Structural basis for the inhibition of the essential *Plasmodium falciparum* M1 neutral aminopeptidase

Sheena McGowan[a,b,1], Corrine J. Porter[a,b,1], Jonathan Lowther[c,1], Colin M. Stack[c,2], Sarah J. Golding[a], Tina S. Skinner-Adams[e], Katharine R. Trenholme[e], Franka Teuscher[e], Sheila M. Donnelly[c], Jolanta Grembecka[f], Artur Mucha[g], Pawel Kafarski[g], Ross DeGori[a], Ashley M. Buckle[a,b], Donald L. Gardiner[e], James C. Whisstock[a,b,3,4], and John P. Dalton[c,3,4]

[a]Department of Biochemistry and Molecular Biology and [b]Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton, Victoria, 3800, Australia; [c]Institute for the Biotechnology of Infectious Diseases, University of Technology Sydney, Level 6, Building 4, Corner of Thomas and Harris Street, Ultimo, Sydney, NSW 2007, Australia;[e]Malaria Biology Laboratory, Australian Centre for International and Tropical Health, Queensland Institute of Medical Research, 300 Herston Road, Herston, Brisbane, Queensland 4006, Australia; [f]Department of Molecular Physiology and Biological Physics, University of Virginia, 1300 Jefferson Park Avenue, Jordan Hall, Charlottesville, VA 22908; and [g]Department of Bioorganic Chemistry, Faculty of Chemistry, Wroclaw University of Technology, Wybrzeze, Wyspianskiego 27, 50-370 Wroclaw, Poland

Edited by Robert M. Stroud, University of California, San Francisco, CA, and approved December 19, 2008 (received for review July 31, 2008)

Plasmodium falciparum **parasites are responsible for the major global disease malaria, which results in >2 million deaths each year. With the rise of drug-resistant malarial parasites, novel drug targets and lead compounds are urgently required for the development of new therapeutic strategies. Here, we address this important problem by targeting the malarial neutral aminopeptidases that are involved in the terminal stages of hemoglobin digestion and essential for the provision of amino acids used for parasite growth and development within the erythrocyte. We characterize the structure and substrate specificity of one such aminopeptidase, *Pf*A-M1, a validated drug target. The X-ray crystal structure of *Pf*A-M1 alone and in complex with the generic inhibitor, bestatin, and a phosphinate dipeptide analogue with potent in vitro and in vivo antimalarial activity, hPheP[CH_2]Phe, reveals features within the protease active site that are critical to its function as an aminopeptidase and can be exploited for drug development. These results set the groundwork for the development of antimalarial therapeutics that target the neutral aminopeptidases of the parasite.**

drug design | malaria | structural biology | protease

There are 300–500 million cases of clinical malaria annually, and 1.4–2.6 million deaths. Malaria is caused by apicomplexan parasites of the genus *Plasmodium*, with *Plasmodium falciparum* the most lethal of the 4 species that infect humans. Clinical manifestations begin when parasites enter erythrocytes, and most antimalaria drugs, such as chloroquine, exert their action by preventing the parasite development within these cells (1). As a result of the rapid spread of drug-resistant parasites, there is a constant need to identify and validate new antimalarial targets.

Intraerythrocytic parasites have limited capacity for de novo amino acid synthesis and rely on degradation of host hemoglobin (Hb) to maintain protein metabolism and synthesis, and an osmotically stable environment within the erythrocyte (1–4). Within the erythrocytes, malaria parasites consume as much as 75% of the cellular Hb (1). Hb is initially degraded by the concerted action of cysteine-, aspartyl-, and metallo-endoproteases, and a dipeptidase (cathepsin C) within a digestive vacuole (DV) to di- and tripeptide fragments (5, 6). These fragments are suggested to be exported to the parasite cytoplasm, where further hydrolysis to release free amino acids takes place [supporting information (SI) Fig. S1; see refs. 7 and 8].

The release of amino acids involves 2 metallo-exopeptidases: an alanyl aminopeptidase, *Pf*A-M1 (9, 10), and a leucine aminopeptidase, *Pf*A-M17 (7, 11, 12). We have demonstrated that the aminopeptidase inhibitor bestatin, an antibiotic and natural analogue of the dipeptide Phe-Leu derived from the fungus *Streptomyces olivoreticul*, prevents *P. falciparum* malaria growth in culture (13, 14). More recently, it was shown not only that synthetic phosphinate dipeptide analogues that inhibit metallo-aminopeptidases prevented the growth of wild-type and the chloroquine-resistant parasites in culture but also that one compound, hPheP[CH_2]Phe (termed compound 4, Co4), reduced a murine infection of *Plasmodium chabaudi chabaudi* by 92% compared with controls (15, 16). Importantly, Co4 was found to cause no toxicity in these in vivo studies (16). Here, we functionally characterize *Pf*A-M1 and validate it as a target for the inhibitory activity of bestatin and hPheP[CH_2]Phe. We also present its 3D structure alone and complexed with both of these inhibitors, which sets the groundwork for the development of a previously undescribed class of antimalarial drugs.

Results and Discussion

Substrate Specificity of the *P. falciparum* Alanyl Aminopeptidase *Pf*A-M1. The M1 alanyl aminopeptidase gene (MAL13P1.56) (9), as annotated by PlasmoDB, is present in single copy and located on chromosome 13 of *P. falciparum*. The gene is 3,257 bp in length and encodes a protein, known as *Pf*A-M1, of 1,085 aa and 126 kDa. The full-length amino acid sequence of *Pf*A-M1 exhibits ≈70% sequence similarity with M1 aminopeptidase orthologues of the various rodent malaria species (*Plasmodium berghei*, *P. chabaudi chabaudi*, and *Plasmodium yoelii*); the sequences are most divergent at the large nonconserved N-terminal extension (≈194 aa), which contains 3 asparagine-rich low-complexity regions (LCRs) and a putative transmembrane domain (Fig. S2).

Author contributions: S.M., J.G., A.M., P.K., D.L.G., J.C.W., and J.P.D. designed research; S.M., C.J.P., J.L., C.M.S., S.J.G., T.S.S.-A., K.R.T., F.T., S.M.D., R.D., and J.P.D. performed research; J.G., A.M., and P.K. contributed new reagents/analytic tools; S.M., C.J.P., J.L., A.M.B., D.L.G., and J.P.D. analyzed data; and S.M., A.M.B., D.L.G., J.C.W., and J.P.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 3EBG, 3EBH, and 3EBI).

[1]S.M., C.J.P., and J.L. contributed equally to this work.

[2]Present address: School of Biomedical Health, Department of Medical Microbiology, University of Western Sydney, Narellan Road, Campbelltown, NSW 2560, Australia.

[3]J.C.W. and J.P.D. contributed equally to this work.

[4]To whom correspondence may be addressed. E-mail: james.whisstock@med.monash.edu.au or john.dalton@uts.edu.au.

This article contains supporting information online at www.pnas.org/cgi/content/full/0807398106/DCSupplemental.

BIOCHEMISTRY

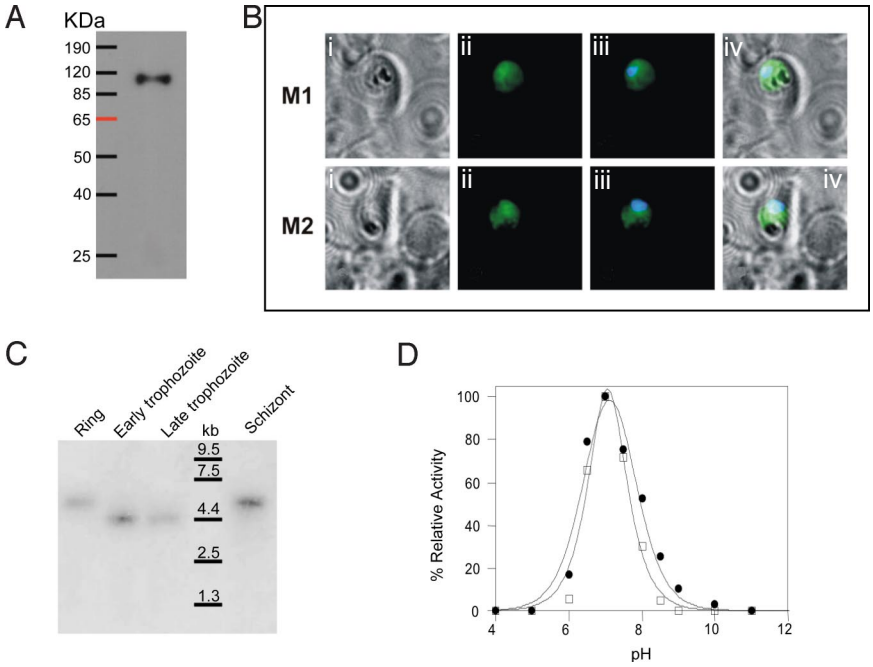

Fig. 1. Analysis of *Pf*A-M1 expression and activity. (*A*) Western blot of transgenic parasites expressing the product of the inserted transgene encoding *Pf*A-M1. The blot was probed with a monoclonal anti-c-myc primary antibody followed by horseradish peroxidase-conjugated anti-mouse immunoglobulin antibodies and visualized by enhanced chemiluminescence. (*B*) Indirect immunofluorescence of transgenic parasites stained with monoclonal anti c-myc primary antibody followed by anti-mouse cy2. (*i*) Bright field; (*ii*) anti-c-myc antibody; (*iii*) anti-c-myc/nuclear stain merged; (*iv*) merge of *i* and *iii*. The data show that the *Pf*A-M1 transgenic protein is localized to the parasite cytosol. (*C*) Northern blot analysis of stage-specific parasite RNA reveals that the endogenous *Pf*A-M1 is expressed by parasites at all developmental stages within the erythrocyte. Developmental stages are indicated at the top: R, ring stage parasites; ET, early trophozoite parasites; LT, late trophozoite parasites; and S, schizont stage parasites. Size difference observed between trophozoite and ring/schizont is due to different transcription initiation sites. (*D*) The pH optima for activity of r*Pf*A-M1 (circles) and native *Pf*A-M1 in soluble extracts of parasites (squares) measured by using the fluorogenic peptides substrate H-Arg-NHMec.

Using peptide antisera, Florent *et al.* (9) and Florent and coworkers (10), detected a 122-kDa *Pf*A-M1 in membrane fractions of malaria parasites. However, this membrane form appeared to be further processed to smaller active soluble forms of 96 and 68 kDa. In our studies, expression of the full-length *Pf*A-M1 was not successful, so a truncated form of the *P. falciparum* M1 aminopeptidase (residues 195-1085) correlating with the start of the M1 aminopeptidase of *Escherichia coli* PepN was prepared (17, 18). The construct lacked the 3 asparagine-rich LCRs and the putative transmembrane domain (Fig. S2) but was successfully expressed in *E. coli* and extracted as a soluble functional enzyme. The protein resolved as a major band at ≈100 kDa (predicted molecular mass 104.678 kDa) with a minor N-terminally truncated breakdown product of ≈55 kDa on reducing SDS/PAGE (Fig. S3).

The purified recombinant *Pf*A-M1 (r*Pf*A-M1) exhibited a similar retention time on HPLC size analysis as the *Pf*A-M1 in soluble extracts of *P. falciparum* parasites, both eluting between 80 and 100 kDa (Fig. S3). In support of the studies by Florent *et al.* (9), and Florent and coworkers (10), *Pf*A-M1 antisera localized the enzyme to the parasite cytosol (data not shown). Also, *P. falciparum* D10 parasites transfected with a plasmid (pHTB-*Pf*A-M1-cmycB) carrying the truncated *Pf*A-M1 gene expressed and trafficked an ≈115-kDa product (Fig. 1*A*) to the parasite cytosol (Fig. 1*B*). These transgenic parasites exhibited a 2.8-fold higher level of *Pf*A-M1 activity compared with D10 wild-type parasites, showing that the transgene product was functionally active within the parasite. Northern blotting experiments showed that the *Pf*A-M1 is expressed at all stages in the development of the malaria parasites in the erythrocyte (Fig. 1*C*).

The r*Pf*A-M1 protein displayed a broad specificity for amino acids, demonstrated by its ability to cleave peptide bonds involving hydrophobic, basic, and aromatic amino acids, which is consistent with M1 aminopeptidases of *E. coli* PepN (17, 18) and *Salmonella typhimurium* PepN (19). The most efficiently cleaved P1 substrates were Leu (k_{cat}/K_m = 4,607 $M^{-1}{\cdot}s^{-1}$), Ala (k_{cat}/K_m = 2,295 $M^{-1}{\cdot}s^{-1}$), Arg (k_{cat}/K_m = 1,491 $M^{-1}{\cdot}s^{-1}$), and Phe (k_{cat}/K_m = 924 $M^{-1}{\cdot}s^{-1}$) (Table S1). However, proline was a poor substrate with k_{cat}/K_m a least 1,000-fold less than Leu (Table S1), consistent with the catalytic mechanism described by Ito *et al.* (20), who suggest that 2 conserved glutamate residues stabilize the reaction intermediate through the formation of hydrogen bonds with the terminal amino group, NH_3. The presence of a proline in the P1 position may prevent such interactions forming due to the constraints imposed by the cyclized side chain. The r*Pf*A-M1 protein did not cleave after the acidic amino acids Glu and Asp, which is consistent with our suggestion that another enzyme, a cytosolic M18 aspartyl aminopeptidase, is responsible for the specific removal of these residues from peptides (21).

The efficient cleavage of the amino acid Arg is of particular interest, because this amino acid is not cleaved by the other major malaria aminopeptidase, *Pf*A-M17 (12). This observation facilitated the development of a specific assay for measuring *Pf*A-M1 in malaria extracts. Thus, by using Arg as a substrate, we showed that r*Pf*A-M1 and native *Pf*A-M1 activity within soluble extracts of malaria parasites exhibited optimal activity at pH 7.0, with <20% activity below pH 6.0, which is consistent with a cytosolic function for the enzyme (Fig. 1*D*) (12). Also, both r*Pf*A-M1 and native *Pf*A-M1 activity depends on a divalent metal ion as metal chelators (1 mM EDTA or *o*-phenanthroline) were found to inhibit their activity. Collectively, the above data demonstrate that despite lacking the N-terminal domain, r*Pf*A-M1 exhibits

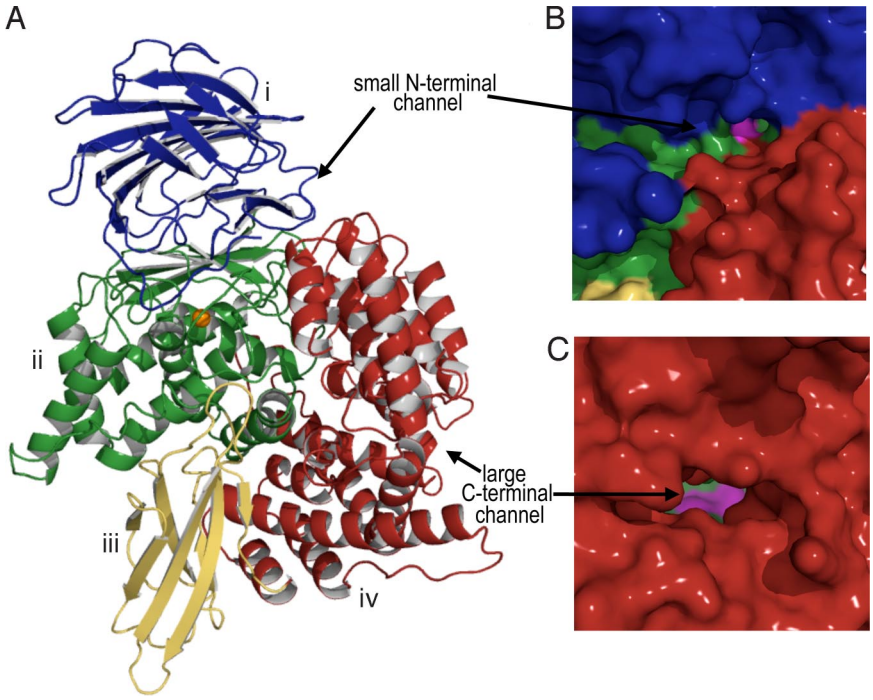

Fig. 2. The structure of r*Pf*A-M1. (*A*) Diagram of unbound r*Pf*A-M1 colored by domain: I (blue), II (green), III (yellow), and IV (red). (*B* and *C*) Molecular surface of *Pf*A-M1 (colored as in *A*) showing small (*B*) and large (*C*) openings to active site (active site cavity shown in magenta).

physicobiochemical characteristics identical to those of the native soluble malaria *Pf*A-M1 enzyme.

X-Ray Crystal Structure of r*Pf*A-M1 Confirms Bacterial Aminopeptidase Fold. Outside malaria, *Pf*A-M1 shows closest identity (≈35%) to the bacterial M1 aminopeptidases, which do not contain the extended N-terminal domain. We determined the X-ray crystal structure r*Pf*A-M1 to 2.1 Å (Table S2). The structure confirmed that r*Pf*A-M1 adopts the bacterial aminopeptidase N-fold (20, 22, 23), and comprises 26 α-helices and 7 β-sheets divided into 4 domains (Fig. 2*A*). The catalytic domain II (residues 392–649) adopts a thermolysin-like fold and contains the active site, incorporating the zinc-binding motif $H^{496}EYFHX_{17}KE^{519}$ and the well-conserved $G^{490}AMEN$ motif involved in substrate recognition (20, 22, 23). The catalytic zinc ion is coordinated by Nε2 atoms of His^{496} and His^{500}, the carboxyl Oε atom of Glu^{519}, and a water molecule that acts as the nucleophile that attacks the carbonyl carbon of the substrate (20). This water molecule forms a slightly longer metallo-bond with the zinc ion and is also coordinated by Glu^{497} and Glu^{463}. It is not clear what the catalytic base of the reaction is, but the structure suggests that it may be the Glu^{497} residue.

Inspection of the molecular surface of *Pf*A-M1 reveals 2 openings to the active site cavity. The first opening (N-terminal channel) comprises a shallow 8-Å-long groove at the junction of domains I and IV (Fig. 2*B*). The second and larger opening (C-terminal channel) is formed by the C-terminal domain IV, which comprises 8 pairs of α-helices arranged in 2 layers to form a cone-shaped superhelical structure. This domain interacts with the catalytic domain II and contains an ≈30-Å-long channel leading toward the active site (Fig. 2*C*). At the entrance, a helix (α14) with a 90° bend confines the pore size to ≈13-Å diameter.

The r*Pf*A-M1 Has a Buried Active Site That Does Not Require Conformational Change to Bind Substrate. Bestatin is an antibiotic originally isolated from filtrates of the fungus *S. olivoreticul*, but is now available in synthetic form (24). Its structure *N*-[(2*S*,3*R*)-3-amino-2-hydroxy-4-phenylbutanoyl]-L-leucine is an analogue of the dipeptide Phe-Leu. Thus, bestatin is a competitive inhibitor of many metallo-aminopeptidases (24). Accordingly, kinetic studies performed with r*Pf*A-M1 revealed that bestatin is an effective inhibitor of r*Pf*A-M1, with an inhibitory constant K_i = 478.2 nM.

More recently, a new class of metallo-aminopeptidase inhibitors, phosphinate dipeptide analogues, were designed by using a computer-aided approach by Grembecka *et al.* (15); 1 of these hPheP[CH_2]Phe, Co4, exhibited more potent inhibition of r*Pf*A-M1 (K_i = 79 nM) than bestatin. Both bestatin and Co4 had similar killing activity against malaria parasites in culture; IC_{50} bestatin range 8–14 μM and Co4 is 24–62 μM (Fig. 3). However, bestatin only showed low-level (33%) antimalarial activity in murine (*P. c. chabaudi*) models, whereas Co4 was much more potent and reduced infection by 92% compared with controls (16).

To elucidate the mechanism by which bestatin and Co4 bound to

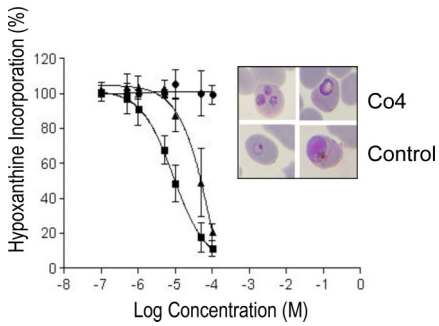

Fig. 3. The inhibitory effect of the aminopeptidase inhibitor bestatin (squares), and Co4 (triangles) on *P. falciparum* clone 3D7 growth in culture compared with parasites grown in the absence of inhibitor (circles). Data are presented as mean ± SD of 3 independent, triplicate experiments. Parasites grown in the presence of bestatin or Co4 for 24 h exhibit cellular damage and stunted development compared with control parasites grown in the absence of drug (*Inset*).

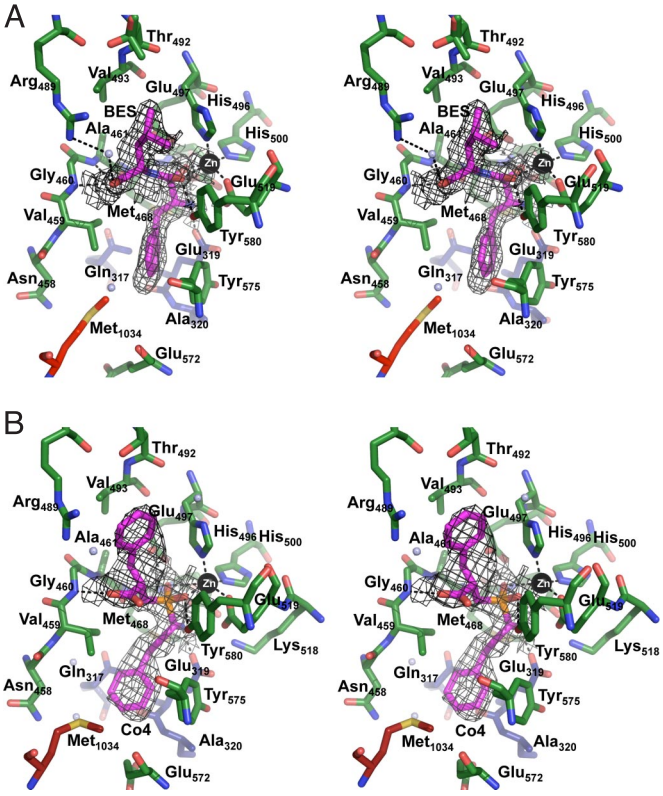

Fig. 4. Stereo diagram of inhibitors binding to active site of r*Pf*A-M1. (*A*) 1.65-Å r*Pf*A-M1-BES and (*B*) 2.0-Å r*Pf*A-M1-Co4 active site showing inhibitors bound in the active site. Inhibitors (BES/Co4) are colored in magenta. Carbon atoms of residues are colored by domain: I (blue), II (green), IV (red), and Zn is shown as black sphere. Water molecules are shown as blue spheres. Hydrogen bonds are indicated (dashed lines). Electron density is a composite omit map contoured at 1.0 σ calculated by using a model containing only *Pf*A-M1 atoms (no zinc, ligand, or water was included in the calculation).

the active site of r*Pf*A-M1, we determined the X-ray crystal structure of r*Pf*A-M1 complexed to each inhibitor to 1.65 and 2.0 Å, respectively (Table S2). Given that the active site of the enzyme is buried in the middle of the protein, we were particularly interested in understanding whether conformational change is required for substrate/inhibitor entry. Interestingly, comparison of the unbound and ligand-bound structures revealed that the enzyme does not undergo any global conformational rearrangements on binding either inhibitor [rmsd of 0.14 (bestatin) and 0.16 Å (Co4) >890 Cα atoms]. The omit electron density of ligands within the active site was well defined (Fig. 4), and the dipeptide analogues slot neatly into the large catalytic cavity without causing any major alterations to active site residues.

In *E. coli* aminopeptidase N protein Met^{454}, positioned immediately preceding the GAMEN exopeptidase motif, is postulated to function as a cushion to accept substrates (20), altering the size of the active site pocket. This residue is conserved in bacterial aminopeptidases; however, the equivalent position (469) in eukaryote M1 aminopeptidases is commonly a smaller valine or alanine residue. We noted no movement of the r*Pf*A-M1 Val^{469} residue to accommodate either the bestatin or Co4 molecule. Indeed, only 2 residues showed significant movement after ligand binding, the Glu^{526} side chain that moves away from the active site, removing what would otherwise form a close contact with P1′ position of either inhibitor (Leu moiety of bestatin or the Phe ring of Co4). However, in the unbound form of r*Pf*A-M1, the Met^{1034} residue adopts 2 alternative conformations in the ligand-bound structures, this side chain adopts a single conformation and forms packing interactions with the inhibitor moieties (P1 position).

The lack of any large scale conformational changes in the bound structure suggests that it is unlikely a conformational change is required for active site access. Indeed, initial attempts to obtain data from a ligand-bound complex involved soaking the inhibitor into r*Pf*A-M1 crystals. This procedure caused no deterioration of crystal quality (however, it did result in variable occupancy of the active site, hence cocrystallization was used), also indicating that it is unlikely that large-scale rearrangements of the protein molecules are required for substrate/inhibitor binding. Therefore, we suggest that the active site of this monomeric aminopeptidase is buried, substrate access is achieved by means of the C-terminal domain vortex, and that control of substrate hydrolysis can be achieved, and depends on, the size of this channel.

Both inhibitors primarily interact with the enzyme through a pentahedral coordination with the catalytic zinc ion (Fig. 4). Aside from bonds with His^{496}, His^{500}, and Glu^{519}, the hydroxyl group (O2) and the carbonyl group (O3) of bestatin and the central phosphoryl O atoms of Co4 are coordinated to the zinc ion (Fig. S4). Pentahedral zinc coordination, rather than the tetrahedral geometry observed in the unbound structure, is required for the transition state of the enzyme that exists after the nucleophilic attack at the carbonyl carbon of the substrate. The bestatin carbonyl carbon (O3) and the Co4 central phosphoryl O atoms (O3 and O4; Fig. S4) also form hydrogen bonds with the side chain of Tyr^{580}, stabilizing this reaction intermediate. A *cis*-peptide (Glu^{319}-Ala^{320}) allows the side chain of Glu^{319} to extend into the active site, where it forms a hydrogen bond with the N2 atom of bestatin and the amino group (NH_2) of Co4 (Fig. 4). The GAMEN recognition motif residues also contribute hydrogen bonds to ligand binding with the side chain of Glu^{463} and main-chain amide of Gly^{460}, with both inhibitors (Fig. 4). Bestatin also forms a hydrogen bond with the main-chain amide of Ala^{461} (Table S3). When bound, Co4 resides in an extended conformation (≈11 Å from Phe-ring centroids) and has a total contact area of 19.9Å^2 (compared with bestatin total buried surface area of 17.7 Å^2). The 2 Phe rings of Co4 form favorable hydrophobic interactions; the Phe ring at P1′ position (C1-C6, Fig. S4) packs against side-chains of residues Arg^{489}, Thr^{492}, and Val^{493}, whereas the Phe ring at position P1 (C10-C16, Fig. S4) forms hydrophobic contacts with side chains of Gln^{317}, Val^{459}, Met^{462}, Tyr^{575}, and Met^{1034}. In bestatin, the Leu moiety is flanked by Val^{459} and Thr^{492}; 2 water molecules also coordinate the O1 atom (bestatin) with the Nη atoms of Arg^{489}. The Phe-like amino acid is cushioned by hydrophobic interactions with Gln^{317}, Met^{462}, Tyr^{575}, and Met^{1034}. In total, bestatin makes 8 hydrogen bonds with the active site, whereas Co4 forms only 6 hydrogen bonds (Table S3). However, Co4 is further stabilized by 10 residues contributing to its hydrophobic packing within in the active site core in contrast to bestatin, which has only 7 such interactions with only 1 residue (Thr^{492}), contributing to the stability of the leucyl moiety in the P1′ position.

Bestatin is a natural analogue of the dipeptide Phe-Leu, but is a much weaker inhibitor of *Pf*A-M1 than Co4. When analyzing the active site contacts of the 2 ligands, a possible distinction between potency to Co4 of bestatin may be due to a lack of contacts made by the leucyl side chain of bestatin, leaving a open hydrophobic cleft that is filled when Co4 is bound. Interestingly, the phosphinate dipeptide hPheP[CH_2]Gly (Co2) is a significantly poorer inhibitor than Co4 [having a 13-fold greater K_i of 1,030.0 nM than Co4 (hPheP[CH_2]Phe)] (16). We suggest that this reduced potency is probably also due to the lack of contacts at the P1′ position within this large active-site cavity. Because of its high K_i, compound hPheP[CH_2]Gly does not exhibit antimalarial activity in vitro. However, these data clearly highlight the impact such ligand-binding analysis may have for structure-based drug design in conjunction with biological analysis.

The active site of *PfA*-M1 is buried deep in the structure, providing a solution to the problem of controlling appropriate substrate entry, as well as preventing hydrolysis of unwanted substrates. By comparison with the available bacterial aminopeptidases structures (20, 22, 23), *PfA*-M1 can, thus, be considered to adopt a "closed" conformation. Interestingly, structural studies on the homologous tricorn-interacting factor 3, F3, reveals this protease undergoes a substantial conformational changes on substrate entry, where domain IV swings away from domain II, exposing the active site. However, for the aminopeptidase N structures available, including *Pf*A-M1, there is no movement of the C-terminal domain noted, even in the presence of inhibitors. We argue that in *PfA*-M1, the large C-terminal channel functions to permit substrate entry whereby Hb-derived peptides access the buried active site leaving the smaller sized opening for exit of released amino acids. A maximal pore size of 10 Å has been proposed to serve as the entry point in other multimeric proteases structures (22, 25); here, we suggest that for such an opening is sufficient to allow access to the buried active site of this monomeric protease.

Given our detailed knowledge of the active site of the *PfA*-M1 enzyme, high-throughput screening of chemical libraries followed by medicinal chemistry would allow the development of additional lead antimalarial compounds. The static nature of this family of proteases allows the opportunity for a 2-fold approach to rational drug design; discovery of lead compounds that bind to and irreversibly block the active site, and/or compounds that block egress of digestion products by occluding the N-terminal channel, both essentially trapping the molecule in an inactive state.

Conclusions

There is a paucity of new antimalarial drugs entering the development pipeline. Our report of the unique active site structure of *Pf*A-M1, in complex with the antimalarial inhibitors bestatin and Co4, provides the groundwork for the de novo discovery of a previously unrecognized class of antimalarials by using high-throughput chemical screening and medicinal chemistry. Co4 also inhibits the second important neutral aminopeptidase of malaria, *Pf*A-M17 (12, 26), creating the possibility of developing a 2-target or combination therapy that would be more resilient to the emergence of drug-resistant malaria parasites.

Materials and Methods

Parasite Preparation. *P. falciparum* clone D10 was cultured as described (27). For experiments investigating the stage specific expression of *Pf*A-M1, parasites were synchronized by using 2 rounds of sorbitol treatment (28), and stage specific parasites harvested at ring, trophozoite, and schizont stage. Details of parasites, and methods of immunoblotting, Northern blotting, and transfection can be found in *SI Methods*.

Enzymatic Analysis. Aminopeptidase activity was determined by measuring the release of the fluorogenic leaving group, 7-amino-4-methyl-coumarin (NHMec) from the fluorogenic peptide substrates (*SI Methods*).

In Vitro Sensitivity of *P. falciparum* Malaria Parasites to Aminopeptidase Inhibitors. The in vitro sensitivity of each parasite population to bestatin, Co4 and Co2 was determined by using [^{3}H]-hypoxanthine incorporation (for further details see *SI Methods*) (29). IC_{50} values were determined by linear interpolation of inhibition curves (*SI Methods*).

Crystallization, X-Ray Data Collection, Structure Determination, and Refinement. A truncated form of *Pf*A-M1 (residues 195-1085, r*Pf*A-M1) was purified from *E. coli* (*SI Methods*); r*Pf*A-M1 at 5 mg/mL in 50 mM Hepes pH 8.5; 300 mM NaCl 5% (vol/vol) glycerol was crystallized in 22% (vol/vol) polyethylene glycol 8000, 10% (vol/vol) glycerol, 0.1 M Tris (pH 8.5), and 0.2 M $MgCl_2$. Crystals of the ligand bound r*Pf*A-M1 complexes were obtained by cocrystallization under similar conditions in the presence of 1 mM ligand. The diffraction data for the unbound, bestatin-bound, and Co4-bound protease were collected to 2.1-, 1.65-, and 2.0-Å resolution, respectively (Table S2). Crystallographic analysis was performed by using CCP4i (30–33). The structure was determined by using the program PHASER (34) (using 2GTQ.pdb as a search probe; see ref. 23). Refinement was performed by using REFMAC (35). All model building was done by using COOT (for further crystallization details, see *SI Methods*) (36).

ACKNOWLEDGMENTS. J.C.W. is an Australian Research Council (ARC) Federation Fellow; A.M.B. is a National Health and Medical Research Council (NHMRC) Senior Research Fellow; and C.J.P. is a NHMRC Training (Peter Doherty) Fellow. J.P.D., K.R.T., and D.G. are supported by the ARC Discovery Project Grant DP0666128; K.R.T. and D.G. by the NHMRC Program Grant 290208 and by generous donation from Mark Nicholson, Alice Hill, and the Tudor Foundation; and T.S-A. by a University of Queensland Postdoctoral Fellowship and a Ramaciotti Development grant.

1. Rosenthal PJ (2003) Antimalarial drug discovery: Old and new approaches. *J Exp Biol* 206:3735–3744.
2. Divo AA, Geary TG, Jensen JB (1985) Oxygen- and time-dependent effects of antibiotics and selected mitochondrial inhibitors on *Plasmodium falciparum* in culture. *Antimicrob Agents Chemother* 27:21–27.
3. Lew VL, Macdonald L, Ginsburg H, Krugliak M, Tiffert T (2004) Excess hemoglobin digestion by malaria parasites: A strategy to prevent premature host cell lysis. *Blood Cells Mol Dis* 32:353–359.
4. Liu J, Istvan ES, Gluzman IY, Gross J, Goldberg DE (2006) *Plasmodium falciparum* ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proc Natl Acad Sci USA* 103:8840–8845.
5. Klemba M, Gluzman I, Goldberg DE (2004) A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J Biol Chem* 279:43000–43007.
6. Rosenthal PJ (2002) Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr Opin Hematol* 9:140–145.
7. Curley GP, *et al.* (1994) Aminopeptidases from *Plasmodium falciparum*, *Plasmodium chabaudi chabaudi* and *Plasmodium berghei*. *J Eukaryot Microbiol* 41:119–123.
8. Kolakovich KA, Gluzman IY, Duffin KL, Goldberg DE (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.
9. Florent I, *et al.* (1998) A *Plasmodium falciparum* aminopeptidase gene belonging to the M1 family of zinc-metallopeptidases is expressed in erythrocytic stages. *Mol Biochem Parasitol* 97:149–160.
10. Allary M, Schrevel J, Florent I (2002) Properties, stage-dependent expression and localization of *Plasmodium falciparum* M1 family zinc-aminopeptidase. *Parasitology* 125:1–10.
11. Gavigan CS, Dalton JP, Bell A (2001) The role of aminopeptidases in hemoglobin degradation in *Plasmodium falciparum*-infected erythrocytes. *Mol Biochem Parasitol* 117:37–48.
12. Stack CM, *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.
13. Gavigan CS, Machado SG, Dalton JP, Bell A (2001) Analysis of antimalarial synergy between bestatin and endoprotease inhibitors using statistical response-surface modelling. *Antimicrob Agents Chemother* 45:3175–3181.
14. Nankya-Kitaka MF, Curley GP, Gavigan CS, Bell A, Dalton JP (1998) *Plasmodium chabaudi chabaudi* and *P. falciparum*: Inhibition of aminopeptidase and parasite growth by bestatin and nitrobestatin. *Parasitol Res* 84:552–558.
15. Grembecka J, Mucha A, Cierpicki T, Kafarski P (2003) The most potent organophosphorus inhibitors of leucine aminopeptidase. Structure-based design, chemistry, and activity. *J Med Chem* 46:2641–2655.
16. Skinner-Adams TS, *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.
17. Chandu D, Nandi D (2003) PepN is the major aminopeptidase in *Escherichia coli*: Insights on substrate specificity and role during sodium-salicylate-induced stress. *Microbiology* 149:3437–3447.
18. Golich FC, Han M, Crowder MW (2006) Over-expression, purification, and characterization of aminopeptidase N from *Escherichia coli*. *Protein Expr Purif* 47:634–639.
19. Kumar A, Nandi D (2007) Characterization and role of peptidase N from *Salmonella enterica* serovar Typhimurium. *Biochem Biophys Res Commun* 353:706–712.
20. Ito K, *et al.* (2006) Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from *Escherichia coli* and conformational change of methionine 260 involved in substrate recognition. *J Biol Chem* 281:33664–33676.
21. Teuscher F, *et al.* (2007) The M18 aspartyl aminopeptidase of the human malaria parasite *Plasmodium falciparum*. *J Biol Chem* 282:30817–30826.
22. Addlagatta A, Gay L, Matthews BW (2006) Structure of aminopeptidase N from *Escherichia coli* suggests a compartmentalized, gated active site. *Proc Natl Acad Sci USA* 103:13339–13344.
23. Nocek B, Mulligan R, Bargassa M, Collart F, Joachimiak A (2007) Crystal structure of aminopeptidase N from human pathogen *Neisseria meningitidis*. *Proteins* 273:273–279.
24. Scornik OA, Botbol V (2001) Bestatin as an experimental tool in mammals. *Curr Drug Metab* 2:67–85.

25. Russo S, Baumann U (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J Biol Chem* 279:51275–51281.
26. Gardiner DL, Trenholme KR, Skinner-Adams TS, Stack CM, Dalton JP (2006) Overexpression of leucyl aminopeptidase in *Plasmodium falciparum* parasites. Target for the antimalarial activity of bestatin. *J Biol Chem* 281:1741–1745.
27. Trager W, Jensen JB (1976) Human malaria parasites in continuous culture. *Science* 193:673–675.
28. Lambros C, Vanderburg JP (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol* 65:418–420.
29. Geary TG, Delaney EJ, Klotz IM, Jensen JB (1983) Inhibition of the growth of *Plasmodium falciparum* in vitro by covalent modification of hemoglobin. *Mol Biochem Parasitol* 9:59–72.
30. Leslie AGW (1992) *Joint CCP4 + ESF-EAMCB Newslett Prot Crystallogr* (Daresbury Laboratory, Warrington, UK), 26.
31. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62:72–82.
32. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D* 59:1131–1137.
33. CCP4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
34. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D* 61:458–464.
35. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
36. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
37. Davis IW, *et al.* (2007) Molprobity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–W383.

Supporting Information

McGowan *et al.* 10.1073/pnas.0807398106

SI Methods

Parasites and Preparation of Parasite Extract. *Plasmodium falciparum* clone D10 was cultured as described (1). For experiments investigating the stage-specific expression of *Pf*A-M1, parasites were synchronized by using 2 rounds of sorbitol treatment (2), and stage-specific parasites were harvested at ring, trophozoite, and schizont stage.

The *P. falciparum* M1 Alanyl Aminopeptidase Gene. Because of the high AT-richness of genes encoding malarial proteins *Pf*A-M1 was chemically synthesized by GeneArt using codons for optimized gene expression in the yeast *Pichia pastoris*. Potential N-linked glycosylation sites were removed during gene synthesis by replacing the asparagine of all Asn-Xaa-(Thr/Ser) motifs with glutamine (Gln, Q). However, expression with the codon optimized *Pf*A-M1 in *P. pastoris* was unsuccessful; therefore, we turned to *Escherichia coli* as the expression host. In this host, expression of the full-length gene did not result in a functionally-active enzyme and, therefore, a truncated form of the *P. falciparum* M1 aminopeptidase (residues 195-1085, r*Pf*A-M1; see Fig. 1), which removed N-terminal low-complexity regions and a putative transmembrane region was prepared by PCR amplification using the synthesized gene as a template followed by directional cloning into the bacterial expression vector pTrcHis2B (Invitrogen). The primers used were M1 forward 5′-CTGCAGAACCAAAGATCCAC-3′, and M1 reverse 5′-GGTACCTCA*ATGATGATGATGATGATG*TGGGCCCA-ACTTGTTTGT-3′. Unique PstI and KpnI sites (underlined) were introduced at the 5′ and 3′ ends of the amplified product. A C-terminal His-tag was introduced into the M1 reverse primer (italics). The construct housing the truncated PfA-M1 gene was verified by DNA sequencing. The verified construct was then transformed into competent BL21 cells. Expression was carried out as per the supplier's instructions (Invitrogen).

Purification and Molecular Analysis of Recombinant M1 Alanine Aminopeptidase (r*Pf*A-M1). The r*Pf*A-M1 was purified from BL21 cells by using a Ni NTA-agarose column as previously described (3). Native *Pf*A-M1 from soluble extracts of malaria parasites (>80% trophozoites) was prepared as described previously (3). Size analysis was performed by (*i*) SDS/12% PAGE analysis, and (*ii*) gel-permeation chromatography using a Amersham Pharmacia Biotech Smart System HPLC equipped with a Superdex-200 gel-filtration column. The proteins (40 μL containing 10–20 μg of protein) were passed through the column at a flow rate of 40 μL/min, and fractions were analyzed for aminopeptidolytic activity toward the fluorogenic substrate H-Arg-7-amino-4-methylcoumarin (NHMec) as described below. Separation of the molecular-size standards apoferritin (440 kDa), β-amylase (232 kDa), BSA (67 kDa), and carbonic anhydrase (29 kDa) was monitored by A_{280}.

Enzymatic Analysis. Aminopeptidase activity was determined by measuring the release of the fluorogenic leaving group, NHMec, from the fluorogenic peptide substrates H-Leu-NHMec, H-Ala-NHMec, H-Arg-NHMec, H-Met-NHMec, H-Phe-NHMec, H-Gly-NHMec, H-Val-NHMec, H-Ile-NHMec H-Glu-NHMec, H-Asp-NHMec, and H-Pro-NHMec. Reactions were carried out in 96-well microtiter plates (200 μL total volume, 30 min, 37 °C) by using a spectrofluorimeter (Bio-Tek KC4), with excitation at 370 nm and emission at 460 nm. Enzyme was first added to 50 mM Tris·HCl, pH 8.0, before the addition of 10 μM H-Leu-NHMec. Initial rates were obtained at 37 °C over a range of substrate concentrations spanning K_m (0.2–500 μM), and at fixed enzyme concentrations in 50 mM Tris·HCl, pH 8.0. Experiments to characterize the pH optimum and metal dependency of the r*Pf*AM1 were carried out by using various buffers over the pH range 3–10, and in the presence or absence of metal chelators EDTA and *o*-phenathroline as described (3).

Northern Blotting. Total RNA was extracted and Northern blotting was performed essentially as described by Kyes *et al.* (4) with the following modifications: 100-μL pellet volumes of infected red blood cells were collected from cultures at ≈5% parasitemia, lysed, and stored in TRIzol (Invitrogen). Samples were separated on a 1% TBE agarose gel containing 10 mM guanidine thiocyanate (Sigma–Aldrich), soaked in 50 mM NaOH for 30 min, and transferred onto a Hybond N+ membrane (Amersham Biosciences).

Blots were probed with a 1,500-bp PCR product amplified from a full-length *Pf*A-M1 pGem clone by using primers *Pf*A-M1IntF (TACAATGGGCTTTAGAATGTC), and *Pf*A-M1IntR (AATTCATCATCTTTTGA). This product was labeled with [α-^{32}P]dCTP by random priming by using a Decaprime II kit (Ambion). The probe was hybridized overnight at 40 °C in a hybridization buffer containing formamide (Northern Max; Ambion). The filter was washed once at low stringency and twice at high stringency (Northern Max; Ambion), then exposed overnight to Super Rx Medical X-Ray film (Fuji), and developed by using a Kodak X-Omat 3000RA processor (Kodak).

Immunoblotting. Parasite protein fractions were extracted by using 0.03% saponin (Sigma–Aldrich) and prepared as described previously (5). SDS/PAGE was performed by using 10% acrylamide gels and run on Miniprotein II rigs (BioRad). Equal loading was estimated by using the Bradford method (6), and by staining gels with Coomassie Brilliant Blue (Bio-Rad) with protein proportions visually estimated.

Protein was transferred onto Hybond C+ membranes (Amersham Biosciences), which were blocked in 5% skim-milk powder for 1 h at 37 °C or overnight at 4 °C. Anti-c-myc (Sigma–Aldrich) were used as primary antibody to label transgenic *Pf*A-M1 protein at a 1/3,000 dilution. The secondary antibody was an anti-mouse IgG (Chemicon) used at a dilution of 1/5,000. Blots were incubated with ECL Detection Reagents (Amersham Biosciences), with exposure times ranging from 5 to 10 min.

In Vitro Sensitivity of *P. falciparum* Malaria Parasites to Aminopeptidase Inhibitors. The in vitro sensitivity of *P. falciparum* (3D7) parasites to bestatin, Co4 (hPheP[CH_2]Phe), and Co2 hPheP[CH_2]Gly was determined by using [^{3}H]hypoxanthine incorporation (7). Briefly, serial dilutions of each inhibitor were prepared in culture media (0.2–200 μM) and added with [^{3}H]hypoxanthine (0.5 μCi per well) to asynchronous cultures. After a 48-h incubation, the amount of [^{3}H]hypoxanthine incorporation was measured and IC_{50} values were determined by linear interpolation of inhibition curves (8). Each assay was performed in triplicate on at least 2 separate occasions.

Crystallization, X-Ray Data Collection, Structure Determination, and Refinement. Before crystallization, the purified r*Pf*A-M1 enzyme was dialyzed against gel filtration buffer [50 mM Hepes (pH 8.5)

300 mM NaCl, 5% (vol/vol) glycerol] before size-exclusion chromatography using a Superdex S200 10/30 column and concentrated to 5 mg/mL. The crystals were grown by using the hanging drop vapor diffusion method, with 1:1 (vol/vol) ratio of protein to mother liquor (0.5 mL well volume). The crystals appeared overnight in 22% (vol/vol) polyethylene glycol 8000, 10% (vol/vol) glycerol, 0.1 M Tris (pH 8.5), and 0.2 M $MgCl_2$ and reached full size in 3 days. Crystals of the rP*f*A-M1-Co4 and rP*f*A-M1-bestatin complex were obtained by cocrystallization under similar conditions in the presence of the ligand at 1 mM. Crystals were dehydrated against reservoir buffer with 15% (vol/vol) glycerol for 16 h. Crystals were equilibrated for 5 min in reservoir buffer in the presence of 20% (vol/vol) glycerol. Cryoannealing was performed 3 times by blocking the cryostream (100 K) for 5 sec. Cryoannealing substantially improved the diffraction quality observed. Crystal quality was variable and a large number had to be screened.

Data were collected at 100 K by using both in-house and synchrotron radiation. The diffraction data for the ligand-free, Co4-bound, and bestatin-bound protease were collected to 2.1-, 2.0-, and 1.65-Å resolution, respectively. Diffraction images were processed by using MOSFLM (9) and SCALA (10) from the CCP4 suite (11); 5% of each dataset was flagged for calculation of R_{free} (12) with neither a σ nor a low-resolution cut-off applied to the data. A summary of statistics is provided in Table S2. Subsequent crystallographic and structural analysis was performed by using the CCP4i interface (13) to the CCP4 suite (11), unless stated otherwise.

Structure solution preceded by using the molecular replacement method and the program PHASER (14). A search model was constructed from the crystal structure of aminopeptidase N from *Neisseria meningitides* (PDB ID 2GTQ), the closest structural homolog identified by using the FFAS server (15). A "mixed" model, consisting of conserved side chains (all other non-alanine/glycine residues truncated at Cγ atom), was then created by using the SCRWL server (15). A single clear peak in both the rotation and translation functions was evident with 1 PfA-M1 monomer placed in the asymmetric unit. Unbiased ($F_o - F_c$) density in initial electron density maps and initial refinement cycles indicated that this solution was correct.

Maximum likelihood refinement using REFMAC (16), incorporating translation, liberation, and screw-rotation displacement (TLS) refinement was carried out, by using a bulk solvent correction (Babinet model with mask). Building was guided by manual inspection of the model and R_{free}. Simulated annealing composite omit maps were generated by using CNS (17) omitting 5% of the model. All model building and structural validation were done by using COOT (18). Water molecules were added to the model by using ARP/wARP (19) when the R_{free} reached 25%. Solvent molecules were retained only if they had acceptable hydrogen-bonding geometry contacts of 2.5–3.5 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 (www.pymol.org). Hydrogen bonds (excluding water-mediated bonds), were calculated by using the CONTACT (11). The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 3EBG.pdb; 3EBH.pdb; and 3EBI.pdb).

1. Trager W, Jensen JB (1976) Human malaria parasites in continuous culture. *Science* 193:673–675.
2. Lambros C, Vanderburg JP (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol* 65:418–420.
3. Stack CM, *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.
4. Kyes S, Pinches R, Newbold C (2000) A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol* 105:311–315.
5. Spielmann T, Gardiner DL, Beck HP, Trenholme KR, Kemp DJ (2006) Organization of ETRAMPs and EXP-1 at the parasite-host cell interface of malaria parasites. *Mol Microbiol* 59:779–794.
6. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254.
7. Geary TG, Delaney EJ, Klotz IM, Jensen JB (1983) Inhibition of the growth of *Plasmodium falciparum* in vitro by covalent modification of hemoglobin. *Mol Biochem Parasitol* 9:59–72.
8. Huber W, Koella JC (1993) A comparison of three methods of estimating EC50 in studies of drug resistance of malaria parasites. *Acta Trop* 55:257–261.
9. Leslie AGW (1992) *Joint CCP4 + ESF-EAMCB Newslett Prot Crystallogr* (Daresbury Laboratory, Warrington, UK), 26.
10. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62:72–82.
11. CCP4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
12. Brunger AT (1993) Assessment of phase accuracy by cross validation: The free R value. Methods and applications. *Acta Crystallogr D* 49:24–36.
13. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D* 59:1131–1137.
14. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D* 61:458–464.
15. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) FFAS03: A server for profile–profile sequence alignments. *Nucleic Acids Res* 33:W284–288.
16. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
17. Brunger AT, *et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54:905–921.
18. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
19. Cohen SX, *et al.* (2008) ARP/wARP and molecular replacement: The next generation. *Acta Crystallogr D* 64:49–60.
20. Davis IW, *et al.* (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–W383.

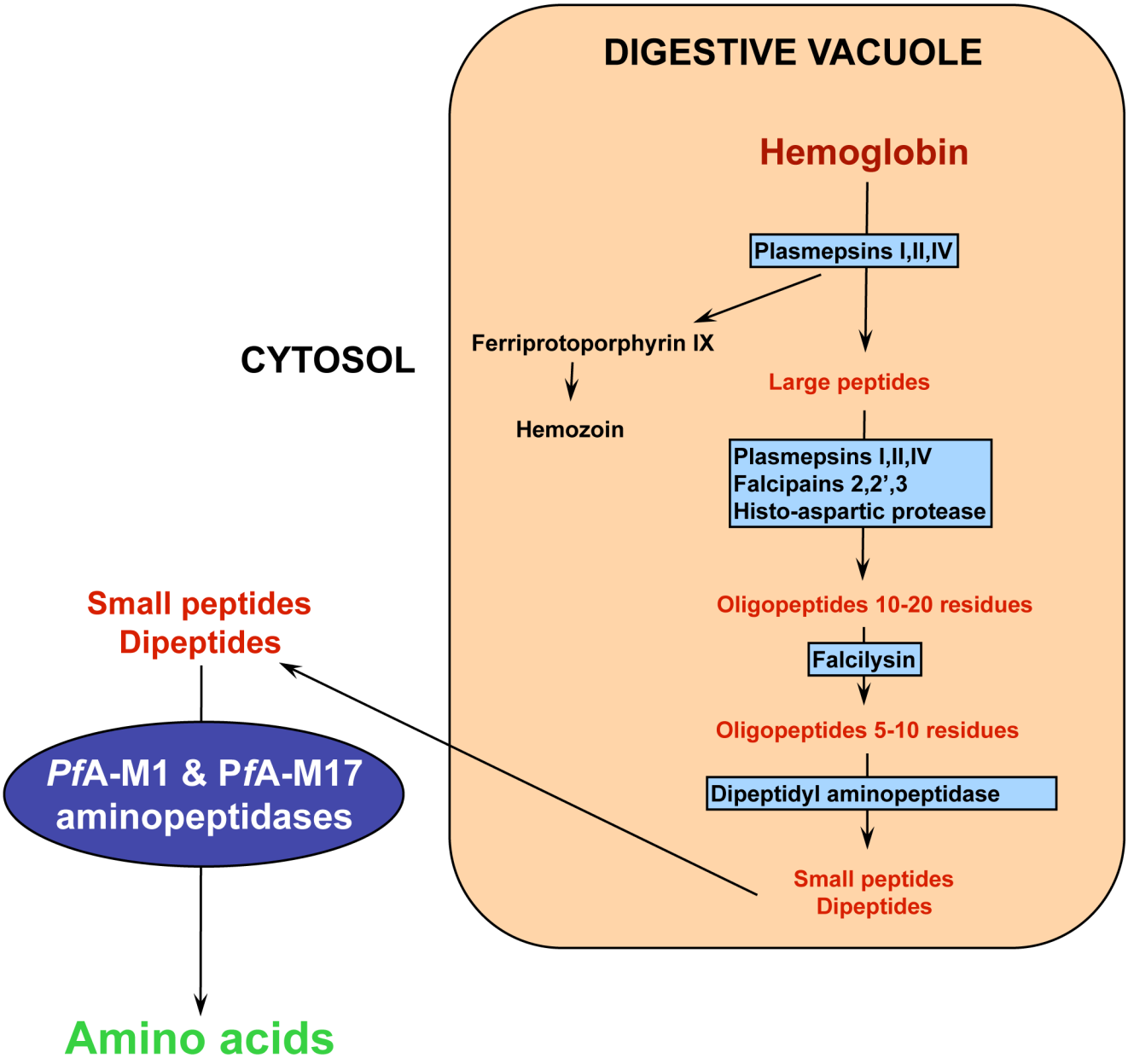

Fig. S1. Aminopeptidases *Pf*A-M1 and *Pf*A-M17 release amino acids from dipeptides that are transported from the digestive vacuole, where hemoglobin digestion is initiated by various proteases.

Fig. S2. Alignment of *Plasmodium* species M1 neutral aminopeptidases. Sequence alignment was prepared by using ClustalW and Espript. Identical residues are highlighted in red and conservatively substituted amino acids are written in red. The *Plasmodium* species are listed on the left. Putative transmembrane domain is boxed in green and asparagine-rich low complexity regions are located within black hashed boxes. The GAMEN substrate recognition motif is boxed in solid black lines and the zinc binding motif is underlined with catalytic residues indicated (asterisk). The first amino acid of the truncated r*Pf*A-M1 that was successfully purified from *E. coli* is circled.

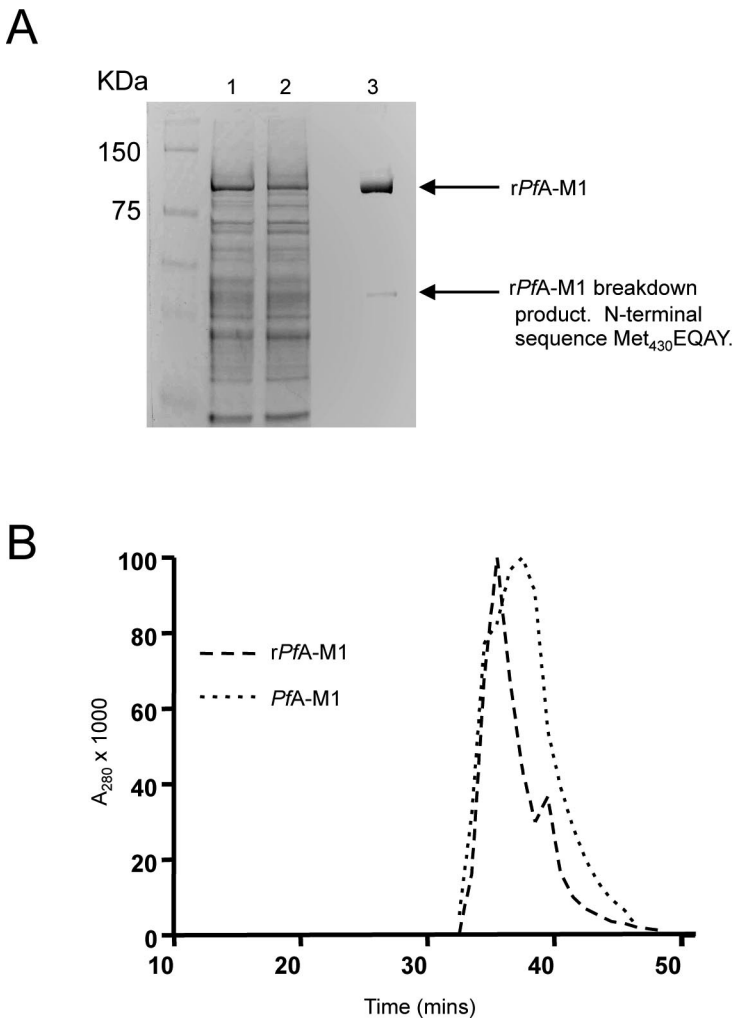

Fig. S3. (*A*) Purification of r*Pf*A-M1 by Ni-NTA-agarose affinity chromatography. Lane 1, bacterial extracts; lane 2, unbound protein; lane 3, purified r*Pf*A-M1. (*B*) Size separation of r*Pf*A-M1 (0.04 mg/mL) and parasite soluble extracts (0.05 mg/mL) on a Superdex-200 column.

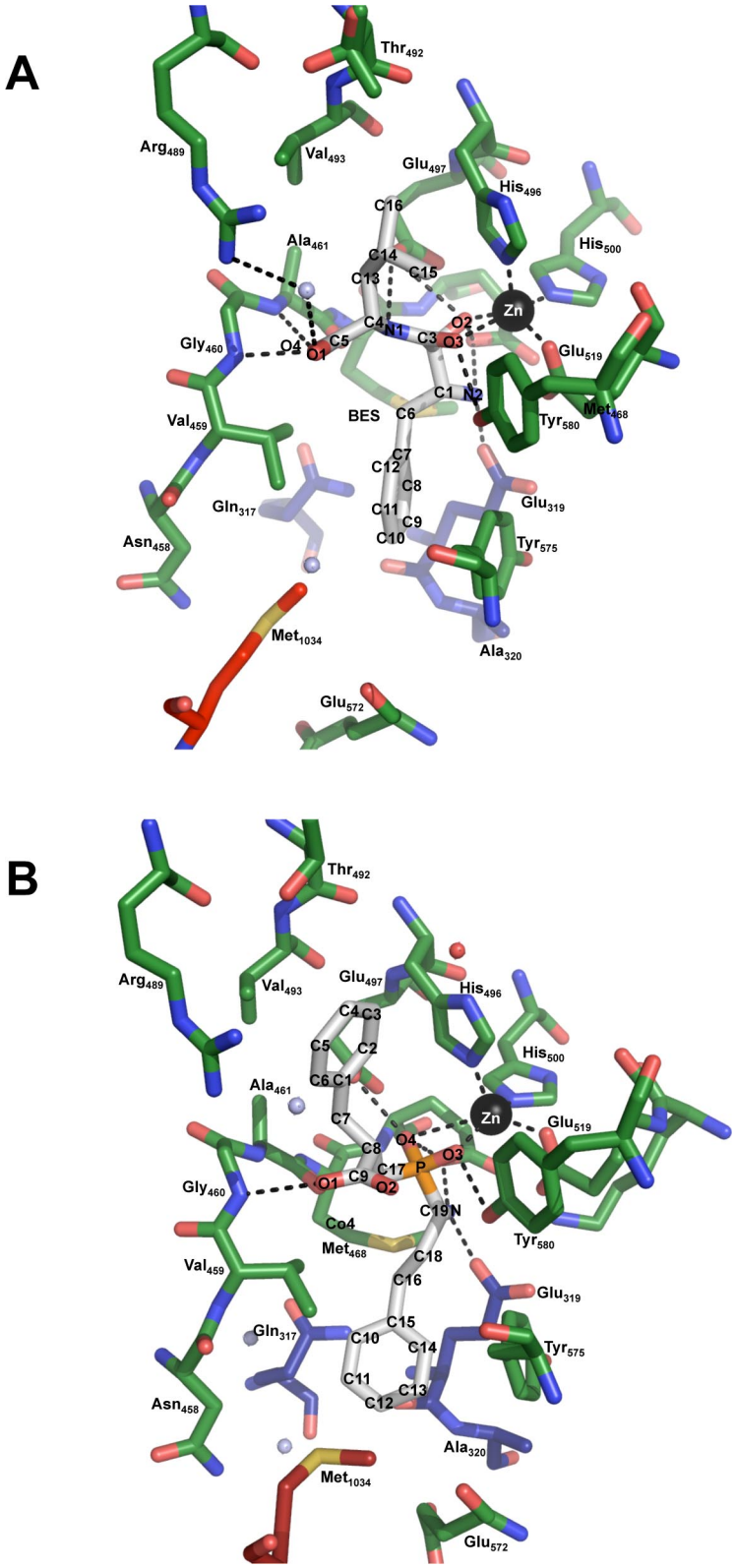

Fig. S4. Binding of inhibitors to active site of r*Pf*A-M1. (*A*) The 1.65-Å rPfA-M1-BES and (*B*) 2.0 Å rPfA-M1-Co4 active site showing inhibitors bound in the active site. Carbon atoms of inhibitors (BES/Co4) are colored in gray and atom numbers are indicated. Carbon atoms of *Pf*A-M1 residues are colored by domain: I (blue), II (green), IV (red), and Zn shown as black sphere. Water molecules are shown as blue spheres. Hydrogen bonds are indicated (dashed lines).

Table S1. Comparison of specificity constants for various N-terminal amino acids for r*Pf*A-M1 at pH 7.5

Substrate	k_{cat}, s^{-1}	K_m, μM	k_{cat}/K_m, $M^{-1}{\cdot}s^{-1}$	Abundance in human Hb, %
H-Leu-NHmec	1.52	329.9	4,607	12.46
H-Ala-NHmec	2.04	888.9	2,295	12.46
H-Arg-NHmec	1.07	717.4	1,491	2.08
H-Phe-NHmec	0.18	194.8	924	5.19
H-Gly-NHmec	0.116	348.6	333	6.92
H-Val-NHmec	0.036	1,068.1	34	10.73
H-Ile-NHmec	0.040	1,706	23	0
H-Pro-NHmec	0.0032	734.4	4	4.84

NHMec, 7-amino-4-methylcoumarin.

Table S2. Data collection and refinement statistics

	r*Pf*A-M1	r*Pf*A-M1-Co4	r*Pf*A-M1-bestatin
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions, Å	a = 75.7, b = 108.7, c = 118.4	a = 75.9, b = 108.6, c = 118.3	a = 75.7, b = 108.6, c = 118.0
Resolution, Å	34.99–2.1 (2.21–2.1)	28.61–2.0 (2.11–2.0)	46.5–1.65 (1.74–1.65)
Total no. of observations	289,352	432,263	440,194
No. of unique observations	56,863	60,523	95,094
Multiplicity	5.1 (3.8)	7.1 (5.5)	4.6 (4.1)
Data completeness, %	98.7 (94.0)	91.0 (77.5)	81.5 (68.1)
$\langle I/\sigma\, I\rangle$	16.7 (2.8)	22.0 (3.0)	18.3 (2.8)
R_{pim}, %*	4.2 (21.3)	3.0 (24.5)	2.9 (22.0)
Structure refinement			
No. of nonhydrogen atoms			
Protein	7,233	7,332	7,337
Solvent	762	739	835
Ligand	—	26	23
R_{free}, %	22.0	22.1	20.2
R_{cryst}, %	17.0	17.5	17.7
rms deviations from ideality			
Bond lengths, Å	0.010	0.009	0.006
Bond angles, °	1.13	1.33	0.97
Ramachandran plot			
Favored, %	98.0	98.2	98.4
Allowed, %	100	100	100
B factors, Å^2			
Mean main chain	21.8	17.4	15.4
Mean side chain	22.6	19.8	18.2
Mean ligand	—	23.5	18.2
Mean water molecule	30.8	33.3	29.7
rmsd bonded *B* values			
Main chain	0.48	0.77	0.57
Side chain	1.1	2.3	1.9
MolProbity score (ref. 20)	1.38 (99th percentile†)	1.45 (98th percentile†)	1.24 (98th percentile‡)

Values in parentheses refer to the highest-resolution shell.

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

†N = 12,522, 1.75 Å (ref. 20).

‡N = 9,033, 1.4–1.9 Å (ref. 20).

Table S3. The r*Pf*A-M1–ligand interactions

Metal core or protein ligand	Bestatin atom	Co4 atom	Type of bond
Metal			
Zn^{2+}		r*Pf*A-M1 His^{496} $N\varepsilon_2$	Metal–protein
Zn^{2+}		r*Pf*A-M1 His^{500} $N\varepsilon_2$	Metal–protein
Zn^{2+}		r*Pf*A-M1 Glu^{519} $O\varepsilon_1$	Metal–protein
Zn^{2+}	O3	O3	Metal–ligand
Zn^{2+}	O2	O4	Metal–ligand
Protein ligand			
Glu^{319} $O\varepsilon_1$	N2	N	H bond
Gly^{460} N	O4	O1	H bond
Ala^{461} N	O4	—	H bond
Glu^{463} $O\varepsilon$	N2/O2	N/O4	H bond
Glu^{497} $O\varepsilon$	O2/N1	O4	H bond
Tyr^{580} OH	O3	O3	H bond
Water HOH	O1	—	H bond
Gln^{317}	P1′ (Phe)	P1	van der Waals
Ala^{320}	—	N	van der Waals
Val^{459}	O4/O1	P1	van der Waals
Ala^{461}	—	C17	van der Waals
Met^{462}	P1′ (Phe)	P1	van der Waals
Thr^{492}	P1 (Leu)	P1′	van der Waals
Val^{493}	—	P1′	van der Waals
Lys^{518}	N2	C19	van der Waals
Tyr^{575}	P1′ (Phe)	C19	van der Waals
Met^{1034}	P1′ (Phe)	P1	van der Waals