., Krungkrai, S. R., Kusakari, Y., Konishi, S., Satoh, A., Matsumura, H., Krungkrai, J., Horii, T. & Inoue, T. (2012). *Acta Cryst.* **F68**, 244–246.  
 Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.  
 Vedadi, M. *et al.* (2007). *Mol. Biochem. Parasitol.* **151**, 100–110.  
 Zhang, K. & Rathod, P. K. (2002). *Science*, **296**, 545–547.  
 Zilversmit, M. M., Volkman, S. K., DePristo, M. A., Wirth, D. F., Awadalla, P. & Hartl, D. L. (2010). *Mol. Biol. Evol.* **27**, 2198–2209.