Crystal structure of *Plasmodium falciparum* proplasmepsin IV: the plasticity of proplasmepsins

Rosario Recacha,^{a*} Kristaps Jaudzems,^a Inara Akopjana,^b Aigars Jirgensons^a and Kaspars Tars^b

^aLatvian Institute of Organic Synthesis, Aizkraukles 21, Riga, LV-1006, Latvia, and ^bBiomedical Research and Study Centre, Ratsupites 1, Riga, LV-1067, Latvia. *Correspondence e-mail: rosario@biomed.lu.lv

Received 6 May 2016

Accepted 18 July 2016

Edited by M. S. Weiss, Helmholtz-Zentrum Berlin für Materialien und Energie, Germany

Keywords: malaria; proplasmepsin IV; *Plasmodium falciparum*; aspartic protease zymogen.

PDB reference: proplasmepsin IV, 5jod

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

Plasmepsin IV from *Plasmodium falciparum* (PM IV) is a promising target for the development of novel antimalarial drugs. Here, the crystal structure of the truncated zymogen of PM IV (pPM IV), consisting of the mature enzyme plus a prosegment of 47 residues, has been determined at 1.5 Å resolution. pPM IV presents the fold previously described for studied proplasmepsins, displaying closer similarities to proplasmepsin IV from *P. vivax* (pPvPM) than to the other two proplasmepsins from *P. falciparum*. The study and comparison of the pPM IV structure with the proplasmepsin structures described previously provide information about the similarities and differences in the inactivation–activation mechanisms among the plasmepsin zymogens.

1. Introduction

The malaria parasite continues to be one of the leading causes of death in many developing countries (World Health Organization, 2013). With the development of resistance against currently available treatments, the discovery of new therapeutics is imperative. There are four species of malaria parasite that infect humans: *Plasmodium falciparum* (Pf), *P. malariae* (Pm), *P. ovale* (Po) and *P. vivax* (Pv). Pf is responsible for the majority of deaths associated with malarial infections, while Pm, Po and Pv are associated with chronic infections, in some cases for an individual's entire life.

During the asexual stage of the parasite life cycle, it invades red blood cells and degrades haemoglobin. This is performed by plasmodial aspartic proteases, plasmepsins, in a specialized compartment called the food vacuole. It has long been known that treating the parasites in culture with aspartic protease inhibitors kills them (Francis *et al.*, 1994). This fact, along with their key function in haemoglobin degradation, makes the plasmepsin enzymes viable drug targets (Coombs *et al.*, 2001). *P. falciparum* contains four plasmepsin enzymes in the food vacuole: PM I, PM II, histoaspartic protease (HAP) and PM IV. *In vitro* studies have shown that PM IV is capable of digesting haemoglobin, but is more active against denatured globin (Banerjee *et al.*, 2002). Another function that has been attributed to PM II and PM IV is the degradation of spectrin and other components of the erythrocyte cytoskeleton (Le Bonniec *et al.*, 1999; Wyatt & Berry, 2002). PM IV is the orthologue of the unique plasmepsin present in the food vacuole of the other three species infecting man (PmPM, PvPM and PoPM; Dame *et al.*, 2003). The sequence similarity among the different species is between 70 and 87% (Bernstein

© 2016 International Union of Crystallography

et al., 2003). Therefore, there is a growing interest in the structural characterization of this class of enzymes to aid structure-based drug design (Jaudzems *et al.*, 2014; Recacha *et al.*, 2015; Rasina *et al.*, 2016).

Plasmeprins, like many other proteases, are synthesized as zymogens, inactive precursors of the enzymes which provide protection from proteolysis, stabilize the inactive form and prevent the entry of the substrate into the active site. *In vivo*, a maturase converts *Pf* proplasmepsins (pPMs) to the active proteinases, but recombinant pPM II and pPM IV can also be activated autocatalytically *in vitro* at acidic pH (Hill *et al.*, 1994).

The Protein Data Bank contains four crystal structures of three proplasmepsins: pPvPM from *P. vivax* (PDB entry 1miq; Bernstein *et al.*, 2003) and pPM II [PDB entries 1pfz (Bernstein *et al.*, 1999) and 5bwy (R. Recacha, I. Akopjana, K. Tars & K. Jaudzems, unpublished work)] and pHAP (PDB entry 3qvc; Bhaumik *et al.*, 2011) from *P. falciparum*. All of these proenzymes are folded into three topologically similar domains (Fig. 1*a*): the N- and C-terminal domains are anchored to the central motif, a six-stranded antiparallel β -sheet.

In this paper, we present the crystal structure of pPM IV and compare it with the structures of the mature enzyme and the proplasmepsins pPvPM, pPM II and pHAP. pPM IV has a higher sequence homology (Fig. 2*a*) to pPvPM (72%) than to pPM II (64%) and to pHAP (59%), showing that the zymogens pPM IV and pPvPM have greater structural similarity. Structural studies of the similarities and differences between the structure of pPM IV and the structures of its mature enzyme and of the other proplasmepsins will provide more

information about the inactivation mechanism by the propeptide and insight for future drug design.

2. Materials and methods

2.1. Expression and purification

The proenzyme of plasmeprsin IV, with a prosegment of 47 residues, was overexpressed using the pET-3a vector (Novagen) in *Escherichia coli* BL21 (DE3) cells. The protein was isolated from inclusion bodies. Refolding and purification were performed as described by Beyer *et al.* (2004).

The purified pPM IV was concentrated to 14 mg ml⁻¹ in 20 mM Tris-HCl pH 8.0 using a 10 kDa cutoff Amicon concentrator. Stocks of 50 μ l protein solution were flash-frozen in liquid nitrogen and stored at -80°C.

2.2. Crystallization and data collection

pPM IV at 14 mg ml⁻¹ in 20 mM Tris-HCl pH 8.0 was crystallized by vapour diffusion using MRC 96-well sitting-drop plates (Molecular Dimensions) by mixing 1 μ l protein solution with 1 μ l reservoir solution consisting of 100 mM sodium citrate pH 4.7, 25% (w/v) PEG 3350, 200 mM ammonium acetate (Table 1). The obtained crystals were cryo-protected in reservoir solution containing 30% (v/v) glycerol and flash-cooled in liquid nitrogen.

Data from pPM IV crystals were collected on beamline I911-3 at the MAX-lab synchrotron, Lund, Sweden. Images were processed by *iMosflm* (Leslie *et al.*, 2002; Powell *et al.*, 2013) and scaled using *SCALA* (Evans, 2006) from the CCP4 suite (Winn *et al.*, 2011).

Figure 1

Structural view of the conversion of pPM IV (*a*) to PM IV (*b*). The regions discussed in the main text are labelled. The prosegment is shown in blue, the N-terminal 13 amino acids of the mature sequence in red, the flap loop in yellow and the catalytic Asp34-Asp214 dyad as sticks.

Figure 2

Comparison of proplasmepsins: pPM IV, pPvPM, pPM II and pHAP. (a) Structure-based sequence alignment (performed with *ESPrPr*; Robert & Gouet, 2014) of the four proplasmepsins. The prosegment is underscored with a blue bar, the N-terminal fragment (residues 1–13) with a red bar and the flap loop with a yellow bar. (b) Superposition of the zymogen structures of PM IV (green), PvPM (orange), PM II (magenta) and HAP (lilac) in two orientations related by a rotation of 90°. The regions discussed in the text are shown in darker colours and labelled, and the catalytic Asp34/His34–Asp214 is shown as sticks. The Tyr122p–Asp4 loop and the flap loop are highlighted with a black circle and a black rectangle, respectively. (c, d) Close-up views showing the conformational differences among proplasmepsins in the loop comprised of residues Tyr122p–Asp4 and the flap loop, respectively. Salt bridges and hydrogen bonds are shown as dashed lines and the side chains of the residues involved in the interactions are shown as sticks. (c) The main chain of residues Ser1–Asp4 (pPM IV) and Asn3–Asp4 (pPvPM) are shown as sticks. (d) The angle between the flap loop of pPM IV and of pPM II is shown at the bottom.

Table 1
Crystallization.

Method	Vapour diffusion
Plate type	MRC 96-well sitting-drop plate
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	14
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0
Composition of reservoir solution	100 mM sodium citrate pH 4.7, 25% (w/v) PEG 3350, 200 mM ammonium acetate
Volume and ratio of drop	2 µl, 1:1
Volume of reservoir (µl)	100

2.3. Structure solution and refinement

The orientation and position of pPM IV were solved using the molecular-replacement program *Phaser* (McCoy *et al.*, 2007) in the *CCP4* suite (Winn *et al.*, 2011), using pPM II (PDB entry 1pfz; Bernstein *et al.*, 2003), which presents 67% sequence homology, as a model (Fig. 2a). Rigid-body refinement followed by Cartesian simulated annealing was initially applied to the structure using *PHENIX* (Echols *et al.*, 2014), which resulted in a high-quality electron-density map for the manual rebuilding of the pPM IV structure. The model was improved by iterative cycles of manual rebuilding with *Coot* (Emsley & Cowtan, 2004) in combination with σ_A -weighted $2F_o - F_c$ and $F_o - F_c$ maps and automated refinement with *PHENIX*, applying a hybrid TLS-isotropic ADP model.

Statistics of data collection, final refinement and validation by *MolProbity* (Chen *et al.*, 2010) are summarized in Tables 2 and 3. The superposition of models and the r.m.s. deviations were calculated using *LSQKAB* in the *CCP4* suite (Winn *et al.*, 2011). Buried surfaces and residues at the intermolecular contacts in the crystals were identified with the *PISA* server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html; Krissinel & Henrick, 2007), and the *CONTACT* program in the *CCP4* suite (Winn *et al.*, 2011) was used to study the interactions. The figures were generated with *UCSF Chimera* (Pettersen *et al.*, 2004) and *CCP4mg* (McNicholas *et al.*, 2011).

The coordinates and structure factors for pPM IV have been deposited in the Protein Data Bank as entry 5jod.

3. Results and discussion

The fragment of the PM IV zymogen that was crystallized consisted of the C-terminal prosegment truncated to 47 residues, denominated 78p–124p, followed by the sequence corresponding to the mature enzyme. This truncated proplasmepsin IV is inactive at neutral pH, and undergoes autoactivation at pH 4.4 when incubated at 37°C (unpublished data; Fig. 1b; PM IV; PDB entry 1ls5; Asojo *et al.*, 2003).

3.1. Overall structure and comparison to other proplasmepsins

pPM IV crystallized in space group *P2*₁ with two molecules in the crystallographic asymmetric unit. The monomers superpose very well for the trans-motif and C-terminal domain, with the major differences present in the N-terminal domain, with monomer *B* in a more closed conformation (data

Table 2
Data collection and processing.

Values in parentheses are for the outer resolution shell.

Diffraction source	I911-3, MAX-lab
Wavelength (Å)	0.97
Temperature (K)	100
Detector	MAR CCD
Crystal-to-detector distance (mm)	153.9
Rotation range per image (°)	0.20
Total rotation range (°)	130
Space group	<i>P2</i> ₁
Unit-cell parameters (Å, °)	<i>a</i> = 82.91, <i>b</i> = 62.78, <i>c</i> = 94.03, α = 90, β = 113.50, γ = 90
Mosaicity (°)	0.56
Resolution (Å)	48.41–1.53 (1.55–1.53)
<i>R</i> _{meas} (%)	7.9 (7.9)
<i>R</i> _{p.i.m.} (%)	4.5 (5.0)
Total No. of reflections	342658 (13264)
No. of unique reflections	129149 (6237)
Completeness (%)	96.3 (95.1)
$\langle I/\sigma(I) \rangle$	7.5 (1.2)
Multiplicity	2.7 (2.1)
Wilson <i>B</i> factor (Å ²)	18.2
<i>CC</i> _{1/2}	0.995 (0.330)

Table 3
Structure refinement.

Values in parentheses are for the outer resolution shell.

Resolution (Å)	36.68–1.53
Completeness (%)	96.3 (95.9)
No. of reflections, working set	122621 (5965)
No. of reflections, test set	6883 [5.33%]
<i>R</i> _{work} / <i>R</i> _{free}	0.179/0.209
No. of atoms	
Protein	5947
Water	689
Ligand (glycerol)	54
Average <i>B</i> factors (Å ²)	
Overall	30.3
Protein	28.8
Ligand	62.8
Water	40.5
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.070
Ramachandran plot	
Most favoured (%)	98.0
Outliers (%)	0.0
<i>MolProbity</i> overall score	1.36
<i>MolProbity</i> clashcore	3.19

not shown). The r.m.s. deviation between monomers *A* and *B* is around 1.9 Å for the Met15–Val140 C^α atoms. Although the electron density for certain portions of the prosegment of both monomers (Ser115p–Lys120p; Supplementary Fig. S1), and Glu7–Asp10 of the mature fragment of monomer *B*, was poor, it was still possible to model the structure of all of these residues. The Glu7–Asp10 residues correspond to a short 3₁₀-helix (Fig. 1a) and have been described previously in other proplasmepsin structures (Bernstein *et al.*, 1999, 2003; Bhaumik *et al.*, 2011). This short helix is found on the surface of the C-terminal domain (Fig. 1a) and is part of the fragment 1–13 that undergoes conformational changes upon the activation of the proenzyme from a random coil and 3₁₀-helical conformation to a β -strand in concert with its relocation from

the surface of the C-domain into the central β -sheet at the back of the active-site cleft (Fig. 1*b*).

The overall fold of the truncated pPM IV is quite similar to the previously reported zymogens (Fig. 2*b*) pPM II (Bernstein *et al.*, 1999) and pHAP (Bhaumik *et al.*, 2011) from *P. falciparum* and pPvPM from *P. vivax* (Bernstein *et al.*, 2003). Fig. 2(*a*) shows the sequence homology of the proplasmepsins with known structures, and we can see that the major sequence differences among zymogens are localized in the prosegment region, although all of them conserve the secondary-structure motif of a β -strand for the fragment Thr80p–Asp87p, which is followed by two α -helices: α 1, Pro90p–Asn100p, and α 2, Leu102p–Lys114p (Fig. 2*b*). The α 1 helices of pPM II and pHAP superpose very well with each other, but not with the helices from pPM IV or pPvPM, which are shorter. The second α -helix of pPM IV superposes better with α 2 from pPM II than with those from the other zymogens. It is evident that the prosegment structure of pPvPM, including both α -helices and the connecting loop, is rather different from the prosegments of the other zymogens. The loop connecting the two α -helices is very flexible, and is displaced by about 7 Å in pPvPM with respect to the other proplasmepsins, affecting the conformation of the loop Tyr273–Cys285, which assumes an in-conformation in pPMs but is in an out-conformation in pPvPM (Fig. 2*b*). Although the C-terminal domain of the mature enzyme presents high sequence homology, the residues of the loop Tyr273–Cys285 are not so well conserved (Fig. 2*a*).

The pro-mature junction, Gly124p–Ser1, is located in a tight loop delimited by Tyr122p and Asp4, the ‘Tyr–Asp’ loop (Fig. 2*c*). These loops superpose well in the structures of pPM IV and pPM II, but they are displaced by 2.4 Å in relation to that of pPvPM, and in the pHAP structure this loop is disordered (Fig. 2*b*). The fold of the loop is stabilized by hydrogen-bond interactions between the carboxylate of Asp4 and the phenolic hydroxyl of Tyr122p and, in the structure of pPvPM, also with the hydroxyl group of Ser1. Asp4 is also involved in a hydrogen-bond interaction with the main-chain amide group of Phe241 and a salt bridge to the amino group of Lys238 (Fig. 2*c*).

3.2. Flap-loop conformation

For the study and comparison of the flap-loop conformation we used the unpublished structure of pPM II with PDB entry 5bwy, instead of that described in the literature (PDB entry 1pfz; Bernstein *et al.*, 1999), as the flap-loop residues of the 5bwy structure have been modelled and some important interactions of these residues could be compared with the other pPMs.

The structures of pPM IV and pPvPM present a more closed conformation of the flap loop than the zymogens of PM II and HAP (Fig. 2*b*). On superimposition of the structures, the flap loop of pPM IV and the flap loop of pPM II form an angle of approximately 25° (Figs. 2*b* and 2*d*). The different conformations of the flap loops translate into dissimilarities in their interactions (Fig. 2*d*), as well as differences in the conforma-

tion of nearby areas (Fig. 2*b*). The loop Trp128–Asp137 borders the flap on the prime (*Sn'*) side of the substrate-binding site. Although all its residues are conserved among the proplasmepsins studied (Fig. 2*a*), it acquires a different conformation in each of the proplasmepsin structures (Fig. 2*b*). In pPvPM and pHAP it folds towards the cleft, pointing Leu131 into the S2' pocket, while in pPM IV and pPM II this loop points away from the cleft, leaving a deeper pocket (Fig. 2*b*).

The segment from Ala108 to Pro113, which contacts the flap on the nonprime (*Sn*) side, and the following α -helix (residues Ile114–Glu119) has a different sequence in pPM IV compared with all other zymogens. The α -helices of zymogens pPM IV and pPvPM superpose well (Fig. 2*b*), but they do not superpose either with the helix of pPM II or with that of pHAP.

In the structures of pPM IV and pPvPM that show the closed conformation of the flap loop, Tyr77 interacts with Trp41 (Fig. 2*d*). However, in pPMII and pHAP the open conformation of the flap loop accommodates another rotamer of Trp41 that is not possible in the closed-flap conformation owing to clashes with Ile75 (Fig. 2*d*). In the open conformation Trp41 establishes a hydrogen-bond interaction with Ser37, which interacts with the catalytic Asp34 in the proenzymes of PM IV and PvPM. In some structures of the mature PM II both conformations of Trp41 are present (Prade *et al.*, 2005; Rasina *et al.*, 2016). The presence of this pocket in pPM II and pHAP (Fig. 2*d*) makes the binding of an inhibitor to them entropically more favourable than in the mature enzymes (Carcache *et al.*, 2002). This cannot be considered for either pPM IV or pPvPM as they lack this hydrophobic pocket.

3.3. pPM IV prosegment interactions and plasmepsin activation

Proplasmepsins present high sequence and structural homology, with the major differences being localized in the prosegment (Fig. 2*a*). Despite this, the prosegment is involved in numerous well conserved hydrogen bonds and hydrophobic interactions (Fig. 3). The side chains of Lys78p and His79p form salt-bridge interactions with the carboxylic group of Asp150 and their main-chain amide groups form hydrogen bonds to the carboxylic group of Glu174, as has been reported for the other proplasmepsins (Bernstein *et al.*, 1999; Bhaumik *et al.*, 2011). These two residues are followed by the first strand of the β -sheet (Thr80p–Asp87p), which is replaced by Asp4–Val11 of the mature enzyme upon activation (Fig. 1) and forms similar interactions (Fig. 3*a*). These residues are involved in the formation of an antiparallel β -sheet with a string of hydrogen bonds to Ile170–His164 of the mature enzyme, which are highly conserved in the enzymes studied, with the exception of Val165, which differs in each of the enzymes (Fig. 2*a*). The hydroxyl group of Thr80p, which is not conserved in the other proplasmepsins (Fig. 2*a*), interacts with the carboxylic group of Asp95. As described for the other proplasmepsins (Bernstein *et al.*, 1999, 2003; Bhaumik *et al.*, 2011), the hydroxyl group of Thr81p forms a hydrogen bond to the hydroxyl group of Thr169, and the side chains of Ile82p,

Phe84p and Ile86p fit into hydrophobic pockets lined by residues of the mature portion of pPM IV (Fig. 3a). Lys85p and Asp87p are engaged in four antiparallel β -sheet hydrogen bonds to Leu14 and Phe16 as part of an N-domain β -sheet. Lys85p and Asp87p interact with residues of the other monomer of the crystallographic asymmetric unit. The amino group of Lys85p forms salt-bridge interactions with the carboxylate group of Asp279, and the carboxylate group of Asp87p, the last residue of the β -strand, forms a salt bridge with the amine group of Lys101. These two interactions have not been described in the other structures.

Pro88p–Val14 wrap around the C-terminal domain (Fig. 1a). An equivalent feature has been observed in the

other proplasmepsins (Bernstein *et al.*, 1999, 2003; Bhaumik *et al.*, 2011) and has been called the ‘harness’. This ‘harness’ plays a crucial role in keeping the zymogen inactive at neutral pH, as it establishes hydrogen-bond interactions which help to separate the two domains, the N-domain and C-domain, of the proplasmepsins (Fig. 1). In the case of pPM IV these interactions are disrupted at a pH of around 4.0–4.4 (Li *et al.*, 2004) and the enzyme is activated (Fig. 1). The self-activation of pPvPM and pPM II has also been described, but at a pH that is a little higher (Bernstein *et al.*, 2003; Hill *et al.*, 1994). The lower pH causes protonation of the aspartate residue of the Tyr122p–Asp4 loop (Fig. 2d), resulting in disruption of the hydrogen bonds made by its side chain and subsequent

Figure 3
pPM IV prosegment interactions. Interactions made by the β -strand and α 1 and α 2 of the prosegment (C atoms in grey) with the mature portion (C atoms in green) are shown, with salt bridges and hydrogen bonds shown as dashed lines. (a) β -Strand (80p–87p). (b) α 1 (Pro90p–Asn100p). (c) α 2 (Leu102p–Lys114p). (d) Interactions between the α 2 of two symmetry-related molecules. The loops Lys238–Tyr245 and Pro276–Ala283 are shown in green.

opening of the loop (Bernstein *et al.*, 1999). Other hydrogen-bonding interactions between the prosegment and the C-terminal domain of the mature polypeptide chain are also disrupted at low pH (Bernstein *et al.*, 1999, 2003; Bhaumik *et al.*, 2011). Compared with the zymogen, in the mature enzyme reorientation of the N-domain occurs in a more closed conformation (Fig. 1).

The $\alpha 1$ helix (Pro89p–Leu98p) does not present high sequence homology among the four proplasmepsins, with pPM IV showing the greatest difference (Fig. 2*a*). Only Asp91p is present in all of them and forms hydrogen-bond interactions with the imidazole N atom of His164, which is conserved in pPvPM and pPM II but not in pHAP (Asn164). The hydrophobic residues of $\alpha 1$ (Val93p, Leu94p, Val97p and Leu98p) are involved in interactions with residues from the C-terminal domain of the mature PM IV polypeptide and loops (Fig. 3*b*). The hydroxyl group of Ser96p interacts with the side-chain amide of Asn100p. These two residues are only present in pPM IV. The rest of the $\alpha 1$ residues are exposed towards the solvent.

In all structures, the $\alpha 2$ helix is positioned in a hydrophobic groove formed by two loops (Lys238–Tyr245 and Pro276–Ala283), maintaining a hydrogen bond to the mature portion from the hydroxyl group of Tyr105p (conserved in all of the proplasmepsins; Fig. 2*a*) to the main-chain O atom of Pro240 (Fig. 3*c*). The first loop, residues Lys238–Tyr245, presents high sequence homology among the different proplasmepsins (Fig. 2*a*). However, in pHAP the substitution of Pro243 by a serine makes its loop structure very different from the other three structures (Fig. 2*b*). Additionally, Leu244 in pPM IV and pHAP is mutated to a phenylalanine in pPM II and pPvPM. Major differences are present within the loop Lys238–Tyr245 (Fig. 3*c*). In addition to showing greater sequence diversity (Fig. 2*a*), this loop has an in-conformation in pPM II, pPM IV and pHAP, while in pPvPM it adopts an out-conformation (Fig. 2*b*). pPM IV presents the largest sequence differences in the prosegment $\alpha 2$, which are translated to the interactions, so that pPM IV does not present some of the polar interactions that have been observed in pPM II (Bernstein & James, 1999).

Besides the interactions described above, α -helix 2 of the prosegment is also involved in intermolecular interactions with $\alpha 2$ from another crystallographic symmetry-related molecule (Fig. 3*d*). The buried surface area between monomers of pPM IV is 4210 Å². This dimer is stable in solution, as observed by size-exclusion chromatography during purification (data not shown). The same dimerization has been described for the other plasmepsin zymogens (Bernstein & James, 1999; Bernstein *et al.*, 2003; Bhaumik *et al.*, 2011) and they present similar mostly hydrophobic interactions (Fig. 3*d*).

4. Summary

In summary, the crystal structure of the PM IV zymogen revealed a conformation of the prosegment that is similar to those found in its counterparts pPvPM, pPMII and pHAP, suggesting a common mechanism of inactivation/activation. Some differences include the interactions of the nonconserved

residues of the prosegments and mature enzyme, as well as the different conformations adapted by the proplasmepsin structure parts. Similarly to pPvPM, pPM IV shows a closed conformation of the flap loop and does not have the cavity in the substrate-binding site that is present in the other two proplasmepsins, pPM II and pHAP. This cavity, which has been described as a target for the design of inhibitors against pPM II and pHAP, is created by the open conformation of the flap loop featuring a rotamer of Trp41 which cannot be accommodated in the pPM IV and pPvPM structures owing to clashes with Ile75.

Acknowledgements

This study was supported by the InnovaBalt project (EU-Grant-316149) and the FP7 program Biostruct-X proposal 7869. We thank the staff of the MAX-lab synchrotron for their support during data collection.

References

- Asojo, O. A., Gulnik, S. V., Afonina, E., Yu, B., Ellman, J. A., Haque, T. S. & Silva, A. M. (2003). *J. Mol. Biol.* **327**, 173–181.
- Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M. & Goldberg, D. E. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 990–995.
- Bernstein, N. K., Cherney, M. M., Loetscher, H., Ridley, R. G. & James, M. N. G. (1999). *Nature Struct. Mol. Biol.* **6**, 32–37.
- Bernstein, N. K., Cherney, M. M., Yowell, C. A., Dame, J. B. & James, M. N. G. (2003). *J. Mol. Biol.* **329**, 505–524.
- Bernstein, N. K. & James, M. N. G. (1999). *Curr. Opin. Struct. Biol.* **9**, 684–689.
- Beyer, B. B., Goldfarb, N. E. & Dunn, B. M. (2004). *Curr. Protoc. Protein Sci.*, Unit 21.14. doi:10.1002/0471140864.ps2114s32.
- Bhaumik, P., Xiao, H., Hidaka, K., Gustchina, A., Kiso, Y., Yada, R. Y. & Wlodawer, A. (2011). *Biochemistry*, **50**, 8862–8879.
- Carcache, D. A., Hortner, S. R., Bertogg, A., Binkert, C., Bur, D., Marki, H. P., Dorn, A. & Diederich, F. (2002). *Chembiochem*, **3**, 1137–1141.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Coombs, G. H., Goldberg, D. E., Klemba, M., Berry, C., Kay, J. & Mottram, J. C. (2001). *Trends Parasitol.* **17**, 532–537.
- Dame, J. B., Yowell, C. A., Omara-Opyene, L., Carlton, J. M., Cooper, R. A. & Li, T. (2003). *Mol. Biochem. Parasitol.* **130**, 1–12.
- Echols, N., Moriarty, N. W., Klei, H. E., Afonine, P. V., Bunkóczi, G., Headd, J. J., McCoy, A. J., Oeffner, R. D., Read, R. J., Terwilliger, T. C. & Adams, P. D. (2014). *Acta Cryst. D* **70**, 144–154.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst. D* **60**, 2126–2132.
- Evans, P. (2006). *Acta Cryst. D* **62**, 72–82.
- Francis, S. E., Gluzman, I. Y., Oksman, A., Knickerbocker, A., Mueller, R., Bryant, M. L., Sherman, D. R., Russell, D. G. & Goldberg, D. E. (1994). *EMBO J.* **13**, 306–317.
- Hill, J., Tyas, L., Phylip, L. H., Kay, J., Dunn, B. M. & Berry, C. (1994). *FEBS Lett.* **352**, 155–158.
- Jaudzems, K., Tars, K., Maurops, G., Ivdra, N., Otikovs, M., Leitans, J., Kanepe-Lapsa, I., Domraceva, I., Mutule, I., Trapencieris, P., Blackman, M. J. & Jirgensons, A. (2014). *ACS Med. Chem. Lett.* **5**, 373–377.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Le Bonniec, S., Deregnacourt, C., Redeker, V., Banerjee, R., Grellier, P., Goldberg, D. E. & Schrével, J. (1999). *J. Biol. Chem.* **274**, 14218–14223.
- Leslie, A. G. W., Powell, H. R., Winter, G., Svensson, O., Spruce, D., McSweeney, S., Love, D., Kinder, S., Duke, E. & Nave, C. (2002). *Acta Cryst. D* **58**, 1924–1928.

- Li, T., Yowell, C. A., Beyer, B. B., Hung, S.-H., Westling, J., Lam, M. T., Dunn, B. M. & Dame, J. B. (2004). *Mol. Biochem. Parasitol.* **135**, 101–109.
- 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.
- McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst. D* **67**, 386–394.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst. D* **69**, 1195–1203.
- Prade, L., Jones, A. F., Boss, C., Richard-Bildstein, S., Meyer, S., Binkert, C. & Bur, D. (2005). *J. Biol. Chem.* **280**, 23837–23843.
- Rasina, D., Otikovs, M., Leitans, J., Recacha, R., Borysov, O. V., Kanepe-Lapsa, I., Domraceva, I., Pantelejevs, T., Tars, K., Blackman, M. J., Jaudzems, K. & Jirgensons, A. (2016). *J. Med. Chem.* **59**, 374–387.
- Recacha, R., Leitans, J., Akopjana, I., Aprupe, L., Trapencieris, P., Jaudzems, K., Jirgensons, A. & Tars, K. (2015). *Acta Cryst. F* **71**, 1531–1539.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- World Health Organization (2013). *World Malaria Report 2013*. Geneva: World Health Organization. http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/.
- Wyatt, D. M. & Berry, C. (2002). *FEBS Lett.* **513**, 159–162.