



X-ray Crystal Structure and Specificity of the *Plasmodium falciparum* Malaria Aminopeptidase PfM18AAP

Komagal Kannan Sivaraman^{1†}, Christine A. Oellig^{1†},
Kitmun Huynh¹, Sarah C. Atkinson^{2,3}, Marcin Poreba⁴,
Matthew A. Perugini^{2,3}, Katharine R. Trenholme^{5,6},
Donald L. Gardiner⁵, Guy Salvesen⁷, Marcin Drag^{4,7},
John P. Dalton^{8,9}, James C. Whisstock^{1,10} and Sheena McGowan^{1*}

¹Department of Biochemistry and Molecular Biology, Monash University, Building 77, Wellington Road, Clayton Campus, Melbourne, VIC 3800, Australia

²Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, 30 Flemington Road, Parkville, VIC 3010, Australia

³Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne 3086, Australia

⁴Division of Bioorganic Chemistry, Faculty of Chemistry, Wrocław University of Technology, Wrocław, Poland

⁵Malaria Biology Laboratory, The Queensland Institute of Medical Research, 300 Herston Road, Herston, Brisbane, Queensland 4006, Australia

⁶School of Medicine, University of Queensland, St Lucia, QLD 4072, Australia

⁷Program in Apoptosis and Cell Death Research, Sanford Burnham Medical Research Institute, La Jolla, CA 92037, USA

⁸Institute for the Biotechnology of Infectious Diseases (IBID), University of Technology Sydney, Level 6, Building 4, Corner of Thomas and Harris Street, Ultimo, Sydney, NSW 2007, Australia

⁹Institute of Parasitology, McGill University, 21111 Lakeshore Road, Sainte-Anne De Bellevue, QC H9X 3V9, Canada

¹⁰ARC Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton Campus, Melbourne, VIC 3800, Australia

Received 12 April 2012;

received in revised form

2 June 2012;

accepted 5 June 2012

Available online

16 June 2012

Edited by R. Huber

Keywords:

aminopeptidase;

malaria;

protease;

structural biology;

drug design

The malarial aminopeptidases have emerged as promising new drug targets for the development of novel antimalarial drugs. The M18AAP of *Plasmodium falciparum* malaria is a metallo-aminopeptidase that we show demonstrates a highly restricted specificity for peptides with an N-terminal Glu or Asp residue. Thus, the enzyme may function alongside other aminopeptidases in effecting the complete degradation or turnover of proteins, such as host hemoglobin, which provides a free amino acid pool for the growing parasite. Inhibition of PfM18AAP's function using antisense RNA is detrimental to the intra-erythrocytic malaria parasite and, hence, it has been proposed as a potential novel drug target. We report the X-ray crystal structure of the PfM18AAP aminopeptidase and reveal its complex dodecameric assembly arranged via dimer and trimer units that interact to form a large tetrahedron shape that completely encloses the 12 active sites within a central cavity. The four entry points to the catalytic lumen are each guarded by 12 large flexible loops that could control substrate entry into the

*Corresponding author. E-mail address: Sheena.McGowan@monash.edu.

† K.K.S. and C.A.O. are joint first authors.

Abbreviations used: Pf, *Plasmodium falciparum*; Hb, hemoglobin; DAP, diaminopropionic acid; pNA, *p*-nitroaniline; ACC, 7-amino-4-carbamoylmethylcoumarin; DNPEP, human aspartyl aminopeptidase; NCS, noncrystallographic symmetry.

catalytic sites. PfM18AAP thus resembles a proteasomal-like machine with multiple active sites able to degrade peptide substrates that enter the central lumen. The *Plasmodium* enzyme shows significant structural differences around the active site when compared to recently determined structures of its mammalian and human homologs, which provides a platform from which a rational approach to inhibitor design of new malaria-specific drugs can begin.

© 2012 Elsevier Ltd. Open access under [CC BY-NC-ND license](#).

Introduction

Malaria remains a global health issue and accounts for as many as 1 million deaths each year.¹ Significant progress has been made with respect to the development of a malaria vaccine, although the limited efficacy of these and the challenges of widespread vaccination programs dictate that control of malaria will continue to depend on chemotherapeutic intervention.^{2,3} The prevention and treatment of malaria caused by infection with the parasite *Plasmodium falciparum* (*Pf*) are becoming increasingly complicated due to the spread of drug-resistant parasites.^{4,5} The development of next-generation antimalarial agents is a matter of urgency and depends on identifying drugs with new modes and/or targets of action.^{1,6}

The *Pf* parasite has a complex life cycle that requires both human and mosquito hosts.⁷ The asexual intra-erythrocytic stage of *Pf* infection is responsible for all of the clinical symptoms of malaria. Intra-erythrocytic malaria parasites avidly consume the erythrocyte hemoglobin (Hb), exploiting it as a source of amino acids for incorporation into its own proteins. An acidic organelle inside the parasite, the digestive vacuole, is the primary site of Hb proteolysis where the action of a number of endopeptidases, a dipeptidase, and aminopeptidases reduce the molecule to peptides and dipeptides, and possibly some free amino acids.⁷⁻⁹ However, it is believed that the majority of the Hb peptides are exported from the vacuole into the cytoplasm of the cell where several *Plasmodium* metallo-aminopeptidases act in concert to convert the Hb peptides into individual amino acids.

Nine metallo-aminopeptidases have been identified in the *Pf* genome[‡] (see Ref. 10). Four of these are methionine aminopeptidases that presumably have a housekeeping role and may function in the removal of the initiator methionine from newly synthesized polypeptides. The other five enzymes comprise a prolyl iminopeptidase (or post-prolyl aminopeptidase), a prolyl aminopeptidase, a leucine aminopeptidase (*PfA-M17*), an alanine amino-

peptidase (*PfA-M1*), and the *PfM18AAP* aminopeptidase. Given the restricted specificities of each of these enzymes for particular N-terminal amino acids, it is thought that they act in concert to facilitate Hb catabolism. Consistent with this idea, both *PfA-M1* and *PfA-M17* have been shown to be essential for the survival of the parasite.¹¹ The *PfA-M1* enzyme is indeed involved in Hb degradation, while the *PfA-M17* may have additional function since specific inhibition of its activity in *Pf* parasite cultures caused growth retardation early in the intra-erythrocytic life cycle before Hb digestion is initiated.¹² We have previously determined the X-ray crystal structures of both *PfA-M1* and *PfA-M17*.^{13,14}

PfM18AAP is an ~65-kDa aminopeptidase of the M18 family of metalloproteases.¹⁵ The *PfM18AAP* gene is present in single copy on the *Pf* genome and the protease is expressed in the parasite cytoplasm and also exported to the space between the parasite and host cell, known as the parasitophorous vacuole.¹⁶ Gene disruption/truncation experiments of *PfM18AAP* indicated that the enzyme was dispensable for blood-stage replication but had an associated fitness cost to the parasite.¹⁷ Anti-sense RNA knockdown of *PfM18AAP*, however, resulted in parasite growth abnormalities, vascularization, and cellular damage.¹⁶ The cellular localization and role in growth and development thus indicate that the enzyme may have additional functions in addition to Hb digestion. Lauterbach and Coetzer showed that *PfM18AAP* binds to erythrocyte spectrin and other membrane proteins and suggested that the enzyme could play a role in the release of the parasite from the host erythrocyte.¹⁸ Accordingly, from a perspective of new antimalarial drug development, *PfM18AAP*, along with *PfA-M1* and *PfA-M17*, remains a protein of high interest.¹⁹

The M18 aminopeptidases are zinc metalloenzymes that exist as high-molecular-weight complexes; previous gel-filtration analyses of recombinant and native *PfM18AAP* suggested that the enzyme assembly was octameric. *PfM18AAP* exhibits optimal activity at a neutral pH, and studies using chelation agents have shown that altering the type of metal ion in the active site can increase catalytic efficiency but does not affect binding of the substrate.¹⁶ Here, we present the X-ray crystal structure of *PfM18AAP* and

‡ <http://plasmodb.org/plasmo/>

show that it is a canonical member of the human aspartyl aminopeptidase (DNPEP)/M18 aminopeptidase family. Like the bovine and human M18AAP,²⁰ PfM18AAP forms a complex homododecameric tetrahedron quarternary structure but exhibits subtle differences in structure that could be exploited in specific inhibitor design.

Results

Purification of active recombinant PfM18AAP from *Escherichia coli* requires chaperonin

As reported by Teuscher *et al.*, recombinant PfM18AAP had been successfully expressed and purified from insect cells using a synthetic construct.¹⁶ To improve the yield of protein for crystallography studies, we constructed a new DNA construct for expression in *E. coli*. Recombinant PfM18AAP was produced as an N-terminal hexahistidine fusion protein and purified using sequential Ni-NTA affinity and size-exclusion chromatography. Gel-filtration chromatograms consisted of a single peak that was clearly in the void volume. This initial material had no enzymatic activity. SDS-PAGE analysis of the gel-filtration peak showed two protein bands at ~65 kDa. N-terminal sequencing of these two protein species identified that the dominant species was chaperonin GroEL (RefSeq ZP_06938783.1). To remove the chaperone protein from our purified PfM18AAP, we added a dissociation step to our purification protocol. Post-dissociation, the repurified protein was enzymatically active against the simple chromogenic substrate L-glutamic acid γ -(4-nitroanilide) [L-Glu-*p*-nitroaniline (pNA)] and was positively identified as PfM18AAP by N-terminal sequencing.

PfM18AAP forms a dodecamer in solution

Gel filtration of the purified PfM18AAP alone showed that the protein was still retained in the void volume of an S200 10/300 column, indicating that the soluble protein was forming a high-molecular-mass complex. Previous studies had employed analytical gel filtration to estimate the molecular mass of the PfM18AAP and showed that it behaved as an octamer in solution (~560 kDa).¹⁶ To definitively identify the size of the PfM18AAP oligomer in aqueous solution, sedimentation velocity studies were employed in the analytical ultracentrifuge at initial protein concentrations ranging from 0.1 mg/ml to 1.0 mg/ml. Similar data were obtained at all concentrations studied. As an example, the absorbance *versus* radial position profiles at different time points are shown in Fig. 1a (open symbols) for PfM18AAP at 0.5 mg/ml. These data were subse-

quently fitted to a continuous sedimentation coefficient $[c(s)]$ distribution model (Fig. 1b) and a continuous mass $[c(M)]$ distribution model (Fig. 1c). Both models yielded excellent fits, as demonstrated by the random distribution of residuals (Fig. 1b and c, top panels) and low (<0.01) rmsd values (Table 1). The data show that the protein predominantly exists as a 20.4 ± 1.5 S (Svedberg) species with an apparent molar mass of 722 ± 100 kDa, consistent with a dodecameric species (Fig. 1c; Table 1). The frictional coefficient (f/f_0) of the dodecamer is calculated to be 1.51 (Table 1), which indicates that the oligomer is nonspherical or asymmetric in shape. There are also small amounts of other species observed with sedimentation coefficients of ~6 S, ~10 S, ~14 S, and ~30 S (corresponding to estimated molecular masses of ~130 kDa, ~260 kDa, ~400 kDa, and 1300 kDa, respectively). These species are likely to represent other oligomeric forms of PfM18AAP, potentially dimers, tetramers, hexamers, and octadecamers. Similar results are obtained at initial protein concentrations of 0.1 mg/ml and 1.0 mg/ml (Supplementary Fig. 1), thus indicating that the dodecamer of PfM18AAP is stable over the 10-fold concentration range.

Biochemical characterization and primary specificity of PfM18AAP

Recombinant PfM18AAP isolated from insect cells has been characterized previously. To ensure that our *E. coli* material retained similar enzymatic properties, we investigated the enzyme activity of the purified protein prior to crystallization. The presence of metal ions has been shown to enhance and/or inhibit the activity of M18 family members;^{16,20,24} hence, to determine the optimum assay conditions, we tested the enzyme efficiency against a variety of metal cations. We supplemented the activity assay buffer with different concentrations of various cations and assayed the activity of a fixed concentration of PfM18AAP against L-Glu-pNA. The results showed that Co(II) and Mn(II) enhanced activity while Zn(II), Mg(II), and Ni(II) had little effect (Fig. 2). At high concentrations (>1 mM), Zn(II), Mg(II), and Ni(II) were found to be inhibitory to activity (Fig. 2). Given this confirmation, we chose to supplement our assay buffers with 1 mM CoCl₂ to enhance the *in vitro* activity of PfM18AAP. Purified PfM18AAP efficiently cleaved L-Glu-pNA with k_{cat}/K_m values of $2265 \text{ M}^{-1} \text{ s}^{-1}$, consistent with published data.¹⁶

To determine the substrate specificity of PfM18AAP, we screened a recently developed fluorogenic substrate library of 19 natural and 41 nonnatural amino acids to obtain a substrate fingerprint for the enzyme.^{25,26} The results confirm a highly restricted selectivity with enzyme activity observed only in the presence of the natural amino acids Glu, Asp, and

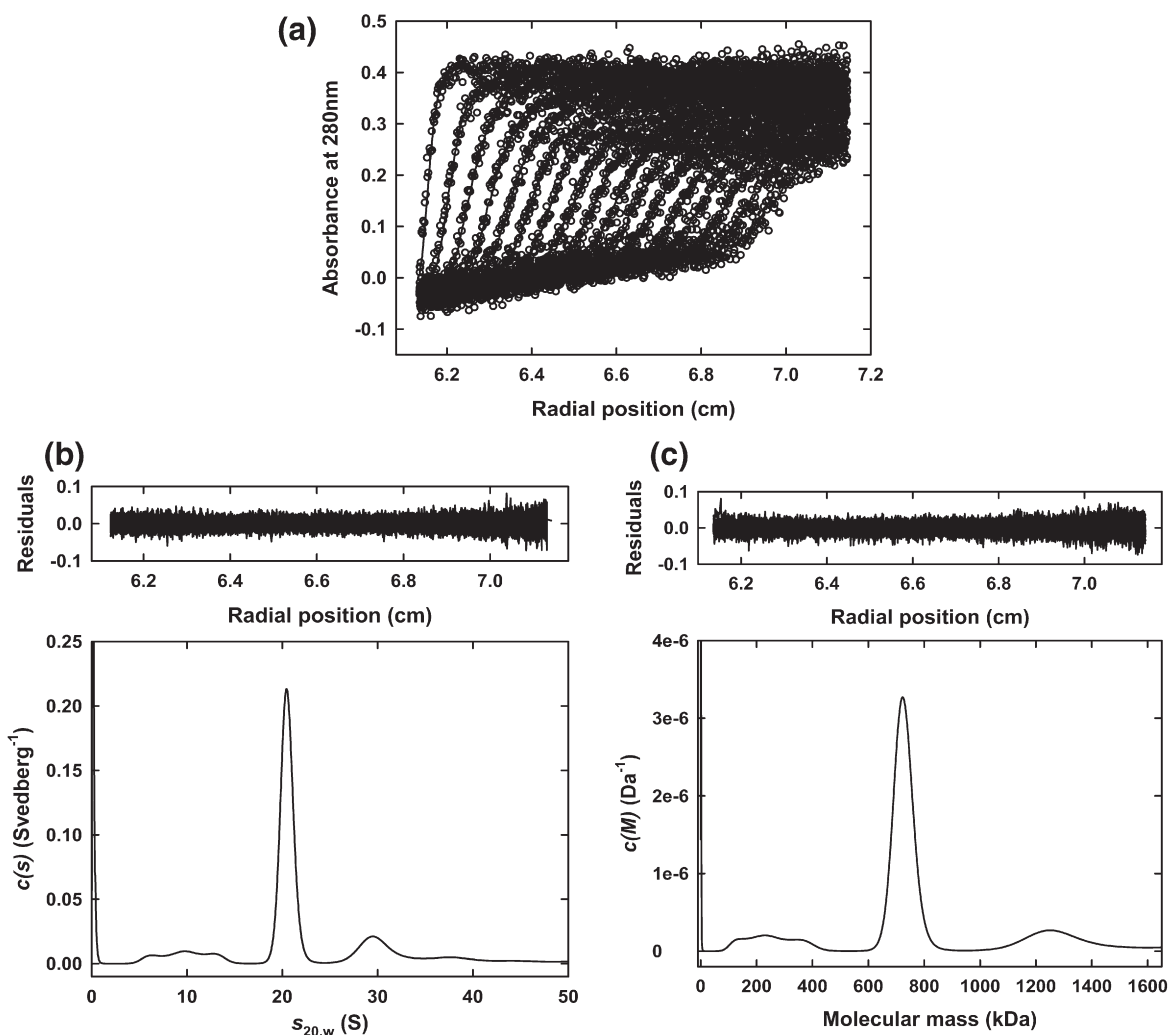


Fig. 1. Sedimentation velocity analysis of 0.5 mg/ml PfM18AAP. (a) Absorbance *versus* radial position profiles of PfM18AAP at different time points with the raw data shown as open symbols overlaid with the $c(s)$ distribution best-fit (continuous lines). (b) Continuous size, $c(s)$, distribution is plotted as a function of standardized sedimentation coefficient ($s_{20,w}$) for PfM18AAP generated using 500 species over a sedimentation coefficient range of 0.01 S to 50 S and at a confidence level (F ratio)=0.95. *Top*: Residuals for the $c(s)$ distribution best fits plotted as a function of radial position (cm). (c) Continuous mass, $c(M)$, distribution plotted as a function of molecular mass for PfM18AAP using 500 species ranging from 2 kDa to 2000 kDa and at a confidence level (F ratio)=0.95. *Top*: Residuals for the $c(M)$ distribution best fits plotted as a function of radial position (cm). The $c(s)$ and $c(M)$ distribution analyses were performed using the program SEDFIT^{21–23}, which is available at www.analyticalultracentrifugation.com.

Table 1. Hydrodynamic properties of PfM18AAP at 0.5 mg/ml, calculated from sedimentation velocity analyses

$s_{20,w}$ ^a (S)	M_r ^b (Da)	M^c (kDa)	f/f_0 ^d
20.4 ± 1.5	65,705	722 ± 100	1.51

^a Standardized sedimentation coefficient taken from the ordinate maximum of the $c(s)$ distribution (Fig. 1b).

^b Relative monomeric molecular mass (M_r) determined from the amino acid sequence.

^c Apparent molar mass (M) taken from the ordinate maximum of the $c(M)$ distribution (Fig. 1c).

^d Frictional ratio calculated using the \bar{v} method²¹ employing the experimentally determined $s_{20,w}$ and theoretical molecular mass of the PfM18AAP dodecamer (i.e., 788,460 Da).

Asn and the nonnatural amino acid 2,3-diaminopropionic acid (DAP) (Fig. 2). Subsequently, we determined the kinetic parameters for each substrate (Table 2) and show that the enzyme has no significant preference between a P1 Glu or Asp residue. Interestingly, a P1 Asn residue had an equivalent K_m value to the other natural amino acids; however, the general rate constant indicates that PfM18AAP turns this substrate over 10 times slower than either Glu/Asp. The nonnatural amino acid DAP was preferred over the natural Asn residue with a tighter K_m ($\sim 7 \times$ lower than any of the natural amino acids) and a faster rate constant (Table 2).

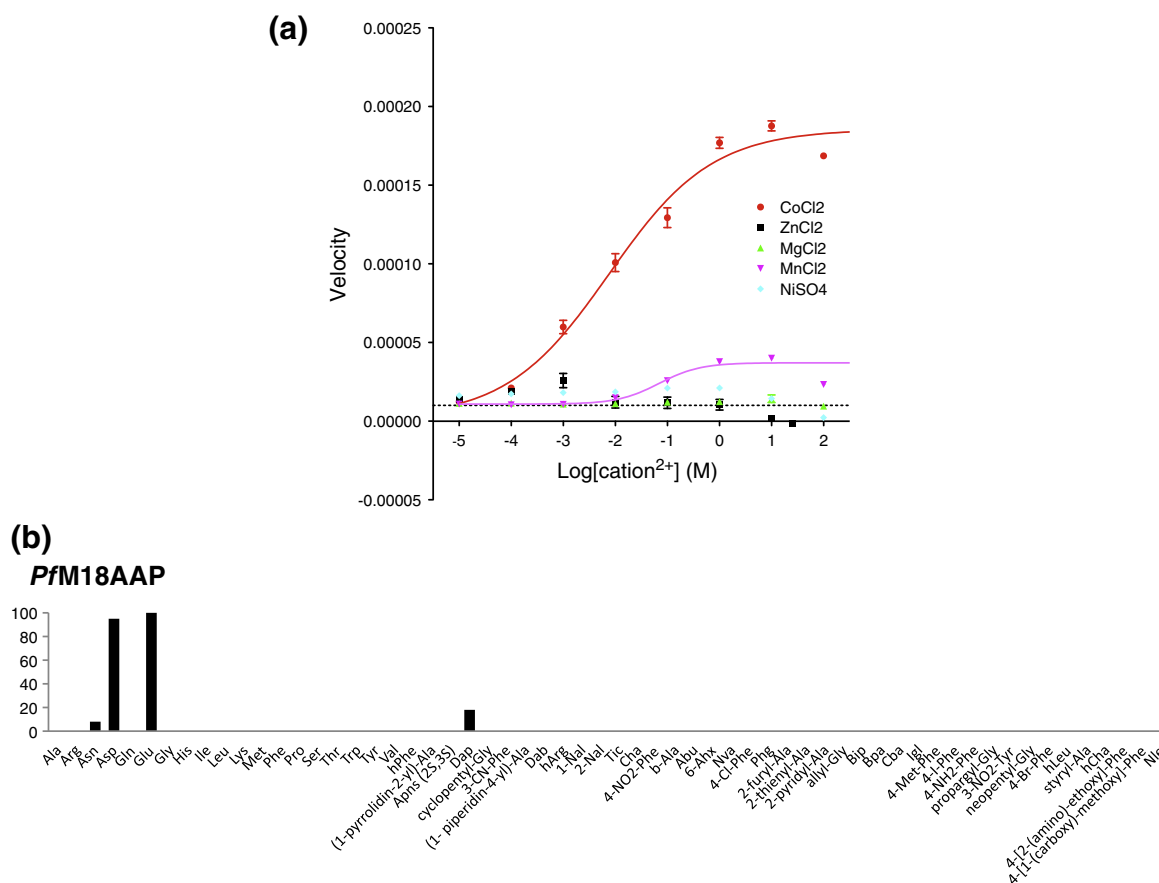


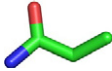
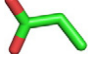
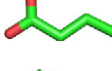

Fig. 2. Effect of various cations on activity of PfM18AAP and individual preference in S1 pocket of PfM18AAP toward natural and nonnatural amino acid substrates. (a) Co²⁺, Zn²⁺, Mg²⁺, Mn²⁺, and Ni²⁺ (0.01–100 mM) were incubated with 125 nM enzyme before the addition of substrate L-Glu-pNA (100 μM). (b) Screening of the 60-membered natural and nonnatural amino acid library. Enzyme activity was monitored via fluorescence at an excitation wavelength of 355 nm and an emission wavelength of 460 nm. The x-axis represents the abbreviated amino acid names. The y-axis represents the average relative activity expressed as a percentage of the best amino acid.

The 2.6-Å X-ray crystal structure of PfM18AAP

The structure of PfM18AAP was solved by molecular replacement and was refined to 2.6 Å and an R_{free} of 19.6% (Table 3; Fig. 3). The initial molecular

replacement solution was completed manually using Coot²⁷ and AutoBuster.²⁸ The final model had 4 molecules within the asymmetric unit, along with 8 zinc ions and 444 solvent molecules. Application of the space group symmetry matrices generated

Table 2. Kinetic parameters of specific PfM18AAP fluorogenic substrates

Substrate	P1 side chain	$k_{\text{cat}}/K_{\text{m}}$ ($\text{M}^{-1} \text{s}^{-1}$)	k_{cat} (s^{-1})	K_{m} (μM)
#3—NH ₂ -Asn-ACC		45.5 ± 3.47	0.028 ± 0.002	625.5 ± 47.2
#4—NH ₂ -Asp-ACC		360.1 ± 28.5	0.203 ± 0.017	563.7 ± 48.1
#6—NH ₂ -Glu-ACC		397.5 ± 30.5	0.200 ± 0.014	503.1 ± 33.6
#23—NH ₂ -Dap-ACC		96.6 ± 8.5	0.070 ± 0.008	72.46 ± 9.43

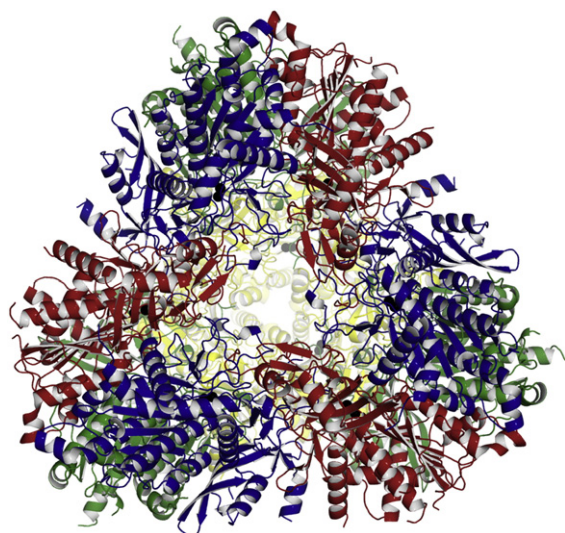


Fig. 3. The 2.6-Å X-ray crystal structure of PfM18AAP. The cartoon model of the 12 PfM18AAP monomers in their dodecameric arrangement. Monomers are colored by chains found in the asymmetric unit A (red), B (Blue), C (green), and D (yellow, at the back of the molecule in this view). The application of space group symmetry matrices generates the dodecamer with three copies of each chain (Chains D, D₁ and D₂ are at the back of the molecule in this view).

the unit cell that contained four complete dodecamers in a tetrahedral arrangement (Fig. 3). The quality of the model was evaluated using the MolProbity server (Table 3 and Refs. 29,30). Residues 166–169, 197–272, and 348–363 were omitted from the model due to absent electron density. Superposition of monomers A, B, C, and D reveals that they are essentially identical (average rmsd, 0.54 Å over 468 C^α atoms). Thus, unless stated otherwise, all structural analysis is confined to monomer A and numbering is as per PfM18AAP.

The overall structure of PfM18AAP monomer consists of two domains, a catalytic domain (residues 1–92 and 307–577) and a regulatory domain (residues 92–306) that together adopt a canonical M18 aminopeptidase fold²⁰ (Fig. 4a). The regulatory domain comprises two three-stranded antiparallel β-sheets and two helices. The arrangement of the β-sheets around helical residues 279–289 appears to interrupt what would otherwise be a central β-sheet through the domain. Extensive loop structures are noted in this domain (Fig. 4a). These loop regions contain two small helical turns and regions of disorder in our model, indicating the flexibility of this region of the protein. In each monomer present in the asymmetric unit, we noted that there was a break in connected density of the loop comprising residues 163–169 and that residues 197–272 also had no visible density (Fig. 4a).

The catalytic domain (residues 1–91 and 307–577) comprises a central seven-stranded β-sheet flanked on both sides by 10 α-helices. A second four-stranded antiparallel β-sheet occupies the space at the top of one side of the αβ arrangement (Fig. 4a). The helices form a funnel shape to the bottom of the molecule while the top of the domain remains largely open and unstructured with loop regions. The loop regions at the top of the domain, combined with residues from the central sheet, form the active-site cleft of each monomer. The active site of each monomer contained two zinc ions that are well coordinated by highly conserved residues His87, Asp325, Glu381, Asp435, and His535 (Fig. 4b; Supplementary Table 1). The active site of each monomer is a small closed cavity that has a small entrance (diameter, 5.6 Å) lined by residues Glu380, Glu381, Gly509, Ser510, and His535. The His161 loop residues of the dimeric partner (see below) further restrict the size of the active site and access to the zinc ions (Fig. 4c). Past the cations is a small substrate binding pocket lined by residues Asp435, Met436, His438, Lys463, Tyr470, Phe500, Thr511, and Met534.

Table 3. Data collection and refinement statistics

	PfM18AAP
<i>Data collection</i>	
Space group	P2 ₁ 3
Cell dimensions (Å)	a = b = c = 200.38 Å, β = 90.0°
Resolution (Å)	100.39–2.60 (2.74–2.60)
Total number of observations	789,865
Number of unique observations	82,063
Multiplicity	9.6 (6.8)
Data completeness (%)	100.0 (99.9)
$\langle I/\sigma_I \rangle$	8.0 (1.5)
R _{pin} (%) ^a	8.3 (58.0)
<i>Structure refinement</i>	
Non-hydrogen atoms	
Protein	15,002
Solvent	444
R _{free} (%)	19.6
R _{cryst} (%)	15.9
rmsd from ideality	
Bond lengths (Å)	0.009
Bond angles (°)	1.14
Ramachandran plot	
Favored (%)	97.9
Allowed (%)	100.0
B-factors (Å ²)	
Mean main chain	41.2
Mean side chain	46.3
Mean water molecule	38.5
rmsd bonded Bs	
Main chain	2.24
Side chain	5.04
MolProbity score	1.78 (99th percentile N = 6237, 2.60 Å ± 0.25 Å)

Values in parentheses refer to the highest-resolution shell.

^a Agreement between intensities of repeated measurements of the same reflections can be defined as: $\sum(I_{h,i} - \langle I_h \rangle) / \sum I_{h,i}$, where $I_{h,i}$ are individual values and $\langle I_h \rangle$ is the mean value of the intensity of reflection h .

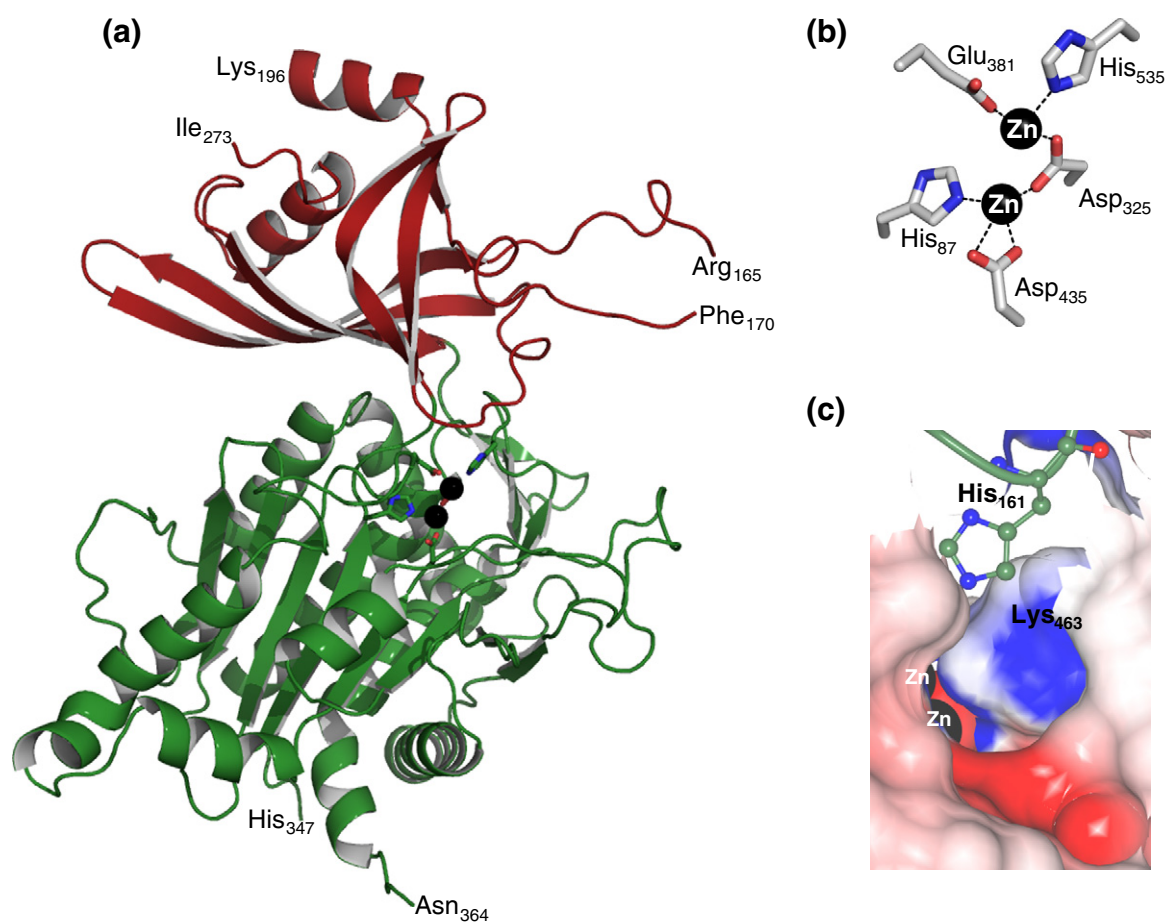


Fig. 4. Cartoon model of the *PfM18AAP* monomer and active site. (a) The *PfM18AAP* monomer with the regulatory domain (residues 92–306) shown in red and the catalytic domain (residues 1–91 and 307–577) shown in green. The zinc ions are shown as black spheres and zinc binding residues are shown. The last ordered residues at gaps in the model are indicated. (b) The zinc ion coordination in the active site of *PfM18AAP* with carbon atoms of zinc binding residues shown in gray and zinc as black spheres. Metallo-bonds are shown as black broken lines. (c) The active-site architecture restricts enzyme specificity. In this view, the electrostatic potential of Chain D is shown and a cartoon model of Chain C (green) and His161 is shown in stick. Chains A and B are excluded from view for clarity. The active-site zinc ions of Chain D are shown as black spheres and indicated. The Lys463 residue at the base of the S1 pocket is indicated.

The architecture of the active site immediately identifies the basis of the highly restricted substrate specificity of *PfM18AAP* for Glu, Asp, and Asn. The S1 pocket is small and buried immediately behind the zinc ions (Fig. 4c). It is lined predominantly with polar residues, and Lys463, which sits at the very base of this pocket (Fig. 4c), presumably provides the counter binding partner for the Glu/Asp substrates. A similar Lys residue is observed at the base of the homologous bovine M18AAP and is suggested to be involved in substrate binding.²⁰

The biological assembly of PfM18AAP

The biological assembly of the *PfM18AAP* is a dodecamer that forms a tetrahedron shape (Fig. 3). The assembly comprises both dimers and trimers of *PfM18AAP*. The four molecules in our asymmetric

unit form one dimeric and one trimeric interaction. The dimer interaction (Chains C and D) is dominated by interactions of the two regulatory domains and the smaller four-stranded sheet of the catalytic domain (Fig. 5a). The dimer interface is extensive and has a total of 57 H-bonds and four salt bridges (Supplementary Table 1) that buries a surface area of 3693 Å². As described for the mammalian enzyme, we also noted that the extended loop formed by residues 155–180 reaches into the active site of the dimer partner and contributes His161 to the catalytic center (Figs. 5a and 4c). Like the bovine and human M18 aminopeptidases, this residue does not contribute directly to zinc binding (distance is >4.5 Å) although it has been shown to be essential for functional activity of the enzyme.³¹ Interestingly, residues 165–170 are also located in this well-coordinated loop despite

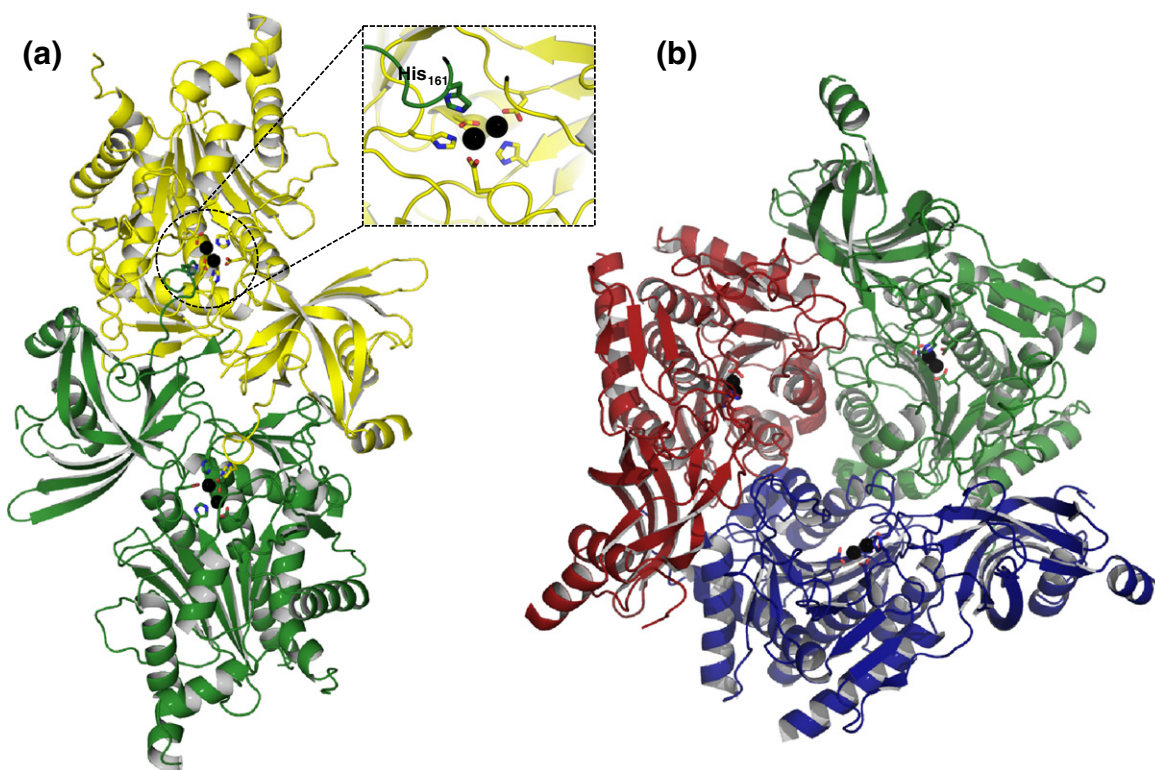


Fig. 5. The *PfM18AAP* enzyme has dimeric and trimeric assemblies. The asymmetric unit of the *PfM18AAP* X-ray crystal structure has four monomers present. Within these molecules, (a) a dimer is formed by Chains C (green) and D (yellow), with the dimeric partner contributing His161 to the active site (inset). (b) The trimer is formed by Chains A (red), B (blue), and C (green).

the fact that these residues have no visible electron density in our structure, which is in direct contrast to the bovine and human enzymes (3VAT and 3L6S, respectively). In these latter enzymes, the sequence of this loop is conserved, while in the *Plasmodium* enzyme, it contains a unique sequence that includes charged and polar residues (Supplementary Fig. 2).

Three catalytic domains interact to form the cone-shaped trimer of *PfM18AAP* (Chains A, B, and C) (Fig. 5b). The interactions of each monomer are predominantly mediated by the helices of each catalytic domain with the regulatory domain acting like a packing buffer between each catalytic domain. The interface between each chain is composed of 31 H-bonds and 10 salt bridges (Supplementary Table 1) and buries a surface area of 2144 Å². The three active sites are located in an equilateral triangle of ~33 Å from each other. The *PfM18AAP* trimers interact with each other to form the final biological dodecamer. These interactions are predominantly mediated by the regulatory domain of each monomer. The interior active-site cavity of *PfM18AAP* is composed therefore of four trimeric cones, each with their three active sites, and has a buried surface area of 36,226.7 Å². The dodecamer

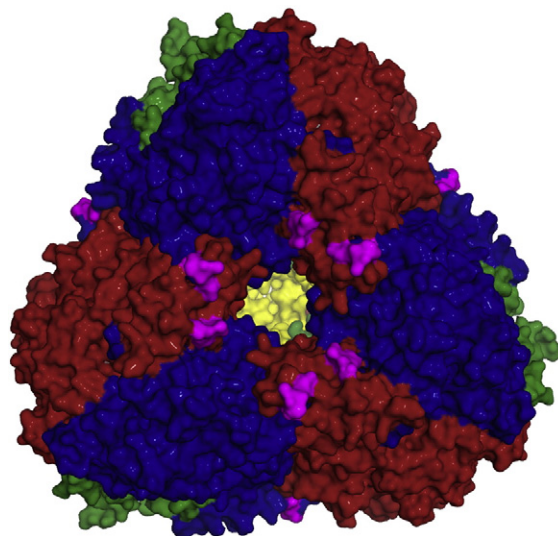


Fig. 6. The molecular surface of *PfM18AAP* shows active-site cavity entry pores. The dodecamer is colored as in Fig. 3. The last ordered residues (195 and 272) at the edge of the loop insertion are located at the periphery of the entrance tunnels and shown as magenta spheres.

has four large “pores”, measuring ~ 23 Å in diameter, which provide a tunnel to the inside of the active-site cavity (Fig. 6). Interestingly, the disordered loop (residues 197–272) of the symmetry-related chain (e.g., A, A₁ and A₂) shows that this loop would be located directly over the entrance of these large pores. The last ordered residue in the electron density clearly shows that the loops extend out from the periphery of the pore formation (Fig. 6). From this location, we propose that these loops would form a barrier to active-site entry, acting to regulate and control the enzyme's function. Secondly, given the size of this loop (>70 residues), we suggest that this extreme flexibility is the reason that chaperonin was required for protein expression and purification.

Structural and sequence comparison of PfM18AAP with DNPEP and bacterial homologs

The DNPEP family is characterized by a canonical fold that is shared in all structures solved to date (four bacterial: 2GLF, 2GLJ, 1YFE, and 2IJZ; one mammalian: 3VAT;²⁰ and one human: 3L6S). The parasite PfM18AAP comprises the same overall monomeric and quaternary structure as the enzyme family; however, the malaria enzyme exhibits unique changes in the loop regions of the regulatory domain and within the helices of the catalytic domain. The predominant difference between the PfM18AAP and other characterized enzymes to date is the presence of two large insertions, one in the regulatory domain (residues 195–260) and another in the catalytic domain (residues 346–363) (Supplementary Fig. 2).

The large insertion in the regulatory domain is unique to Pf and is not present in other *Plasmodium* homologs. In both the human and PfM18AAP structures, the last conserved residue before the insertion (residues 195) is the last ordered residue before a break in connective density. The structure of the human enzyme is missing the next 21 residues while the PfM18AAP enzyme is missing density for the next 76 amino acids. The mammalian enzyme (3VAT) has an equivalent break in connective density.²⁰ The bacterial enzymes vary in this trend and show a loss of the preceding helix but connective density throughout this region. The exception appears to be the *Pseudomonas aeruginosa* (2IJZ) enzyme that shows the same loss of the helix but has connective density.

The second insertion is located in the catalytic domain and is also disordered in the X-ray crystal structure. The last ordered residues are His347 and Asn264, which are located at the bottom of the molecule. They are not located in a position that would interfere with assembly of the trimer. The mammalian and human enzymes both show ordered connective density in this region.

Besides the sequence insertion, the catalytic domain of PfM18AAP shows significant movement of helices, in particular residues 393–414 that form a long helix at the bottom of the catalytic domain and undergo a 90° bend at residue 408 to accommodate a stretch of charged residues. In the trimeric structure, the helical bend in the three helices effectively closes the bottom of the trimer structure. The equivalent helix in the human and mammalian enzymes is not bent and the helix terminates at the position of the bend in PfM18AAP.

Discussion

Intracellular proteolysis requires very intricate biochemical machinery and is subject to spatial and temporal control in order to prevent the damage of proteins not destined for destruction. Some of these high-molecular-weight protease enzymes are self-compartmentalized; that is, they form cellular subcompartments through self-association and enclose inner cavities that harbor the proteolytic active sites.³² The inner active-site cavities localize multiple active sites, allowing proteolysis to occur efficiently. By restricting access to the cavity, the active sites are only accessible to unfolded polypeptides or peptides, and in this way, cytoplasmic proteins are protected from unwanted cleavage. Examples of this are cellular proteasomes and, increasingly, the aminopeptidases.^{33,34}

Aminopeptidases constitute a diverse set of protease enzymes with essential roles in cell maintenance, growth and development, defense, memory, tumor growth and metastasis, and antigen presentation. In Pf, a major malaria parasite of humans, the aminopeptidases are hypothesized to be involved in the catabolism of host Hb, a process essential for parasite survival. The Pf aminopeptidases PfA-M1 and PfA-M17 are validated drug targets; inhibitors have been shown to kill the parasite in culture and *in vivo*.^{11,13,14} In recent years, we have undertaken an extensive structural and functional characterization of these enzymes in an effort to understand their mechanism and specificity.^{13,14,26,35,36} Now, we can further this understanding by the comparison of the aminopeptidase PfM18AAP with these two enzymes with which it may function in concert. The PfA-M1 enzyme has a broad substrate specificity in comparison to PfA-M17 that is restricted to hydrophobic amino acids,^{13,26} although the two enzymes share a predominant preference for Leu, followed by Met, Trp, and Phe.²⁶ The PfM18AAP enzyme can efficiently cleave Glu and Asp and shows limited activity against Asn. The charge of the S1 pocket of PfM18AAP is dominated by the Lys463 residue, providing a negative charge to the small space and accounts for the affinity of both

Glu and Asp (Fig. 4c). The limited activity toward Asn and even the nonnatural DAP residue is presumably the result of the smaller size of these two residues as Gln, which has a similar charge to Asn but is of longer length, was not cleaved at all by the enzyme. Taken together, the specificities of the three enzymes, PfA-M1, PfA-M17, and PfM18AAP, are capable of working in concert to turnover peptides with any P1 residue except Pro (this residue may have its own specialized enzyme, prolyl aminopeptidase).

The 2.6-Å X-ray crystal structure of PfM18AAP reveals that the enzyme adopts an M18 aminopeptidase fold, forming a complex dodecameric biological assembly. The tetrahedron-shaped unit forms an active-site inner cavity that houses 12 active sites, each with a di-zinc metal center. It is interesting to make a parallel with the PfA-M17, which exhibits a highly restricted selectivity for leucine amino acids and exists as a homo-hexameric enzyme with each di-zinc active site surrounding a central cavity. Each site is also gated by a flexible loop, which may regulate substrate entry. Perhaps both enzymes are associated with a subcompartmentalized district of the cytoplasm where peptides are distributed and sorted for processing to free amino acids.

Studies on the mammalian enzyme show that the cations present in the active site can be altered depending on the concentration of metals in the environment. Enzymatic analysis of our PfM18AAP shows similar results with optimal enzyme activity in the presence of Cobalt(II) ions. The di-zinc active site immediately reveals the reason for the highly restricted substrate specificity with the small narrow cavity for the substrate binding pocket. The active site is formed by the contribution of a loop from the dimeric partner molecule. The coordination of this loop provides a functionally conserved histidine residue into the active site of its partner in the dimer. Although not directly contributing to zinc binding, this residue has been shown to be essential for the functional activity of the enzyme.³¹ In PfM18AAP, we noted that this loop had flexibility in all four chains in our model, implying flexibility throughout this region. This is in contrast to the mammalian and human enzymes that show a connected loop containing a small helical turn. This variation in the *Plasmodium* enzyme may provide substrate specificity or act as a “substrate clamp” when substrate is present (Supplementary Fig. 3).

PfM18AAP has a large insertion (residues 195–273) in the regulatory domain of the protein that is unique to *Pf* malaria. Studies show that this region of PfM18AAP is capable of binding to erythrocyte spectrin and other membrane proteins and suggests that PfM18AAP may have functional roles in the parasite that are additional to Hb digestion.¹⁸ Within this large insertion, a putative “spectrin binding domain” has been identified.¹⁸ It is sug-

gested that given PfM18AAP is expressed by mature schizonts prior to cell rupture, the enzyme could play a role in the release of the parasite from the host erythrocyte by unlocking its cytoskeleton. In PfM18AAP and human/mammalian homologs, this region of the protein shows considerable flexibility and cannot be modeled in structures. Interestingly, the last ordered residue in electron density or the “stumps” of this loop/insertion shows that this region would clearly be solvent exposed and is located at the periphery of large “pores” that access the buried active-site cavities (~23 Å in diameter, Fig. 6). We suggest that the access to this interior cavity is tightly controlled and must occur through local areas of flexibility—predominantly thought to be the loops of the regulatory domain.

The PfM18AAP adopts the canonical M18 aminopeptidase fold; however, it possesses unique features that distinguish it from other members of the family. These differences are predominantly found in the loop and surface-exposed regions of the protein, creating a unique enzyme designed for its role in the *Plasmodium*. Studies have shown that interruption of PfM18AAP activity has a fitness cost to the parasite.¹⁷ Specific potent inhibitors of PfM18AAP have not been identified; thus, it has not yet been possible to determine if PfM18AAP represents a novel drug target. Nevertheless, compounds designed to inhibit PfM18AAP may possess novel antimalarial properties that when combined with other agents provide a new avenue for chemotherapeutic control of malaria. The common catalytic core of the *Plasmodium* aminopeptidases PfA-M1, PfA-M17, and PfM18AAP makes them attractive targets for the rational design of antimalarial compounds that target all three enzymes.

Materials and Methods

Purification and molecular analysis of recombinant PfM18AAP

The full-length gene sequence PF11570c was synthesized using codons optimized for gene expression in *E. coli* (DNA2.0). Unique EcoRI and HindIII sites were introduced at the 5' and 3' ends, respectively. The synthesized gene was cloned into the bacterial expression vector pRSET-B, introducing an in-frame N-terminal hexahistidine tag to the construct. The *E. coli* expression construct containing the PfM18AAP gene was verified by DNA sequencing.

This construct was used to overexpress the PfM18AAP in BL21(C41) cells grown in AutoInduction Media. Overexpressed cells were lysed in 50 mM phosphate-buffered saline, pH 8.0, 300 mM NaCl, and 5% glycerol. Preliminary purification of the hexahistidine-tagged PfM18AAP was purified using a two-step purification process of Ni-NTA agarose column followed by size-exclusion chromatography on a Superdex 200 16/60 using an AKTApurify

high-throughput chromatography system§. The enzyme co-purified with chaperonin. Removal of the chaperonin was achieved by rebinding the co-purified proteins to Ni-NTA agarose column and then by extensive washing in dissociation buffer (50 mM Hepes, pH 8.0, 300 mM NaCl, 5% glycerol, 150 mM KCl, 10 mM MgCl₂, and 10 mM ATP). Post-dissociation, the column was returned to 50 mM Hepes, pH 8.0, 300 mM NaCl, and 5% glycerol, and purified Pfm18AAP was eluted by the addition of 500 mM imidazole. Purified Pfm18AAP was desalted and concentrated for use. Biochemical analysis indicated that kinetic parameters (k_{cat} , K_m , k_{cat}/K_m) of material purified and used in subsequent crystallization trials were equivalent to the recombinant material purified from insect cells as previously published.¹⁶

Analytical ultracentrifugation

Sedimentation velocity experiments were conducted in a Beckman model XL-I analytical ultracentrifuge at a temperature of 20 °C. A total of 380 µl of Pfm18AAP at 0.1 mg/ml, 0.5 mg/ml, and 1.0 mg/ml and 400 µl of reference solution [50 mM Hepes, 300 mM NaCl, and 5% (v/v) glycerol, pH 8.0] were loaded into conventional double-sector sapphire cells and mounted in a Beckman 4-hole An-60 Ti rotor. Samples were centrifuged at a rotor speed of 25,000 rpm, and the data were collected continuously at a single wavelength (280 nm), using a step size of 0.003 cm without averaging. Solvent density (1.0224 g/ml at 20 °C) and viscosity (1.159 cp at 20°), as well as estimates of the partial specific volume (0.73 ml/g at 20 °C) and hydration estimate (0.393 g/g), were computed using the program SEDNTERP.³⁷ Sedimentation velocity data at multiple time points were fitted to a continuous sedimentation coefficient [$c(s)$] distribution and a continuous mass [$c(M)$] distribution model^{21–23} using the program SEDFIT||.

Enzymatic analysis

Aminopeptidase activity was determined by measuring the hydrolysis of pNA from the chromogenic substrate L-glutamic acid γ -(4-nitroanilide) (L-Glu-pNA) (Sigma #G1135). Reactions were carried out in 96-well microtiter plates (200 ml total volume, 60 min, 37 °C) using a spectrofluorimeter (BMG LabTech Optima) with excitation at 405 nm. Enzyme was first added to 50 mM Tris-HCl, pH 7.5, supplemented with 1 mM CoCl₂ before the addition of 100 µM L-Glu-pNA. Initial rates were obtained at 37 °C over a range of substrate concentrations spanning K_m (32–2000 µM) and at fixed enzyme concentrations in 50 mM Tris-HCl, pH 7.5, supplemented with 1 mM CoCl₂. Experiments to characterize the metal dependency of the Pfm18AAP were carried out in the presence of various metal cations using 125 nM enzyme and 100 µM L-Glu-pNA.

Substrate profiling was achieved by the use of a fluorogenic substrate library containing 60 amino acids.²⁶

For convenience in solid-phase synthesis, we employed a 7-amino-4-carbamoylmethylcoumarin (ACC) fluorogenic leaving group in this library. Final screening of the library was carried out at 5 µM substrate and 120 nM enzyme as described above and release of free ACC fluorophore was monitored. The substrate concentration ranged from 10 to 1600 µM to measure the K_m values of the selected four substrates. Each experiment was repeated at least three times, and the average value with standard error was calculated. The concentration of dimethyl sulfoxide in the assay was less than 2% (v/v).

Crystallization and X-ray data collection

Purified Pfm18AAP enzyme was concentrated to ~9 mg/ml and crystals were grown using the hanging drop vapor diffusion method, with 1:1 (v/v) ratio of protein to mother liquor (0.5 ml well volume). The crystals appeared in 67% (v/v) methyl-2,4-pentanediol. Crystal morphology was very small cubes (typically <0.015 mm) that did not diffract on in-house X-ray sources or the macromolecular beamline (MX1) at the Australian Synchrotron. Data were collected at 100 K using the Micro Crystallography beamline (MX2) at the Australian Synchrotron and typically required 5-s exposure of 100% flux to achieve medium resolution. Data sets were collected on crystals at the Zn K-peak and at 20 eV above and below the peak to evaluate active-site Zn occupancy.

Diffraction images were processed using MOSFLM,³⁸ pointless,³⁹ and SCALA³⁹ from the CCP4 suite.⁴⁰ Five percent of each data set was flagged for calculation of R_{free} ⁴¹ with neither a sigma nor a low-resolution cutoff applied to the data. A summary of statistics is provided in Table 3. Subsequent crystallographic and structural analysis was performed using the CCP4i interface⁴² to the CCP4 suite,⁴⁰ unless stated otherwise.

Structure determination and refinement

Structure determination proceeded using the molecular replacement method and the program PHASER.⁴³ A search model was constructed from the crystal structure of DNPEP (Protein Data Bank ID: 3L6S) (probe identified using the FFAS server⁴⁴). A “mixed” model consisting of conserved side chains (all other non-alanine/glycine residues truncated at C γ atom) was then created using the SCRWL server.⁴⁴ Four clear peaks in both the rotation and translation functions were evident, and these packed well within the asymmetric unit. Together with the unbiased features in the initial electron density maps, the correctness of the molecular replacement solution was confirmed.

Initial structure refinement and model building proceeded using one molecule in the asymmetric unit [with the other noncrystallographic symmetry (NCS)-related molecules generated using NCS operators]. Maximum likelihood refinement using AutoBuster,²⁸ incorporating translation, liberation, and screw-rotation displacement (TLS) refinement, was carried out. After initial C α backbone trace was completed for one chain, NCS restraints were reduced to identical molecules in each hexamer and retained throughout to improve observation

§ <http://proteinexpress.med.monash.edu.au/index.htm>

|| Available at www.analyticalultracentrifugation.com

to parameter ratio. When restraints were relaxed, no change in electron density was noted. Simulated annealing composite omit maps were generated using CNS⁴⁵ omitting 5% of the model. All model building and structural validation were performed using Coot.²⁷ Solvent molecules were added only if they had acceptable hydrogen-bonding geometry contacts of 2.5 to 3.5 Å with protein atoms or with existing solvent and were in good $2F_o - F_c$ and $F_o - F_c$ electron density.

PyMOL was used to produce all structural representations[¶]. Hydrogen bonds (excluding water-mediated bonds) and salt bridges were calculated using PDBePISA.⁴⁶ The coordinates and structure factors are available from the Protein Data Bank (4EME). Raw data and images are available from TARDIS^{a, 47}

Acknowledgements

S.M. and M.A.P. are Australian Research Council Future Fellows. J.C.W. is an Australian Research Council Federation Fellow and National Health and Medical Research Council Honorary Principal Research Fellow. We thank the National Health and Medical Research Council and the Australian Research Council for funding support. J.P.D. is a Canadian Institutes of Health Research Canada Research Chair (TIER 1) in Infectious Diseases. We thank the Australian Synchrotron (MX-2) and the beamline scientists for beamtime and for technical assistance. We thank the Monash Platforms (Protein Production and Crystallization) for technical assistance.

Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2012.06.006>

References

- (2009). World Malaria Report. World Health Organisation.
- Wells, T. N. C., Alonso, P. L. & Gutteridge, W. E. (2009). New medicines to improve control and contribute to the eradication of malaria. *Nat. Rev. Drug Discov.* **8**, 879–891.
- Good, M. F. (2009). The hope but challenge for developing a vaccine that might control malaria. *Eur. J. Immunol.* **39**, 939–943.
- Enserink, M. (2008). Malaria. Signs of drug resistance rattle experts, trigger bold plan. *Science*, **322**, 1776.
- Enserink, M. (2010). Malaria's drug miracle in danger. *Science*, **328**, 844–846.
- Burki, T. (2009). Artemisinin resistance could endanger fight against malaria. *Lancet*, **9**, 213.
- Rosenthal, P. J. (2002). Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr. Opin. Hematol.* **9**, 140–145.
- Klemba, M., Gluzman, I. & Goldberg, D. E. (2004). A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J. Biol. Chem.* **279**, 43000–43007.
- Kolakovich, K. A., Gluzman, I. Y., Duffin, K. L. & Goldberg, D. E. (1997). Generation of hemoglobin peptides in the acidic digestive vacuole of *Plasmodium falciparum* implicates peptide transport in amino acid production. *Mol. Biochem. Parasitol.* **87**, 123–135.
- Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W. *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, **419**, 498–511.
- Skinner-Adams, T. S., Lowther, J., Teuscher, F., Stack, C. M., Grembecka, J., Mucha, A. *et al.* (2007). Identification of phosphinate dipeptide analog inhibitors directed against the *Plasmodium falciparum* M17 leucine aminopeptidase as lead antimalarial compounds. *J. Med. Chem.* **50**, 6024–6031.
- Harbut, M. B., Velmourougane, G., Dalal, S., Reiss, G., Whisstock, J. C., Onder, O. *et al.* (2011). Bestatin-based chemical biology strategy reveals distinct roles for malaria M1- and M17-family aminopeptidases. *Proc. Natl Acad. Sci. USA*, **108**, E526–E534.
- McGowan, S., Oellig, C. A., Birru, W. A., Caradoc-Davies, T. T., Stack, C. M., Lowther, J. *et al.* (2010). Structure of the *Plasmodium falciparum* M17 aminopeptidase and significance for the design of drugs targeting the neutral exopeptidases. *Proc. Natl Acad. Sci. USA*, **107**, 2449–2454.
- McGowan, S., Porter, C. J., Lowther, J., Stack, C. M., Golding, S. J., Skinner-Adams, T. S. *et al.* (2009). Structural basis for the inhibition of the essential *Plasmodium falciparum* M1 neutral aminopeptidase. *Proc. Natl Acad. Sci. USA*, **106**, 2537–2542.
- Rawlings, N. D., Barrett, A. J. & Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **40**, D343–D350.
- Teuscher, F., Lowther, J., Skinner-Adams, T. S., Spielmann, T., Dixon, M. W., Stack, C. M. *et al.* (2007). The M18 aspartyl aminopeptidase of the human malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **282**, 30817–30826.
- Dalal, S. & Klemba, M. (2007). Roles for two aminopeptidases in vacuolar hemoglobin catabolism in *Plasmodium falciparum*. *J. Biol. Chem.* **282**, 35978–35987.
- Lauterbach, S. B. & Coetzer, T. L. (2008). The M18 aspartyl aminopeptidase of *Plasmodium falciparum* binds to human erythrocyte spectrin in vitro. *Malar. J.* **7**, 161.
- Drag, M. & Salvesen, G. S. (2010). Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **9**, 690–701.
- Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K. & Kiser, P. D. (2012). Insights into substrate specificity and metal activation of mammalian tetrahedral aspartyl aminopeptidase. *J. Biol. Chem.* **287**, 13356–13370.

¶ <http://www.pymol.org>
^a www.tardis.edu.au

21. Perugini, M. A., Schuck, P. & Howlett, G. J. (2000). Self-association of human apolipoprotein E3 and E4 in the presence and absence of phospholipid. *J. Biol. Chem.* **275**, 36758–36765.
22. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619.
23. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys. J.* **82**, 1096–1111.
24. Wilk, S., Wilk, E. & Magnusson, R. P. (1998). Purification, characterization, and cloning of a cytosolic aspartyl aminopeptidase. *J. Biol. Chem.* **273**, 15961–15970.
25. Drag, M., Bogyo, M., Ellman, J. A. & Salvesen, G. S. (2010). Aminopeptidase fingerprints, an integrated approach for identification of good substrates and optimal inhibitors. *J. Biol. Chem.* **285**, 3310–3318.
26. Poreba, M., McGowan, S., Skinner-Adams, T., Trenholme, K. R., Gardiner, D. L., Whisstock, J. C. *et al.* (2012). Fingerprinting the substrate specificity of M1 and M17 neutral aminopeptidases of human malaria, *Plasmodium falciparum*. *PLoS ONE*, **2**, e31938.
27. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
28. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W. *et al.* (2011). BUSTER version X.Y.Z. (Ltd, G. P., ed), Cambridge, U.K.
29. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J. *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 12–21.
30. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
31. Wilk, S., Wilk, E. & Magnusson, R. P. (2002). Identification of histidine residues important in the catalysis and structure of aspartyl aminopeptidase. *Arch. Biochem. Biophys.* **407**, 176–183.
32. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell*, **92**, 367–380.
33. Dura, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M. & Franzetti, B. (2009). The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol. Microbiol.* **72**, 26–40.
34. Matsui, M., Fowler, J. H. & Walling, L. L. (2006). Leucine aminopeptidases: diversity in structure and function. *Biol. Chem.* **387**, 1535–1544.
35. Maric, S., Donnelly, S. M., Robinson, M. W., Skinner-Adams, T., Trenholme, K. R., Gardiner, D. L. *et al.* (2009). The M17 leucine aminopeptidase of the malaria parasite *Plasmodium falciparum*: importance of active site metal ions in the binding of substrates and inhibitors. *Biochemistry*, **48**, 5435–5439.
36. Stack, C. M., Lowther, J., Cunningham, E., Donnelly, S., Gardiner, D. L., Trenholme, K. R. *et al.* (2007). Characterization of the *Plasmodium falciparum* M17 leucyl aminopeptidase. A protease involved in amino acid regulation with potential for antimalarial drug development. *J. Biol. Chem.* **282**, 2069–2080.
37. Laue, T., Shah, B., Ridgeway, T. & SL, P. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. pp. 90–125, Royal Society of Chemistry, Cambridge, UK.
38. Leslie, A. G. W. (1992). Joint CCP4+SF-AMCB Newsletter on Protein Crystallography, No. 26.
39. Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
40. CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D*, **50**, 760–763.
41. Brunger, A. T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **49**, 24–36.
42. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 1131–1137.
43. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **61**, 458–464.
44. Jaroszewski, L., Rychlewski, L., Li, Z., Li, W. & Godzik, A. (2005). FFAS03: a server for profile–profile sequence alignments. *Nucleic Acids Res.* **33**, W284–W288.
45. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
46. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
47. Androulakis, S., Schmidberger, J., Bate, M. A., DeGori, R., Beitz, A., Keong, C. *et al.* (2008). Federated repositories of X-ray diffraction images. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **64**, 810–814.