3D structures of the Plasmodium vivax subtilisin-like drug target SUB1 reveal conformational changes to accommodate a substrate-derived α-ketoamide inhibitor

Mariano Martinez, Fernando A Batista, Manon Maurel, Anthony Bouillon, Laura Ortega Varga, Anne Marie Wehenkel, Lucile Le Chevalier-Sontag, Arnaud Blondel, Ahmed Haouz, Jean-François Hernandez, et al.

STRUCTURAL BIOLOGY

ISSN: 2059-7983
journals.iucr.org/d

3D structures of the *Plasmodium vivax* subtilisin-like drug target SUB1 reveal conformational changes to accommodate a substrate-derived α-ketoamide inhibitor

Mariano Martinez, Fernando A. Batista, Manon Maurel, Anthony Bouillon, Laura Ortega Varga, Anne Marie Wehenkel, Lucile Le Chevalier-Sontag, Arnaud Blondel, Ahmed Haouz, Jean-François Hernandez, Pedro M. Alzari and Jean-Christophe Barale

Acta Cryst. (2023). D**79**, 721–734

STRUCTURAL BIOLOGY
ISSN 2059-7983

3D structures of the *Plasmodium vivax* subtilisin-like drug target SUB1 reveal conformational changes to accommodate a substrate-derived α-ketoamide inhibitor

Mariano Martinez,[a] Fernando A. Batista,[a] Manon Maurel,[b]‡ Anthony Bouillon,[a] Laura Ortega Varga,[c]§ Anne Marie Wehenkel,[a] Lucile Le Chevalier-Sontag,[a] Arnaud Blondel,[c] Ahmed Haouz,[d] Jean-François Hernandez,[b] Pedro M. Alzari[a] and Jean-Christophe Barale[a]*

[a]Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Unité de Microbiologie Structurale, 75015 Paris, France, [b]Institut des Biomolécules Max Mousseron, CNRS, Université Montpellier, ENSCM, 34090 Montpellier CEDEX 5, France, [c]Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Unité de Bioinformatique Structurale, 75015 Paris, France, and [d]Institut Pasteur, Université Paris Cité, Plate-forme de Cristallographie–C2RT, Département de Biologie Structurale et Chimie et CNRS UMR 3528, 75015 Paris, France. *Correspondence e-mail: jean-christophe.barale@pasteur.fr

Received 13 March 2023
Accepted 30 May 2023

Edited by M. Rudolph, F. Hoffmann-La Roche Ltd, Switzerland

‡ Present address: SpiroChem, Rosental Area, WRO-1047-3, Mattenstrasse 22, 4058 Basel, Switzerland.
§ Present address: Alzheimer's Research UK Oxford Drug Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, United Kingdom.

Keywords: subtilisin; malaria; *Plasmodium vivax*; drug design; α-ketoamide inhibitors; 3D structures.

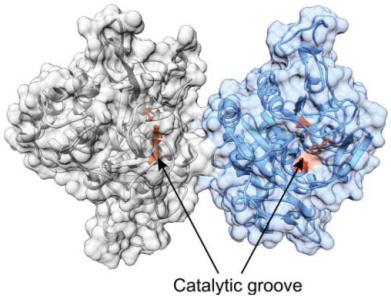
PDB references: catalytic domain of *P. vivax* SUB1, triclinic crystal form, 8coz; trigonal crystal form, 8cp0; complex with inhibitor, 8coy

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

The constant selection and propagation of multi-resistant *Plasmodium* sp. parasites require the identification of new antimalarial candidates involved in as-yet untargeted metabolic pathways. Subtilisin-like protease 1 (SUB1) belongs to a new generation of drug targets because it plays a crucial role during egress of the parasite from infected host cells at different stages of its life cycle. SUB1 is characterized by an unusual pro-region that tightly interacts with its cognate catalytic domain, thus precluding 3D structural analysis of enzyme–inhibitor complexes. In the present study, to overcome this limitation, stringent ionic conditions and controlled proteolysis of recombinant full-length *P. vivax* SUB1 were used to obtain crystals of an active and stable catalytic domain ($PvS1_{Cat}$) without a pro-region. High-resolution 3D structures of $PvS1_{Cat}$, alone and in complex with an α-ketoamide substrate-derived inhibitor (MAM-117), showed that, as expected, the catalytic serine of SUB1 formed a covalent bond with the α-keto group of the inhibitor. A network of hydrogen bonds and hydrophobic interactions stabilized the complex, including at the P1′ and P2′ positions of the inhibitor, although P′ residues are usually less important in defining the substrate specificity of subtilisins. Moreover, when associated with a substrate-derived peptidomimetic inhibitor, the catalytic groove of SUB1 underwent significant structural changes, particularly in its S4 pocket. These findings pave the way for future strategies for the design of optimized SUB1-specific inhibitors that may define a novel class of antimalarial candidates.

1. Introduction

Malaria, which is caused by *Plasmodium* parasites transmitted by mosquitoes, remains a global health issue despite long-term efforts. The World Health Organization reported that the number of clinical cases has not decreased since 2015, and even increased in 2020 compared with 2019 (World Health Organization, 2021). It is not known whether this is a consequence of the SARS-CoV2 pandemic. A first antimalaria vaccine with moderate efficiency has been recommended by the World Health Organization for young children (World Health Organization, 2021), and other candidates are under development (Datoo *et al.*, 2022). Nonetheless, new antimalarial candidates need to be developed due to the continuous selection of multi-resistant *Plasmodium* parasites, including the recent emergence of artemisinin-resistant



parasites in Africa (Balikagala *et al.*, 2021; Uwimana *et al.*, 2020). Phenotypic screening of chemical libraries against *P. falciparum* blood-stage parasites cultured *in vitro* (Gamo *et al.*, 2010) and target-based approaches (recently reviewed in Guerra & Winzeler, 2022) have been used to identify and develop different drug candidates, particularly under the supervision of the Medicines for Malaria Venture (https://www.mmv.org/). Among the many new drug targets identified, *Plasmodium* subtilisin-like serine protease 1 (SUB1) has emerged as a promising candidate because it meets several of the expected prerequisites: (i) the sequence and organization of the SUB1 active site are highly related to bacterial subtilisins and are significantly distant from human subtilisins (Barale *et al.*, 1999; Giganti *et al.*, 2014), suggesting that highly specific SUB1 inhibitors could be produced without unwanted off-target effects, (ii) the *sub1* gene is required for the *Plasmodium* life cycle (Yeoh *et al.*, 2007; Thomas *et al.*, 2018), thus excluding biological redundancy, and (iii) SUB1 is a key player in the finely regulated cascade of proteases that govern egress of the parasite from the host red blood cells (Thomas *et al.*, 2018; Yeoh *et al.*, 2007; Pino *et al.*, 2017; Nasamu *et al.*, 2017; Collins *et al.*, 2013; Tan *et al.*, 2021; Mukherjee *et al.*, 2022; Dvorin & Goldberg, 2022), a key step in the asexual intra-erythrocytic cycle of *Plasmodium* sp. These repeated cycles of parasite growth and egress cause the symptoms of malaria. Moreover, SUB1 plays an equivalent role during the pre-erythrocytic asymptomatic phase in the liver when *Plasmodium* sp. merozoites egress from infected hepatocytes (Tawk *et al.*, 2013) and in intra-erythrocytic male gametocytes, allowing parasite transmission to anopheles vectors (Pace *et al.*, 2019). Thus, SUB1 inhibitors could have therapeutic and also prophylactic properties by reducing *Plasmodium* transmission, as expected for new generations of antimalarial candidates (Leroy *et al.*, 2014). However, two major bottlenecks have impaired significant progress in SUB1 inhibitor design. Initial screening for the identification of small-molecule inhibitors led to limited success (Yeoh *et al.*, 2007; Gemma *et al.*, 2012; Arastu-Kapur *et al.*, 2008; Bouillon *et al.*, 2013). Moreover, apart from 3D homology models (Bouillon *et al.*, 2013; Withers-Martinez *et al.*, 2012), no experimental structural information is currently available for the catalytic domain of SUB1 in complex with inhibitor compounds. X-ray structures are only available for full-length *P. falciparum* SUB1 (PfS1; Withers-Martinez *et al.*, 2014) and for *P. vivax* SUB1 (PvS1; Giganti *et al.*, 2014), and in both the active site is blocked by the prodomain region. To overcome the former issue, rational efforts have been made to synthesize inhibitors derived from SUB1 native substrates but involving an α-ketoamide moiety at the position of the P1–P1′ scissile peptide bond to generate a transient covalent hemiketal link between the ketone and the hydroxyl group of the catalytic serine of SUB1 (recently reviewed in Lidumniece *et al.*, 2022). A similar approach has led to the successful development of efficient drugs, for instance in the case of the hepatitis NS3/4A protease, where peptidomimetic inhibitors combined with other drugs are changing the course of the hepatitis C virus epidemic (de Leuw & Stephan, 2017; Manns & Maasoumy, 2022).

Here, we addressed the second bottleneck by describing high-resolution 3D structures of enzyme–inhibitor complexes. By combining stringent ionic conditions and controlled proteolysis of recombinant full-length *P. vivax* SUB1 ($PvS1_{FL}$), we obtained a fully active PvS1 catalytic domain devoid of the prodomain ($PvS1_{Cat}$) and analysed the 3D structures of $PvS1_{Cat}$ alone and in complex with MAM-117, an α-ketoamide substrate-derived SUB1 inhibitor. These data will pave the way for the structure-driven synthesis of optimized SUB1 inhibitors.

2. Materials and methods

2.1. Production and purification of recombinant PvS1 enzymes

Baculovirus-expressed recombinant $PvS1_{FL\text{-}bac}$ (FL-bac, full-length baculovirus; GenBank JX49148) was prepared as described previously (Bouillon *et al.*, 2013; Giganti *et al.*, 2014). To express PvS1 in *Drosophila* S2 cells, a codon-optimized synthetic gene (GenScript) encoding residues 26–630 of *P. vivax* SUB1 followed by 18 base pairs coding for six histidine residues was inserted into the expression vector pMT/BiP/V5-His (Invitrogen) using the BglII and EcoRI restriction sites downstream of and in phase with the insect BiP signal sequence for extracellular secretion of recombinant proteins. To minimize the risks of unwanted *N*-glycosylation, *NetNGlyc* (Gupta & Brunak, 2002) was used to identify putative modified asparagine residues. Thus, besides wild-type $PvS1_{FL}$ (GenBank MZ344793), two mutants were expressed: $PvS1_{FL\text{-}ng}$ (FL-ng, full length no glycosylation) and $PvS1_{FL\text{-}ng2}$, which contain three (N356S, N427S and N440S) and two (N356S and N440S) substitutions, respectively (GenBank MZ344794 and MZ344795, respectively). *Drosophila* S2 cells were transfected with the recombinant plasmids and stably transfected cells were selected in Insect-XPRESS medium (Lonza) containing 8 µg ml^{-1} puromycin.

For large-scale protein production, cells were grown at 28°C in Insect-XPRESS medium (Lonza) containing 6 µg ml^{-1} puromycin. PvS1 expression was induced at a cell density of 10^7 cells ml^{-1} by adding 5 µ*M* $CdCl_2$. The culture supernatant was harvested 160–170 h after induction, centrifuged at 2150*g* for 30 min and concentrated/diafiltrated against loading buffer (D-PBS; 500 m*M* NaCl, 5 m*M* imidazole) using an ÄKTA-crossflow system (GE Healthcare) supplemented with a 10 kDa, 0.11 m^2 Kvick Lab cassette (GE Healthcare).

Proteins were purified on an ÄKTA purifier system (GE Healthcare) at 4°C on 5 ml TALON metal-affinity resin (Clontech). After extensive washes, the bound proteins were eluted with a linear gradient of 5–200 m*M* imidazole in D-PBS, 500 m*M* NaCl. Fractions containing the expected recombinant PvS1 proteins were pooled and concentrated using Amicon Ultra 15 (10 000 molecular-weight cutoff) and size-fractionated on a HiLoad 16/60 Superdex 75 column equilibrated with 20 m*M* Tris, 500 m*M* NaCl, 1 m*M* $CaCl_2$ pH 8. The fractions were monitored by absorbance (280 nm) and analysed by Coomassie Blue staining of SDS polyacrylamide gels and with

a SUB1-specific activity assay, as described in Bouillon *et al.* (2013). Purified recombinant $PvS1_{FL}$ proteins were stored at −20°C in 10%(*v*/*v*) pure glycerol. Protein concentrations were determined from the A_{280} using the extinction coefficient predicted by ExPASy *ProtParam* (Gasteiger *et al.*, 2005).

Purified recombinant proteins were analysed by Western blotting, as described previously, using horseradish peroxidase-conjugated anti-His tag (Novagen) or anti-SUB1 antibodies (Bouillon *et al.*, 2013).

2.2. Purification of the PvS1 catalytic domain and prodomain

Purified $PvS1_{FL\text{-}ng}$ was loaded onto 2 ml Ni–NTA resin previously equilibrated with binding buffer (50 m*M* Tris, 500 m*M* NaCl pH 8) and diluted tenfold with dissociation buffer (50 m*M* Tris, 2 *M* $MgCl_2$ pH 8). The resin slurry was incubated at 4°C with gentle agitation for 90 min to let the complex dissociate. The flowthrough fraction containing $PvS1_{FL\text{-}ng}$ prodomain (PR-$PvS1_{ng}$) was collected for further purification (see below). Resin containing the catalytic domain of $PvS1_{FL\text{-}ng}$ was re-equilibrated in binding buffer to wash off $MgCl_2$ and the protein was then eluted with 14 ml elution buffer (50 m*M* Tris, 500 m*M* NaCl, 500 m*M* imidazole pH 8). $PvS1_{FL\text{-}ng}$ catalytic domain ($PvS1_{Cat}$) was concentrated using Amicon Ultra 15 (10 000 molecular-weight cutoff) and size-fractionated on a HiLoad 16/60 Superdex 75 column equilibrated with 20 m*M* Tris, 50 m*M* $CaCl_2$ pH 8. Fractions were monitored by absorbance (280 nm) and analysed by Coomassie Blue staining of SDS polyacrylamide gels and a PvS1-specific activity assay. The fractions containing the recombinant enzyme activity were pooled and stored at −20°C with 10%(*v*/*v*) glycerol for activity assays or with 2%(*v*/*v*) PEG 8000 for structural studies.

The flowthrough IMAC fraction containing PR-$PvS1_{ng}$ was dialyzed at 4°C against 20 m*M* Tris, 100 m*M* NaCl pH 8 overnight, concentrated using Amicon Ultra 15 (10 000 molecular-weight cutoff) and size-fractionated onto a HiLoad 16/60 Superdex 75 column equilibrated with 20 m*M* Tris, 50 m*M* $CaCl_2$ pH 8. Fractions were monitored by absorbance (280 nm) and analysed by Coomassie Blue staining of SDS polyacrylamide gels. The fractions containing PR-$PvS1_{ng}$ were pooled, flash-frozen in liquid nitrogen and stored at −80°C until use.

2.3. Optimization of limited proteolysis of $PvS1_{Cat}$ and N-terminal sequencing

Purified $PvS1_{Cat}$ was diluted to 20.91 µ*M* in 50 m*M* HEPES, 500 m*M* NaCl pH 7.5 and supplemented with trypsin, chymotrypsin, subtilisin A (protease:substrate molar ratio of 1:450) or papain (protease:substrate molar ratio of 1:45) (Sigma–Aldrich). After incubation at 25°C, the reaction was stopped by boiling the samples in SDS loading buffer before SDS–PAGE separation and Coomassie Blue staining. After transfer onto polyvinylidene difluoride membranes (Bio-Rad), N-terminal Edman sequencing of the proteins was performed by the Functional Genomics Center proteomics facility, ETH, Zurich, Switzerland.

Preparative purification of trypsin-treated $PvS1_{Cat}$ ($PvS1_{Cat\text{-}Tryps}$) was performed as above, but trypsin digestion at 25°C was extended to 2.5 h. $PvS1_{Cat\text{-}Tryps}$ was concentrated using Amicon Ultra 15 (10 000 molecular-weight cutoff) and size-fractionated on a HiLoad 16/60 Superdex 75 column equilibrated with 20 m*M* Tris, 50 m*M* $CaCl_2$ pH 8. Fractions were monitored by absorbance (280 nm) and analysed by Coomassie Blue staining of SDS polyacrylamide gels and a PvS1-specific enzymatic assay. The fractions containing $PvS1_{Cat\text{-}Tryps}$ were pooled and stored at −20°C with 10%(*v*/*v*) glycerol for activity assays or with 2%(*v*/*v*) PEG 8000 for crystallization studies.

2.4. PvS1-specific enzymatic assay

PvS1-specific enzymatic assays were performed as described previously (Bouillon *et al.*, 2013; Giganti *et al.*, 2014). Briefly, the PvS1 AM-DE Fluorescence Resonance Energy Transfer (FRET) substrate was a decapeptide corresponding to the native automaturation site of PvS1 (KLVGAD/DVSLA) coupled to the Dabsyl (quencher) and EDANS (fluorophore) dyes (excitation and emission at 360 and 500 nm, respectively; Intavis, Germany) at its extremities. The apparent K_m values of the recombinant enzymes for the PvS1 AM-DE FRET substrate were determined as described previously (Bouillon *et al.*, 2013) and the IC_{50} values of MAM-117 and PR-$PvS1_{ng}$ were obtained by testing 9–11 serial dilutions from 20 000 to 3 n*M* for MAM-117 and from 800 to 86 n*M* for PR-$PvS1_{ng}$. Fluorescence was monitored at 37°C every 2 min for 1 h using a Tecan Infinite M1000 reader and the data were collected and analysed using the *Magellan* (Tecan) and *Prism*5 (GraphPad) software tools. Results are expressed as the mean of at least four independent experiments.

2.5. Differential scanning fluorescence (ThermoFluor) assay

Briefly, 2.5 µg $PvS1_{Cat\text{-}Tryps}$ in 20 m*M* Tris pH 8, 150 m*M* NaCl supplemented with 1 m*M* MAM-117 or 5% DMSO was dispensed into 96-well PCR plates (25 µl per well). 2.5 µl of 50× SYPRO Orange solution (Invitrogen) was added to each well, which was then heated from 25 to 95°C (1°C step each 1 min) in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Excitation and emission filters of 492 and 516 nm, respectively, were used to monitor the fluorescence increase resulting from SYPRO Orange binding to exposed hydrophobic regions of the unfolding proteins. The melting temperature (T_m) was defined as the midpoint of the protein-unfolding transition.

2.6. Protein crystallization and crystallographic studies

The initial screening for crystallization conditions was carried out using the vapour-diffusion method and a Mosquito nanolitre-dispensing system (TTP Labtech, Melbourn, United Kingdom) following established protocols (Weber *et al.*, 2019). Optimization of the initial hits was performed manually in Linbro plates with a hanging-drop setup. Optimal crystals of $PvS1_{Cat}$ were obtained by mixing 1.5 µl recombinant protein solution (10.8 mg ml^{-1}) with 1.5 µl reservoir solution

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	$PvS1_{Cat}$	$PvS1_{Cat\text{-}Tryps}$	$PvS1_{Cat\text{-}Tryps}$–MAM-117
Data collection			
Beamline	ID29, ESRF	ID29, ESRF	ID30, ESRF
Wavelength (Å)	0.9763	0.9763	0.9686
Space group	$P3_221$	$P1$	$P1$
a, b, c (Å)	89.15, 89.15, 119.08	53.50, 55.50, 69.04	53.26, 55.10, 69.01
α, β, γ (°)	90, 90, 120	69.2, 78.9, 75.9	69.2, 78.3, 74.8
Resolution (Å)	47.1–3.25 (3.51–3.25)	42.7–1.44 (1.44–1.46)	64.05–1.51 (1.53–1.51)
$R_{p.i.m.}$	0.107 (0.264)	0.047 (0.794)	0.031 (0.337)
$\langle I/\sigma(I)\rangle$	6.9 (2.8)	8.6 (1.5)	12.0 (2.1)
Completeness (%)	98.0 (98.7)	91.2 (79.1)	94.1 (92.8)
Multiplicity	5.6 (5.5)	3.1 (2.8)	1.9 (1.9)
$CC_{1/2}$	0.98 (0.92)	0.99 (0.39)	1.0 (0.78)
Total No. of observations	49091 (9875)	365057 (14374)	202066 (9818)
No. of unique reflections	8758 (1784)	118611 (5123)	104444 (5196)
Refinement			
Resolution (Å)	3.25	1.44	1.51
No. of reflections	8756	118601	104436
R_{work}/R_{free}	0.182/0.233	0.195/0.216	0.176/0.192
No. of atoms			
Protein	2649	5183	5244
Heterogen	3	69	64
Water	—	752	490
Root-mean-square deviations			
Bond lengths (Å)	0.009	0.011	0.012
Bond angles (°)	1.00	1.08	1.11
Ramachandran outliers (%)	0.3	1.2	1.2

consisting of 0.1 *M* MES, 0.2 *M* $MgCl_2$, 25%(*w*/*v*) PEG 4000 pH 6.5 at 18°C. Rod-like crystals (0.25 × 0.05 × 0.05 mm) appeared within 1–2 weeks. Before diffraction data collection, single crystals were flash-cooled in liquid nitrogen using a mixture of 50% Paratone and 50% paraffin oil as a cryoprotectant. Optimal crystals of $PvS1_{Cat\text{-}Tryps}$ were obtained by mixing 1.5 µl recombinant protein solution (8 mg ml^{-1}) with 1.5 µl reservoir solution consisting of 0.5 *M* $LiSO_4$, 15%(*w*/*v*) PEG 8000 at 18°C. Plate-like crystals appeared within 1–2 weeks. Before diffraction data collection, single crystals were flash-cooled in liquid nitrogen using mother liquor containing 25% ethylene glycol as a cryoprotectant.

For the crystallization of $PvS1_{Cat\text{-}Tryps}$ in complex with the inhibitor MAM-117, 1 µl MAM-117 solution (5 m*M* in 50% DMSO) was transferred into a PCR tube and incubated at 37°C for 24–48 h to evaporate the DMSO. 3.5 µl $PvS1_{Cat\text{-}Tryps}$ (14 mg ml^{-1}) was then mixed with dried MAM-117 (final protein:inhibitor ratio of 1:4.4) and incubated at 4°C for 4 h before setting up the crystallization drops. Optimal crystals of $PvS1_{Cat\text{-}Tryps}$ in complex with MAM-117 were obtained at 18°C by mixing 1.5 µl protein complex solution with 1.5 µl reservoir solution consisting of 0.5 *M* $LiSO_4$, 15%(*w*/*v*) PEG 8000 and seeded with microcrystals (strike seeding) of $PvS1_{Cat\text{-}Tryps}$ grown in the absence of the inhibitor in the same crystallization conditions. Plate-like crystals appeared within 1–2 weeks. Before diffraction data collection, single crystals were flash-cooled in liquid nitrogen using mother liquor containing 1.25 m*M* MAM-117 and 25% ethylene glycol as a cryoprotectant.

Diffraction data sets were collected from a single flash-cooled crystal at 100 K using beamlines ID29 and ID30 at the ESRF Synchrotron Radiation Facility, Grenoble, France. Data were processed with *XDS* (Kabsch, 2010) and scaled with *AIMLESS* from the *CCP*4 suite (Agirre *et al.*, 2023). Crystal structures were solved by molecular-replacement methods using *Phaser* (McCoy *et al.*, 2007) using the catalytic domain of $PvS1_{FL\text{-}bac}$ (PDB entry 4tr2; residues Tyr277–Lys611) as a model probe. Structure refinement was performed with *BUSTER* (Bricogne *et al.*, 2009) using a TLS model and noncrystallographic symmetry restraints when appropriate, alternating with manual rebuilding with *Coot* (Emsley *et al.*, 2010). Statistics for the collected data and the final refinement are given in Table 1. Atomic coordinates and structure factors were deposited in the Protein Data Bank with the following PDB codes: 8cp0 ($PvS1_{Cat}$), 8coz ($PvS1_{Cat\text{-}Tryps}$) and 8coy ($PvS1_{Cat\text{-}Tryps}$–MAM-117).

Subpocket volume calculations were performed using *mkgridXf* (Monet *et al.*, 2019) with default parameters, except for the outer radius, which was set to 10 Å, and suppression of the corresponding single amino acid of the bound inhibitor to free the corresponding volume. Orientation and r.m.s. calculations were performed with *VMD* (Humphrey *et al.*, 1996), using the *MultiSeq* module (Roberts *et al.*, 2006) for alignment of the whole catalytic domains and the internal *RMSD_Calculator* for the actual calculation and alignment of the catalytic pocket of PvS1 defined as Tyr371, His372, Leu405, Lys409, Leu410, Gly411, Arg412, Leu413, Met416, Ser434, Phe435, Ser436, Phe437, Ser461, Ser463, Asn464, Leu545, Asn546, Gly547, Thr548, Ser549 and Met550.

The structural figures in this manuscript were prepared using *UCSF Chimera* (Pettersen *et al.*, 2004) and *LigPlot+* (Laskowski & Swindells, 2011) for protein–ligand interactions.

The normalized *B* factors used for mobility comparison were obtained with the *BAN*Δ*IT* server (Barthels *et al.*, 2021).

2.7. Synthesis of MAM-117, an α-ketoamide inhibitor of PvS1

Reagents, resin and solvents for peptide synthesis [fluorenylmethyloxycarbonyl (Fmoc)-protected amino-acid derivatives, 1-{[(dimethylamino)methylene]-*N*-methylmethanaminium}-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium-3-oxide hexafluorophosphate (HATU), diisopropylethylamine (DIEA), Dess–Martin periodinane (DMP), triisopropylsilane (TIS), Fmoc-protected Rink amide polystyrene resin (100–200 mesh, 0.5–0.6 mmol g^{-1}), dimethylformamide (DMF), dichloromethane (DCM), methanol (MeOH), trifluoroacetic acid (TFA), acetonitrile *etc.*] were purchased from IrisBiotech, Novabiochem, Acros Organics, Alfa Aesar, Carlo Erba or Sigma–Aldrich and were used without further purification. HPLC-grade solvents were used for HPLC and LC/MS analyses. The Fmoc-protected norstatine analogue of alanine [(3*S*)-3-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino}-2-hydroxybutanoic acid], which was used as the precursor of the α-ketoacyl residue, was synthesized from Fmoc-Ala-OH as described previously (Buisson *et al.*, 2006).

MAM-117 [Ac-Ile-Thr-Ala-Ala-CO-Asp-Glu(Oall)-NH_2] was prepared in the solid phase following an Fmoc/tBu strategy and using an Fmoc-protected Rink amide resin (0.15 mmol) placed in a plastic syringe equipped with frit. On the resin, stepwise peptide-chain assembly was accomplished manually by a succession of Fmoc deprotection and coupling steps from the C-terminus to the N-terminus. The Fmoc group was eliminated by treatment with piperidine/DMF (20/80) for 10 and 25 min. Fmoc-protected amino acids (0.45 mmol) were coupled in DMF in the presence of HATU (171 mg, 0.45 mmol) and DIEA (105 µl, 0.6 mmol) for 3 h. After each step, excess reagents were eliminated by filtration and the resin was washed with DMF, MeOH and DCM. N-terminal acetylation was performed using acetic acid (26 µl, 0.45 mmol) and the same coupling conditions. After peptide assembly, the secondary alcohol function of the norstatine analogue of alanine was oxidized by two 4 h treatments with DMP (191 mg, 0.45 mmol) in DCM to yield the α-ketoamide group.

The peptide was simultaneously cleaved from the resin and deprotected by 2 h treatment with a mixture of TFA, TIS and water (95:2.5:2.5, 10 ml g^{-1}). After resin removal by filtration, TFA was concentrated *in vacuo*. Diethyl ether was then added to the oily residue, leading to precipitation of the compound. After centrifugation and decantation, the obtained pellet was washed twice with diethyl ether.

Crude MAM-117 was purified by preparative RP-HPLC on a Waters Delta-Pak C18 column (40 × 100 mm, 15 µm, 100 Å) by elution with a linear gradient of eluent *B* in eluent *A* at a rate of 1% min^{-1} (a flow rate of 28 ml min^{-1}). Eluents *A* and *B* were 0.1% solutions of TFA in water and acetonitrile, respectively.

The collected fractions and the purified compound were analysed using the following methods. (i) RP-HPLC using a Waters Alliance 2690 HPLC equipped with a Chromolith SpeedRod C18 column (0.46 × 5 cm), applying a linear gradient (0–100%) of eluent *B* in eluent *A* over 5 min (a flow rate of 3 ml min^{-1}). Detection was made at a wavelength of 214 nm. Eluents *A* and *B* were as above. (ii) LC-MS using a Waters Alliance 2690 HPLC coupled to a Waters Micromass ZQ spectrometer (electrospray ionization mode, ESI+) equipped with a C18 monolithic Onyx Phenomenex column (0.46 × 2.5 cm). Separation was performed with a linear gradient (0–100%) of eluent *B* in eluent *A* over 3 min at a flow rate of 3 ml min^{-1}. Eluents *A* and *B* were 0.1% solutions of formic acid in water and acetonitrile, respectively. Positive ion electrospray mass spectra were acquired at a solvent flow rate of 100–500 µl min^{-1}. Ten scans were summed to obtain the final spectrum.

Characterization of the synthesized MAM-117 ($C_{31}H_{49}N_7O_{13}$, calculated exact mass 727.34): white freeze-dried powder, 24 mg (22% yield), *m*/*z* (ES+) 728.2 (M+H^+). t_R (HPLC), 1.06 min; t_R (LC/MS), 1.04 and 1.11 min [the compound is in equilibrium between the ketone (1.11 min) and its hydrated form (1.04 min)].

3. Results and discussion

3.1. Expression of recombinant active PvS1 enzymes in *Drosophila* S2 cells

Previous studies demonstrated that full-length SUB1 from various *Plasmodium* species can be expressed as soluble and active recombinant enzymes using the baculovirus expression system (Withers-Martinez *et al.*, 2002; Bouillon *et al.*, 2013). However, these secreted recombinant SUB1 enzymes are purified as a tight complex that contains the classical subtilase catalytic domain associated with its cognate pro-region. Analysis of the 3D structure of full-length *P. vivax* SUB1 ($PvS1_{FL\text{-}bac}$; PDB entry 4tr2) showed that together with the classical subtilisin pro-region, SUB1 harbours an additional N-terminal domain (Fig. 1*b*) that stabilizes SUB1 as an inactive complex and that requires a significant shift in calcium ion concentration to be autocatalytically processed to reveal SUB1 activity (Giganti *et al.*, 2014). Within this complex, the active site of SUB1 remains occupied by its primary automaturation site, precluding the crystallization of enzyme–inhibitor complexes.

To circumvent the limited yield obtained when expressing full-length PvS1 in the baculovirus expression system (Table 2), the nonlytic *Drosophila* S2 cell expression system represents an interesting alternative to produce different eukaryotic recombinant proteins for functional and structural studies (Moraes *et al.*, 2012). *Drosophila* S2 cells can be genetically modified to express functional proteins independently of viral infections. Moreover, as the stably transfected cell lines grow in suspension, protein production can easily be scaled up using Erlenmeyer flasks. Full-length PvS1 expressed in S2 cells ($PvS1_{FL}$) showed the same pattern as $PvS1_{FL\text{-}bac}$ (Fig. 1*a*). The catalytic domain appeared as a doublet of 45–50 kDa with a separated pro-region of approximately 25 kDa, indicating that the primary automaturation of full-length PvS1

occurred in a similar way in both expression systems (Withers-Martinez *et al.*, 2002; Bouillon *et al.*, 2013; Giganti *et al.*, 2014). However, the production yield of wild-type $PvS1_{FL}$ remained low and the protein was virtually inactive *in vitro* (Table 2), possibly due to nonspecific *N*-glycosylation. Therefore, we designed a mutant in which three potential *N*-glycosylation sites (Asn361, Asn432 and Asn445) were mutated to serines as a first attempt to improve the production yield and facilitate crystallization (Kozak *et al.*, 2020; Chang *et al.*, 2007). The production of this triple mutant ($PvS1_{FL\text{-}ng}$; Fig. 1*b*) was increased (up to 30 mg of recombinant protein per litre of culture; Table 2). Moreover, $PvS1_{FL\text{-}ng}$ showed a lower apparent molecular weight compared with $PvS1_{FL}$ (Fig. 1*a*), which is consistent with the loss of one or more *N*-glycosylation sites, and displayed a significant enzymatic activity (Table 2). Different constructs of SUB1 could be expressed in

Table 2
Kinetic properties of PvS1 variants.

ND, not determined.

PvS1 variant	Expression system	Mutations	V_{max} (AFU min^{-1})	K_m^{app} (μ*M*)	R^2	Yield (mg per litre of culture)
$PvS1_{FL\text{-}bac}$	Baculovirus	None	391.2 ± 2.7	22.1 ± 2.2	0.91	1–2
$PvS1_{FL}$	*Drosophila* S2 cells	None	ND	ND	—	<1
$PvS1_{FL\text{-}ng}$	*Drosophila* S2 cells	N445S, N361S, N432S	263.3 ± 78	65.8 ± 20†	0.74	30
$PvS1_{Cat}$	*Drosophila* S2 cells	N445S, N361S, N432S	536.2 ± 90.1	66.1 ± 20.1†	0.86	9 (30% starting from $PvS1_{FL\text{-}ng}$)
$PvS1_{Cat\text{-}Tryps}$	*Drosophila* S2 cells	N445S, N361S, N432S	3557.1 ± 77.5	23.36 ± 2	0.99	3.6 (12% starting from $PvS1_{Cat}$)

† Mean values of >4 independent experiments that presented >30% variation across repetitions, as illustrated by the elevated standard error of the means (SEM) and the poor goodness-of-fit (R^2) of linear regressions provided by *GraphPad Prism* (see Section 2).

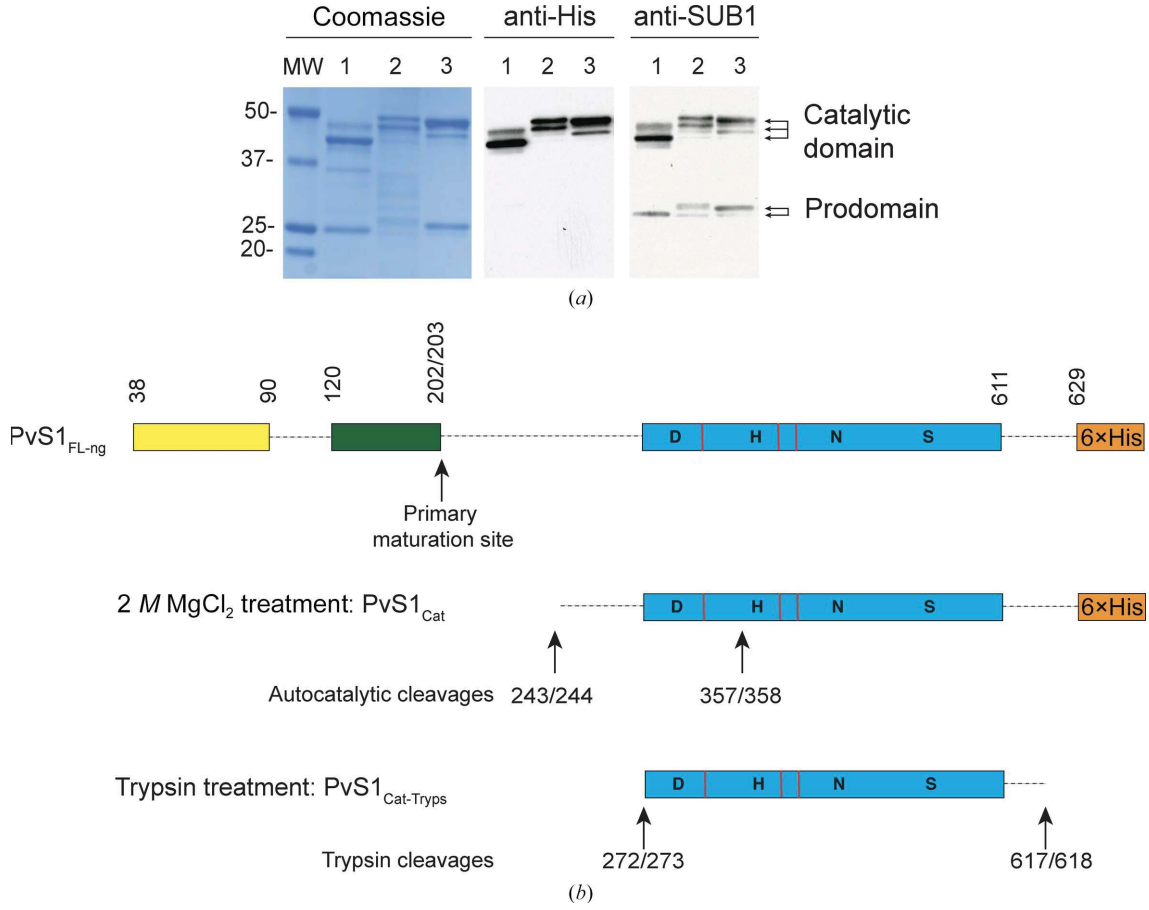


Figure 1
(*a*) SDS–PAGE and Western blot analysis of baculovirus-expressed $PvS1_{FL\text{-}bac}$ (lane 1) and *Drosophila* S2 cell-expressed $PvS1_{FL}$ (lane 2) and $PvS1_{FL\text{-}ng}$ (lane 3). Left panel: Coomassie Blue staining. Middle panel: Western blotting with anti-His-tag antibodies. Right panel: Western blotting with anti-PvS1 antibodies (Bouillon *et al.*, 2013). Protein molecular weights (MW) are indicated in kDa. (*b*) Scheme of full-length PvS1 composed of the prodomain (belt domain in yellow and classical subtilase prodomain in green) and catalytic domain (blue) with the boundary residues indicated and the four catalytic residues (Asp316, His376, Asn464 and Ser549) that define PvS1 as a subtilase. Dotted lines represent unstructured segments that are not visible in the crystal structure of the full-length protein (Giganti *et al.*, 2014). Red bars indicate the locations of the three potential *N*-glycosylation sites (Asn361, Asn432 and Asn445) that are mutated to serines in $PvS1_{FL\text{-}ng}$. Lower panels, the two constructs crystallized in this work; the different proteolytic cleavage sites discussed in the text are indicated by vertical arrows. Residue numbering refers to the native *Plasmodium vivax* subtilisin-like 1 serine protease (GenBank accession No. FJ536584).

S2 cells or other expression systems, including full-length *P. falciparum* S1 and a truncated form of PvS1 lacking the belt domain that stabilizes the complex formed by the pro-region and the catalytic domain (Fig. 1*b*; Giganti *et al.*, 2014). However, due to a limited expression yield, solubility issues or an inability to obtain crystals of PvS1 without the belt domain, we concentrated on $PvS1_{FL\text{-}ng}$ in all subsequent experiments.

3.2. Production of prodomain-free catalytic domain of PvS1

To separate the prodomain of SUB1 from the catalytic domain, we produced $PvS1_{FL\text{-}ng}$ and carried out preliminary experiments to identify stringent native conditions for destabilizing the strong noncovalent protein–protein interactions. For this, we purified $PvS1_{FL\text{-}ng}$ on Ni–NTA resin through its C-terminal 6×His tag and incubated it with various concentrations of NaCl, KCl, $MgCl_2$ and $CaCl_2$ to identify suitable ionic strength conditions to dissociate and wash out the pro-region of PvS1. We then eluted wash-resistant proteins from the resin with imidazole and analysed them by SDS–PAGE (Fig. 2*a*). Among the tested conditions, only $MgCl_2$ could destabilize the $PvS1_{FL\text{-}ng}$ complex in a concentration-dependent manner, leading to a significant reduction in the amount of PvS1 pro-region and the appearance of a putative low-molecular-weight catalytic domain of approximately 35 kDa (Fig. 2*a*). Large-scale treatment of $PvS1_{FL\text{-}ng}$ using 2 *M* $MgCl_2$ followed by a final purification step based on size-exclusion chromatography allowed us to efficiently separate the intact form of the PvS1 prodomain (25 kDa, PR-$PvS1_{ng}$; Fig. 2*b*, lane 2) from the catalytic domain ($PvS1_{Cat}$), which presented two bands at approximately 35 and 45 kDa (Fig. 2*b*, lane 3). N-terminal sequencing of these bands indicated that $PvS1_{FL\text{-}ng}$ was processed like $PvS1_{FL\text{-}bac}$ at the primary maturation site (Asp202/Asp203) and that upon separation from its cognate prodomain the catalytic domain underwent at least two additional autoproteolytic cleavages (Fig. 1*b* and Fig. 2*b*, lane 3). The first site (Ala243/Ser244) was located within the N-terminal unstructured region of the catalytic domain and the second site (Ala357/Asn358) was within a protein loop that defines the SUB1-specific divalent calcium-binding site

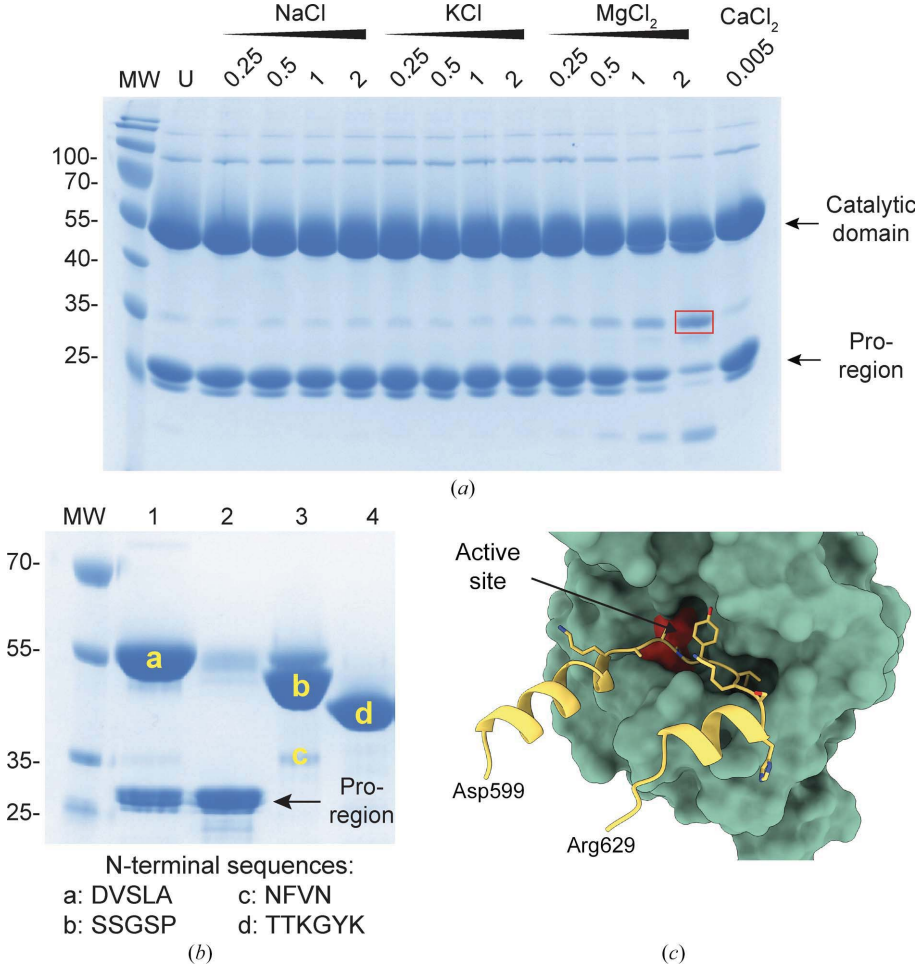


Figure 2
(*a*) Screening of conditions for disassembly of $PvS1_{FL\text{-}ng}$ complexed with its pro-region. SDS–PAGE and Coomassie Blue staining of eluted fractions collected after washing immobilized $PvS1_{FL\text{-}ng}$ with different solutions. The pro-region and catalytic domain of PvS1 are indicated, and the red rectangle indicates $PvS1_{Cat}$ product degradation that corresponds to band c in (*b*). U, untreated $PvS1_{FL\text{-}ng}$; MW, protein molecular weights in kDa. (*b*) SDS–PAGE and Coomassie Blue staining of the purification of $PvS1_{FL\text{-}ng}$ (lane 1), PR-$PvS1_{ng}$ (lane 2), $PvS1_{Cat}$ (lane 3) and $PvS1_{Cat\text{-}Tryps}$ (lane 4). N-terminal sequencing by Edman degradation of the bands indicated on the gel is shown. Protein molecular weights are indicated in kDa. (*c*) Detail of the active site of $PvS1_{Cat}$ (in red) occupied by the C-terminal extension of a neighbouring molecule (yellow).

(residues Glu336–Asp345) (Giganti *et al.*, 2014). Screening for stabilizing conditions, including divalent cation salts, temperature and buffers at different pH values, showed that $CaCl_2$ (25 m*M*) decreased the Ala357/Asn358 cleavage, thus stabilizing $PvS1_{Cat}$. Therefore, we included 25 m*M* $CaCl_2$ in all chromatographic steps. Importantly, removal of the cognate prodomain resulted in a sharp increase in the specific activity of $PvS1_{Cat}$ compared with both $PvS1_{FL\text{-}ng}$ and $PvS1_{FL\text{-}bac}$ (Table 2). Altogether, these results validated the strategy of using *Drosophila* S2 cells for PvS1 expression in order to obtain high yields of pure SUB1 active catalytic domain for kinetic and structural characterization studies.

3.3. Crystal structure of the catalytic domain of PvS1

We could obtain an initial crystal form from freshly purified $PvS1_{Cat}$. These crystals contained one molecule in the asymmetric unit and diffracted to moderate/low (3.25 Å) resolution. We determined the 3D structure using molecular-replacement methods with the catalytic domain of $PvS1_{FL\text{-}bac}$ (residues 277–611; PDB entry 4tr2) as a model probe (Table 1). As expected, the long connecting segment located on the N-terminal side of the catalytic domain (residues 203–276) was not visible in the electron density because this region is disordered in the crystal. The structure of $PvS1_{Cat}$ was very similar to that of the catalytic domain of the full-length auto-inhibited protein ($PvS1_{FL\text{-}bac}$; r.m.s.d. of 0.281 Å on C^{α} atoms; Supplementary Fig. S1), confirming that the three Asn-to-Ser substitutions did not alter the overall structure of the catalytic domain of PvS1 and that the fully active $PvS1_{Cat}$ did not undergo major changes upon removal of the prodomain. The major difference was that the C-terminal extension of the protein, which is not part of the catalytic core (residues 612–629) and is disordered in PDB entry 4tr2, adopted a well defined structure (Supplementary Fig. S1) and bound in a reverse sense (C-terminal to N-terminal residues in the S to S′ enzyme pockets), compared with a substrate, to the active site of a neighbouring molecule in the $PvS1_{Cat}$ crystal (Fig. 2*c*). As this crystal form was not amenable for co-crystallization with inhibitors, we next attempted to proteolytically remove the C-terminus of $PvS1_{Cat}$ to obtain a crystal form with an accessible catalytic site.

We incubated freshly purified $PvS1_{Cat}$ with limited concentrations of trypsin, subtilisin A, chymotrypsin and papain (Fig. 3*a*). With the exception of chymotrypsin, all of these enzymes were able to digest $PvS1_{Cat}$, producing a protease-resistant catalytic core of about 40 kDa. We selected trypsin for further experiments because trypsin digestion led to the production of a homogeneous catalytic domain (hereafter called $PvS1_{Cat\text{-}Tryps}$) after 150 min of incubation. N-terminal sequencing of $PvS1_{Cat\text{-}Tryps}$ revealed that trypsin cleaved $PvS1_{Cat}$ between Lys272 and Thr273 (Fig. 2*b*), removing the disordered N-terminal connecting segment (Fig. 1*b*). Moreover, Western blot analysis using an anti-His antibody confirmed that trypsin also removed the C-terminal extension of $PvS1_{Cat}$ (Fig. 3*b*).

Enzymatic activity assays revealed that $PvS1_{Cat\text{-}Tryps}$ was significantly more active than $PvS1_{Cat}$ and full-length PvS1 (Table 2). This strongly suggested that the N-terminal connecting segment and the C-terminal extension of PvS1 are dispensable and minimize its catalytic activity. Consequently, we could precisely and reproducibly measure the apparent V_{max} and K_m of $PvS1_{Cat\text{-}Tryps}$, unlike those of $PvS1_{FL\text{-}ng}$ and $PvS1_{Cat}$, which varied greatly (Table 2). Importantly, the apparent K_m of 23.36 ± 2 µ*M* of $PvS1_{Cat\text{-}Tryps}$ for the PvS1 Dabsyl-EDANS substrate was similar to the K_m values obtained for $PvS1_{FL\text{-}bac}$ and $PvS1_{FL\text{-}ng}$, suggesting that the $MgCl_2$ and trypsin treatments did not impair the enzymatic properties of $PvS1_{Cat\text{-}Tryps}$ (Table 2). Thus, we scaled up the trypsin-based proteolysis and further purified $PvS1_{Cat\text{-}Tryps}$ by gel filtration. This revealed a shift in retention time compared with $PvS1_{FL\text{-}ng}$ and $PvS1_{Cat}$ (Fig. 3*c*), in agreement with a lower molecular weight form that was also observed by SDS–PAGE (Fig. 2*b*, lane 4).

$PvS1_{Cat\text{-}Tryps}$ crystallized in the triclinic space group *P*1 with two molecules in the asymmetric unit. We determined the crystal structure at 1.8 Å resolution (Table 1). Close inspection revealed that the active sites of both molecules in the asymmetric unit were solvent-accessible (Fig. 3*d*), suggesting that this crystal form can be exploited for the structural characterization of protein–inhibitor complexes. Analysis of the $PvS1_{Cat\text{-}Tryps}$ structure showed that trypsin cleaved the C-terminus of PvS1 at Lys617 because the small peptide extension (residues 612–617), engaged in crystal contacts, was well defined in the electron-density map (Supplementary Fig. S2). The entire catalytic domain, from Lys278 to Lys617, was well defined in the electron-density map, with the only exception of a solvent-protruding loop (residues 468–473) that was disordered in both molecules and was missing from the model. The high-resolution model had four Ramachandran outliers (1.2%), but all residues were well defined in density and three of them (Asp316, Tyr371 and Ser463) were part of the substrate-binding site, where they interacted with the bound inhibitor (see below). The overall structure of triclinic $PvS1_{Cat\text{-}Tryps}$ was very similar to those of the trigonal $PvS1_{Cat}$ and the $PvS1_{FL\text{-}bac}$ catalytic domain (PDB entry 4tr2; residues Tyr277–Lys611), with overall r.m.s.d. values of 0.237 and 0.286 Å for the equivalent residues Lys278–Lys617 and Asn281–Lys617, respectively. In particular, superposition of the residues defining the catalytic groove between $PvS1_{Cat\text{-}Tryps}$ and $PvS1_{FL\text{-}bac}$ (r.m.s.d. of 0.308 Å) only revealed minor differences: Asp316, His372, Leu405, Leu413, Phe437 and Ser461 displayed a similar conformation (Fig. 3*e*), while the side chains of Ser436, Asn464, Ser549 showed a different orientation due to interactions with the cognate prodomain in the full-length protein (Giganti *et al.*, 2014).

3.4. Crystal structure of PvS1 in complex with the MAM-117 inhibitor

For most subtilases, enzyme specificity is primarily dictated by interactions of the P1–P4 substrate-residue side chains that tightly bind to the corresponding active-site subpockets S1–S4,

termed according to the Schechter and Berger nomenclature, also defining the scissile peptide bond between the P1 and P1′ positions (Siezen & Leunissen, 1997; Schechter & Berger, 1967).

In the framework of our search for specific inhibitors based on the sequence of native SUB1 substrates, we initially synthetized the peptidomimetic α-ketoamide inhibitor JMV5126 (derived from SERA4, a member of the SERA proteins that belong to the natural SUB1 substrates; Yeoh *et al.*, 2007). JMV5126 inhibits both PvS1 and PfS1 (Giganti *et al.*, 2014). Compared with JMV5126, from which it is derived (Fig. 4*a*), MAM-117 is shorter, corresponding to positions P4–P2′ of the SERA4 cleavage site, rather than the P5–P4′ pseudopeptide JMV5126, and contains an Ala-α-ketoacyl and a Glu(Oall) in positions P1 and P2′, respectively, compared with a Gln-α-ketoacyl and an Asp in JMV5126. MAM-117 was 100 times more potent than JMV5126 in inhibiting the activity of $PvS1_{FL\text{-}bac}$ (Supplementary Fig. S3*c*, Table 3) and $PvS1_{Cat\text{-}Tryps}$ (IC_{50} values of 165 ± 27 and 226 ± 33 n*M*, respectively). Notably, MAM-117 inhibited PvS1 more potently than its pro-region PR-$PvS1_{ng}$ (IC_{50} values of 326 ± 19 and 430 ±40 n*M* for $PvS1_{FL\text{-}bac}$ and $PvS1_{Cat\text{-}Tryps}$, respectively; Table 3 and Supplementary Figs. S3*a* and S3*b*). In addition, the strong interaction of MAM-117 with the active site of $PvS1_{Cat\text{-}Tryps}$ was illustrated by a ΔT_m of 13°C as measured by a ThermoFluor assay (Supplementary Fig. S3*c*).

Table 3
Kinetic inhibition.

	IC_{50} (n*M*)	
	PR-$PvS1_{ng}$	MAM-117
$PvS1_{FL\text{-}bac}$	325.9 ± 19.4	164.8 ± 27.3
$PvS1_{Cat\text{-}Tryps}$	430.5 ± 40.3	225.5 ± 32.8

To elucidate the structural basis of this inhibition, we determined the crystal structure of $PvS1_{Cat\text{-}Tryps}$ in complex with MAM-117 (Fig. 4*a*) at 1.5 Å resolution (Table 1). The crystal structure of the complex showed clear electron density for the $PvS1_{Cat\text{-}Tryps}$ active site of both protein molecules present in the asymmetric unit and for both bound MAM-117 molecules. Their similar structure (the r.m.s.d. values for MAM-117 in chains *A* and *B* were 0.1825 and 0.1023 Å for all atoms and for the C^{α} atoms of P4–P1′, respectively) was easily modelled (Figs. 4*b* and 4*c* and Supplementary Fig. S4*a*). *PISA* analysis (Krissinel & Henrick, 2007) revealed that 72% of the solvent-accessible area of MAM-117 was involved in interaction with the catalytic groove of $PvS1_{Cat\text{-}Tryps}$, forming an extensive network of hydrophobic interactions with Thr548, Leu545, Met550, Met416, Phe435, Leu,405, Leu413, Phe437 and Gly547 (Figs. 4*c* and 4*d*). MAM-117 was further stabilized by forming hydrogen bonds to Tyr371, Lys409, His372, Ser549, Ser434, Gly411, Ser436, Asn464 and Asn546.

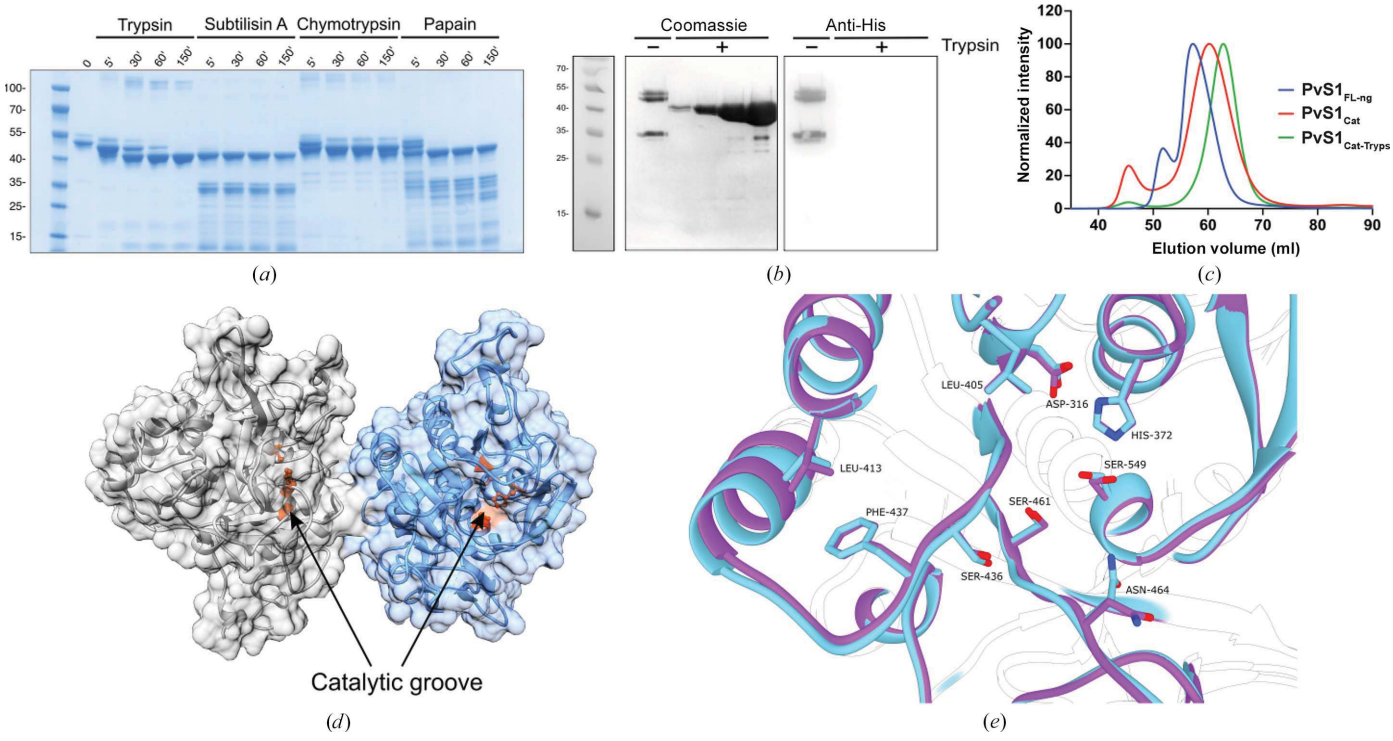


Figure 3
(*a*) Limited proteolysis of $PvS1_{Cat}$ with trypsin, subtilisin A, chymotrypsin and papain. Samples were taken at different time points and analysed by SDS–PAGE followed by Coomassie Blue staining. (*b*) Coomassie Blue staining and Western blotting using anti-His antibodies of purified $PvS1_{Cat\text{-}Tryps}$. Protein molecular weights are indicated in kDa. (*c*) Superposition of the size-exclusion chromatography profiles of $PvS1_{FL\text{-}ng}$, $PvS1_{Cat}$ and $PvS1_{Cat\text{-}Tryps}$ on a Superdex 75 16/60 column. (*d*) Crystal structure of $PvS1_{Cat\text{-}Tryps}$ showing the two polypeptide chains (blue and grey) that are present in the asymmetric unit. Active-site residues (Asp316, His372 and Ser549) are shown in red and solvent-exposed empty catalytic grooves are indicated by arrows. (*e*) Structural alignment of $PvS1_{FL\text{-}bac}$ (PDB entry 4tr2, residues Tyr277–Lys611, cyan) and $PvS1_{Cat\text{-}Tryps}$ (magenta) catalytic grooves.

Considering the interaction with the ligands, comparison of the $PvS1_{Cat\text{-}Tryps}$–MAM-117 and $PvS1_{FL\text{-}bac}$ structures (Giganti *et al.*, 2014) showed a slight backbone displacement in the S1 pocket of $PvS1_{Cat\text{-}Tryps}$ to accommodate the α-keto moiety of MAM-117, while the MAM-117 P4–P1 residues superposed well with the PvS1 automaturation peptide Val199-Gly200-Ala201-Asp202 (Supplementary Fig. S4*b*). The main difference was that P4 of MAM-117 nested more deeply in the S4 pocket of $PvS1_{Cat\text{-}Tryps}$ compared with Val199 (P4) within the S4 pocket of $PvS1_{FL\text{-}bac}$. This was shown by an increased r.m.s.d. value of the two P4 C^{α} atoms of 1.13 Å compared with those of the P3, P2 and P1 residues: 0.6, 0.65 and 0.58 Å, respectively. A comparative *LigPlot* analysis of the complex crystal structures indicated that most of the residues involved in the interaction with both ligands were shared. Specifically, the main hydrogen bonds were conserved along the P4–P1

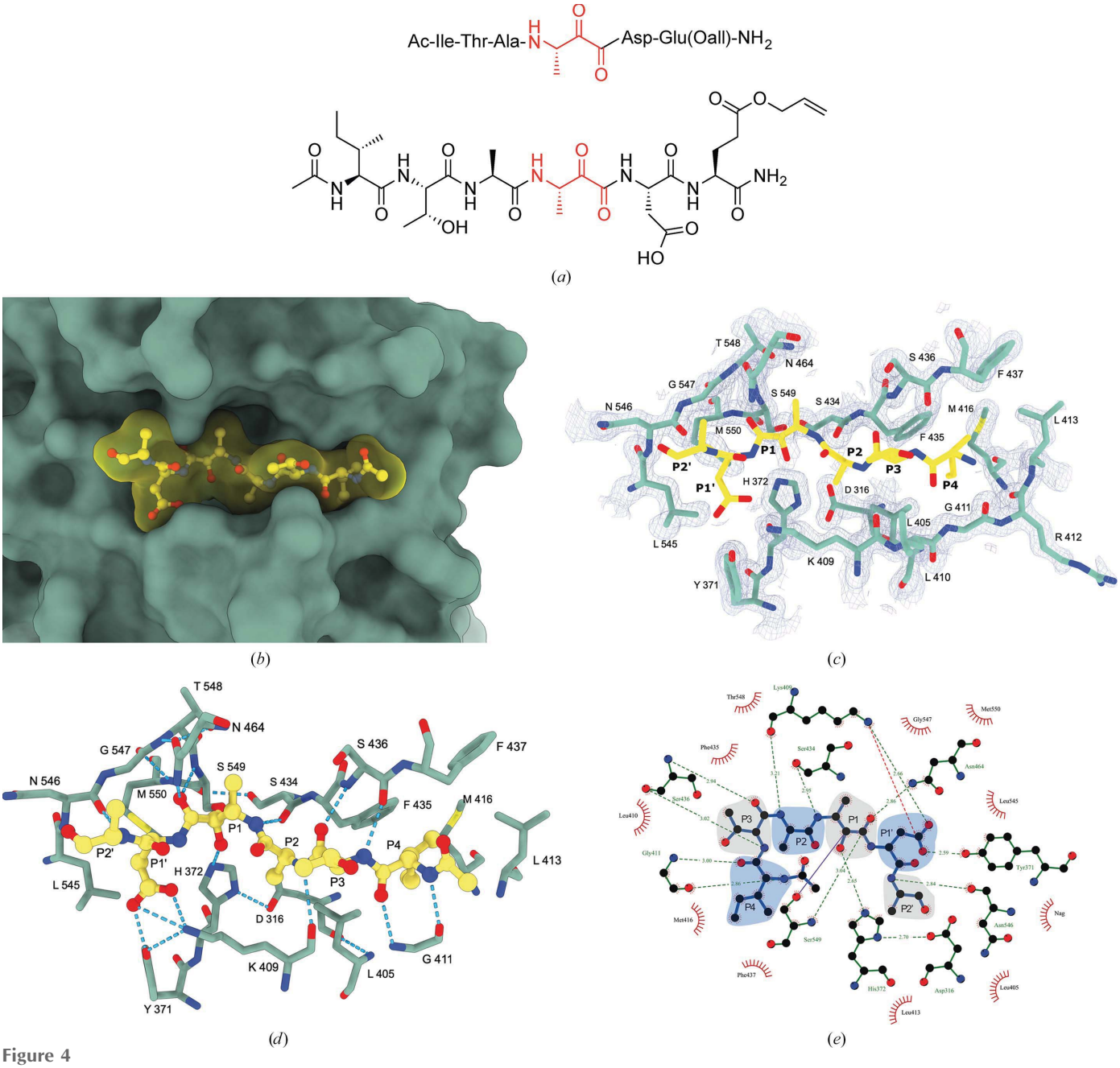


Figure 4

(*a*) Sequence and chemical structure of the MAM-117 inhibitor. The P1 residue that contains the α-ketoamide moiety is in red. Ac and all correspond to acetyl and allyl, respectively. (*b*) Crystal structure of $PvS1_{Cat\text{-}Tryps}$ (depicted as a green surface) in complex with the inhibitor MAM-117 (yellow). (*c*) Electron density of chain *A* of the receptor ($PvS1_{Cat\text{-}Tryps}$ catalytic groove) coupled to its ligand (MAM-117 inhibitor) within the asymmetric unit of $PvS1_{Cat\text{-}Tryps}$–MAM-117. (*d*) $PvS1_{Cat\text{-}Tryps}$ residues (in green) involved in the interaction with MAM-117 (in yellow). Hydrogen bonds are depicted as blue dotted lines. The P4–P2′ positions of MAM-117 are indicated. (*e*) *LigPlot*+ analysis of the interactions of the MAM-117 inhibitor (blue lines) with the catalytic active site of $PvS1_{Cat\text{-}Tryps}$ (green residues) (Laskowski & Swindells, 2011). The covalent bond between the catalytic Ser549 residue of $PvS1_{Cat\text{-}Tryps}$ and the α-ketoamide moiety of MAM-117 is shown as a purple line. Hydrogen bonds are depicted as green dotted lines and hydrophobic interactions as red semicircles. The P4 to P2′ positions of MAM-117 are shown in grey or light blue.

portion, while the tetrahedral intermediate was also stabilized by the oxyanion hole (Asn464; Supplementary Fig. S4*c*). However, we also observed significant structural differences between the $PvS1_{Cat\text{-}Tryps}$–MAM-117 complex and $PvS1_{FL\text{-}bac}$ and $PvS1_{Cat\text{-}Tryps}$, particularly in the binding mode, due to the different chemical composition of the ligands. Close inspection of the active site of $PvS1_{Cat\text{-}Tryps}$ showed that the catalytic Ser549 residue was engaged, forming, as expected, a covalent bond to the C atom of the α-keto group of MAM-117 and trapping SUB1 in a transition state-like configuration (Fig. 4*d*). Conversely, in $PvS1_{FL\text{-}bac}$ the P1 residue Asp202 of the automaturation site of PvS1 made hydrogen bonds to Ser436 and Ser461 (Supplementary Fig. S4*c*). In addition, the $PvS1_{Cat\text{-}Tryps}$–MAM-117 complex was stabilized by more interactions than between $PvS1_{FL\text{-}bac}$ and the automaturation site through the recruitment of new residues to form hydrogen bonds and hydrophobic interactions. This was illustrated by the closer distance between Leu413 and the bulkier Ile residue in the P4 position of MAM-117 (distance of 3.69 Å; Fig. 4*d*) compared with the P4 Val residue present in the SUB1 automaturation peptide (4.16 Å; Supplementary Fig. S4*c*). Interestingly, several MAM-117-specific contacts were located at the other extremity of the PvS1 binding groove within the S1′–S2′ region defined by Gly547, Tyr371, Met550, Asn546 and Leu545 (Fig. 4*d*). MAM-117 P1′ was engaged with the side chains of Tyr371 and Lys409 of PvS1 via a hydrogen bond and a salt bridge, respectively. This illustrated the previously observed contribution of the SUB1 S1′ pocket to interaction with substrate-derived inhibitors (Kher *et al.*, 2014). Concerning P2′, Kher and coworkers reported compound 12 (Ac-Ile-Thr-Ala-Ala*-CO-βAla-OH, where Ala* is the racemic residue), a close MAM-117 analogue (Kher *et al.*, 2014) that only lacks a P2′ residue and the α-carbonyl of P1′. Compared with MAM-117, compound 12 is 15-fold less potent (IC_{50} value of 2.5 ± 0.1 μM) towards $PfS1_{FL\text{-}bac}$, the *P. falciparum* ortholog of $PvS1_{FL\text{-}bac}$ with a virtually identical catalytic groove (Withers-Martinez *et al.*, 2012; Bouillon *et al.*, 2013; Giganti *et al.*, 2014). Therefore, the addition of Glu(Oall) in position P2′ of MAM-117, the NH group of which interacts with the backbone carbonyl of Asn546 of PvS1 via a hydrogen bond (Fig. 4*e* and Supplementary Fig. S4*c*), led to a 15-fold gain in affinity. Interestingly, the P2′ Glu(Oall) extremity of MAM-117 was not clearly defined in the crystal structure and showed significant flexibility due to the absence of any stabilizing interaction with SUB1. Therefore, the optimization of future SUB1 inhibitors might involve new interactions with the neighbouring residues of P2′, such as Leu545. Altogether, these observations highlight the importance of P′ positions for the development of potent SUB1-specific inhibitors, whereas the substrate specificity of eukaryotic subtilisins is mainly triggered by their P4–P1 residues, or P6–P1 residues for some pro-hormone convertases (Siezen & Leunissen, 1997; Seidah & Chrétien, 1999; Seidah & Prat, 2012).

Lastly, the most significant differences in the catalytic groove of PvS1 revealed by solution of the $PvS1_{Cat\text{-}Tryps}$–MAM-117 structure were a series of conformational changes that involved the side chains of the oxyanion hole (Asn464) and also the S4 pocket of PvS1. Unlike in $PvS1_{Cat\text{-}Tryps}$, where the whole lateral chain of Asn464 in the PvS1 oxyanion hole pointed outwards, when interacting with its automaturation substrate or with MAM-117 Asn464 was placed in a catalytically competent state, thus stabilizing the intermediate of the reaction by forming a hydrogen bond to the O atom of the β-carbonyl group of the P1 residue of MAM-117 (Figs. 5*a* and 5*b*). In addition, compared with $PvS1_{Cat\text{-}Tryps}$, the Phe435 and Met416 lateral chains moved apart within the S4 pocket of $PvS1_{Cat\text{-}Tryps}$–MAM-117 to accommodate the bulky isoleucine side chain of MAM-117 P4. Conversely, in $PvS1_{FL\text{-}bac}$ the P4 valine only changed the conformation of Phe435 (Fig. 5*c*). Therefore, the volume of the S4 pocket of $PvS1_{Cat\text{-}Tryps}$–MAM-117 (52.1 $Å^3$) doubled compared with its equivalent in $PvS1_{FL\text{-}bac}$ (25 $Å^3$; Fig. 5*d*). This significant increase is likely to be the consequence of the larger size of the Ile side chain in MAM-117. Moreover, this residue was more deeply buried in the S4 pocket of $PvS1_{Cat}$ compared with the P4 Val residue of the automaturation site of $PvS1_{FL\text{-}bac}$. These observations illustrate how the plasticity of the SUB1 S4 pocket could accommodate the large P4 side chains present in some native SUB1 substrates, such as the isoleucine found in SERA4 and SERA5 (Yeoh *et al.*, 2007; Bastianelli *et al.*, 2014) and also the bulky cyclopentyl moiety of the cyclopentylglycine present in P4 of PfSUB1 inhibitors (Lidumniece *et al.*, 2021).

4. Conclusion

The egress of *Plasmodium* sp. merozoites from host cells is tightly regulated and sequentially involves the aspartate protease plasmepsin X (PmX). PmX participates in activating the subtilisin-like protease SUB1, which then activates the SERA cysteine and serine proteases involved in rupture of the parasitophorous and hepatocyte or erythrocyte cell membranes, thus allowing blood-cell invasion by newborn merozoites (recently reviewed by Dvorin & Goldberg, 2022). Inhibitors of PmX and SUB1 impair merozoite egress *in vivo* and *in vitro*, respectively (Pino *et al.*, 2017; Nasamu *et al.*, 2017; Yeoh *et al.*, 2007; Gemma *et al.*, 2012; Arastu-Kapur *et al.*, 2008; Bastianelli *et al.*, 2011; Bouillon *et al.*, 2013), thus defining these enzymes as promising new drug targets. Importantly, targeting different crucial enzymes within the same pathway could reduce the risk of resistance because this would require several independent genetic events within the same parasite.

In this work, we have described the first high-resolution 3D structures of $PvS1_{Cat}$, a pro-region-free catalytic form of *P. vivax* SUB1, alone and complexed with MAM-117, a native substrate-derived pseudopeptide inhibitor. Besides the expected covalent bond formed between the α-ketoamide moiety of MAM-117 and the catalytic Ser549 of SUB1, analysis of the 3D structure of $PvS1_{Cat}$–MAM-117 demonstrated that the catalytic groove of SUB1 undergoes subtle conformational changes induced by the substrate-derived inhibitor. These changes could not have been anticipated from the structures of full-length PvS1 (Giganti *et al.*, 2014) and PfS1 (Withers-Martinez *et al.*, 2014) and from the structure of

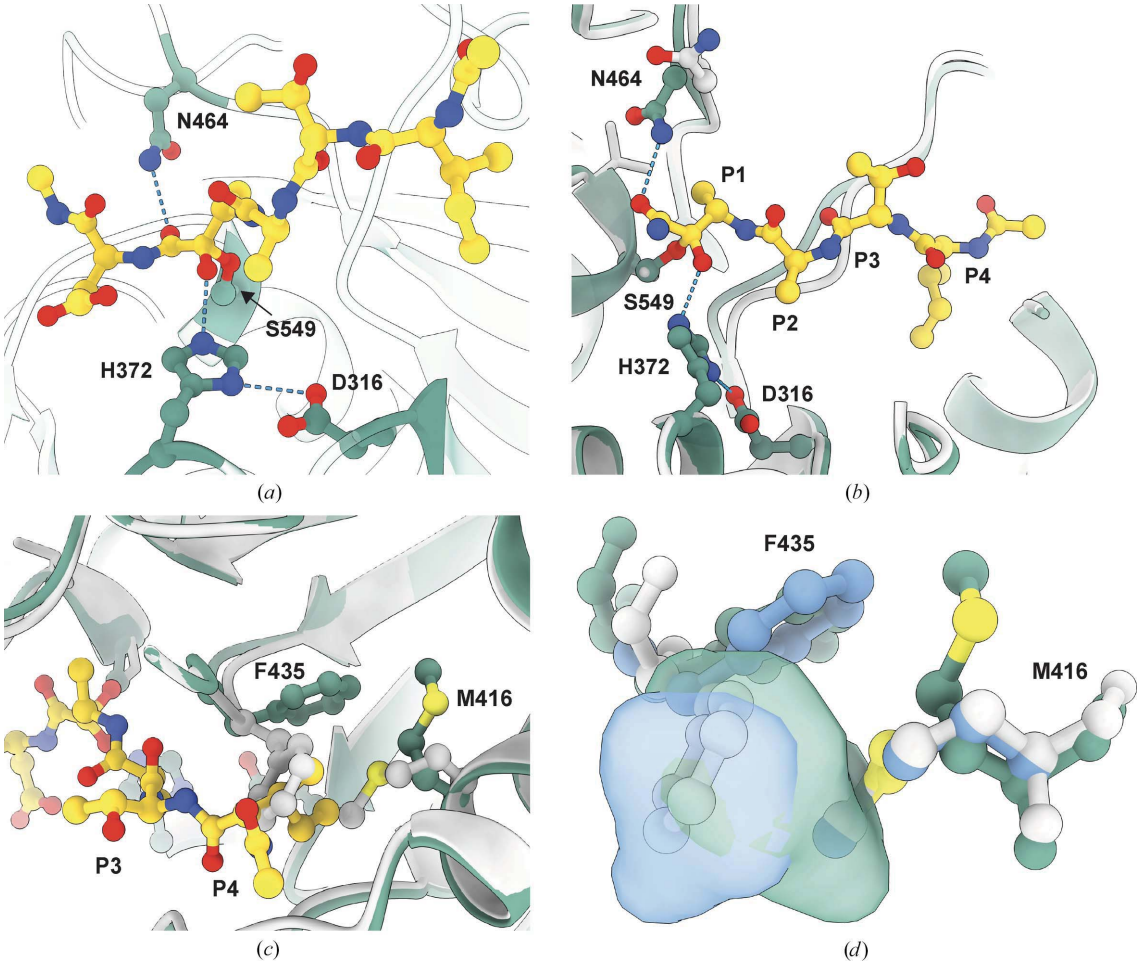


Figure 5
(*a*) Catalytic triad residues of $PvS1_{Cat\text{-}Tryps}$ (Asp316, His372 and Ser549) and Asn464 in the oxyanion hole in the transition-state analogue with MAM-117 (shown in yellow). The covalent bond between Ser549 and the α-keto group of MAM-117 in position P1 is shown by an arrow. Hydrogen bonds are depicted as blue dotted lines. (*b*, *c*) Conformational changes of Asn464 in the oxyanion hole of $PvS1_{Cat\text{-}Tryps}$ (*b*) and the Phe435 and Met416 residues of the S4 hydrophobic binding pocket (*c*) in $PvS1_{Cat\text{-}Tryps}$ in the apo form (grey) or in complex with MAM-117 (green). For clarity only the P1–P4 residues of MAM-117 are shown in (*b*). (*d*) Conformational changes of the Phe435 and Met416 residues of the S4 hydrophobic binding pocket of $PvS1_{Cat\text{-}Tryps}$ in $PvS1_{FL\text{-}Bac}$ (white) and $PvS1_{Cat\text{-}Tryps}$ in the apo form (blue) or in complex with MAM-117 (green) and the consequences for the S4 pocket volumes of $PvS1_{FL\text{-}Bac}$ (blue) and $PvS1_{Cat\text{-}Tryps}$–MAM-117 (green).

free $PvS1_{Cat}$ in the apo form. Together with the contribution of the P′ moieties of MAM-117 that clearly participate in the network of inhibitor–enzyme interactions, the plasticity of the catalytic pockets of SUB1 defines new possible pharmacophores for the development of future SUB1 inhibitors. As part of a target-based project, $PvS1_{Cat}$ represents an experimental platform to provide novel and reliable 3D structural information to accelerate the discovery of optimized SUB1 inhibitors. Increasing their affinity and selectivity and also their antiparasite activity could ultimately lead to compounds with potent *in vivo* activity that could act synergically with PmX inhibitors and define a new generation of antimalarial candidates.

Acknowledgements

We thank the staff of the Crystallography Core Facility at the Institut Pasteur for carrying out robot-driven crystallization screening. We acknowledge ESRF for the provision of synchrotron-radiation facilities and we thank the staff members of the ID29 and ID30 beamlines for support during data collection. Authorship contributions were as follows. M. Martinez, A. Blondel, M. Maurel, F. Batista, A. M. Wehenkel, L. Ortega, A. Bouillon, A. Haouz, J.-F. Hernandez and J.-C. Barale designed, performed and analysed the experiments. P. M. Alzari and J.-C. Barale designed the study. M. Martinez, A. Bouillon, F. Batista, J.-F. Hernandez, P. M. Alzari and J.-C. Barale wrote the manuscript. All authors approved the final version of the manuscript. The authors declare they have no competing interests.

Funding information

For this work, A. Bouillon and M. Martinez were supported by Agence Nationale pour la Recherche (ANR) grant ANR-11-RPIB-002 and FB and LO were supported by grant ANR-19-CE18-0010. LO was also supported by ANR-17-CE11-0030 and by the 'URGENCE COVID-19' fundraising campaign of

Institut Pasteur. Part of this work and M. Maurel were supported by Carnot Chimie Balard Institute through ANR program 11 CARN 0001-01.

References

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst.* D**79**, 449–461.

Arastu-Kapur, S., Ponder, E. L., Fonović, U. P., Yeoh, S., Yuan, F., Fonović, M., Grainger, M., Phillips, C. I., Powers, J. C. & Bogyo, M. (2008). *Nat. Chem. Biol.* **4**, 203–213.

Balikagala, B., Fukuda, N., Ikeda, M., Katuro, O. T., Tachibana, S. I., Yamauchi, M., Opio, W., Emoto, S., Anywar, D. A., Kimura, E., Palacpac, N. M. Q., Odongo-Aginya, E. I., Ogwang, M., Horii, T. & Mita, T. (2021). *N. Engl. J. Med.* **385**, 1163–1171.

Barale, J.-C., Blisnick, T., Fujioka, H., Alzari, P. M., Aikawa, M., Braun-Breton, C. & Langsley, G. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6445–6450.

Barthels, F., Schirmeister, T. & Kersten, C. (2021). *Mol. Inf.* **40**, e2000144.

Bastianelli, G., Bouillon, A., Nguyen, C., Crublet, E., Pêtres, S., Gorgette, O., Le-Nguyen, D., Barale, J.-C. & Nilges, M. (2011). *PLoS One*, **6**, e21812.

Bastianelli, G., Bouillon, A., Nguyen, C., Le-Nguyen, D., Nilges, M. & Barale, J.-C. (2014). *PLoS One*, **9**, e109269.

Bouillon, A., Giganti, D., Benedet, C., Gorgette, O., Pêtres, S., Crublet, E., Girard-Blanc, C., Witkowski, B., Ménard, D., Nilges, M., Mercereau-Puijalon, O., Stoven, V. & Barale, J.-C. (2013). *J. Biol. Chem.* **288**, 18561–18573.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Smart, O. S., Vonrhein, C. & Womack, T. O. (2009). *BUSTER* version 2.9.3. Cambridge: Global Phasing.

Buisson, M., Rivail, L., Hernandez, J.-F., Jamin, M., Martinez, J., Ruigrok, R. W. & Burmeister, W. P. (2006). *FEBS Lett.* **580**, 6570–6578.

Chang, V. T., Crispin, M., Aricescu, A. R., Harvey, D. J., Nettleship, J. E., Fennelly, J. A., Yu, C., Boles, K. S., Evans, E. J., Stuart, D. I., Dwek, R. A., Jones, E. Y., Owens, R. J. & Davis, S. J. (2007). *Structure*, **15**, 267–273.

Collins, C. R., Hackett, F., Strath, M., Penzo, M., Withers-Martinez, C., Baker, D. A. & Blackman, M. J. (2013). *PLoS Pathog.* **9**, e1003344.

Datoo, M. S., Natama, H. M., Somé, A., Bellamy, D., Traoré, O., Rouamba, T., Tahita, M. C., Ido, N. F. A., Yameogo, P., Valia, D., Millogo, A., Ouedraogo, F., Soma, R., Sawadogo, S., Sorgho, F., Derra, K., Rouamba, E., Ramos-Lopez, F., Cairns, M., Provstgaard-Morys, S., Aboagye, J., Lawrie, A., Roberts, R., Valéa, I., Sorgho, H., Williams, N., Glenn, G., Fries, L., Reimer, J., Ewer, K. J., Shaligram, U., Hill, A. V. S. & Tinto, H. (2022). *Lancet Infect. Dis.* **22**, 1728–1736.

Dvorin, J. D. & Goldberg, D. E. (2022). *Annu. Rev. Microbiol.* **76**, 67–90.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Gamo, F. J., Sanz, L. M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J. L., Vanderwall, D. E., Green, D. V., Kumar, V., Hasan, S., Brown, J. R., Peishoff, C. E., Cardon, L. R. & Garcia-Bustos, J. F. (2010). *Nature*, **465**, 305–310.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.

Gemma, S., Giovani, S., Brindisi, M., Tripaldi, P., Brogi, S., Savini, L., Fiorini, I., Novellino, E., Butini, S., Campiani, G., Penzo, M. & Blackman, M. J. (2012). *Bioorg. Med. Chem. Lett.* **22**, 5317–5321.

Giganti, D., Bouillon, A., Tawk, L., Robert, F., Martinez, M., Crublet, E., Weber, P., Girard-Blanc, C., Petres, S., Haouz, A., Hernandez, J.-F., Mercereau-Puijalon, O., Alzari, P. M. & Barale, J.-C. (2014). *Nat. Commun.* **5**, 4833.

Guerra, F. & Winzeler, E. A. (2022). *Curr. Opin. Microbiol.* **70**, 102220.

Gupta, R. & Brunak, S. (2002). *Pac. Symp. Biocomput.*, pp. 310–322.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kher, S. S., Penzo, M., Fulle, S., Finn, P. W., Blackman, M. J. & Jirgensons, A. (2014). *Bioorg. Med. Chem. Lett.* **24**, 4486–4489.

Kozak, S., Bloch, Y., De Munck, S., Mikula, A., Bento, I., Savvides, S. N. & Meijers, R. (2020). *Acta Cryst.* D**76**, 1244–1255.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.

Leroy, D., Campo, B., Ding, X. C., Burrows, J. N. & Cherbuin, S. (2014). *Trends Parasitol.* **30**, 478–490.

Leuw, P. de & Stephan, C. (2017). *GMS Infect. Dis.* **5**, Doc08.

Lidumniece, E., Withers-Martinez, C., Hackett, F., Blackman, M. J. & Jirgensons, A. (2022). *J. Med. Chem.* **65**, 12535–12545.

Lidumniece, E., Withers-Martinez, C., Hackett, F., Collins, C. R., Perrin, A. J., Koussis, K., Bisson, C., Blackman, M. J. & Jirgensons, A. (2021). *Proc. Natl Acad. Sci. USA*, **118**, e2022696118.

Manns, M. P. & Maasoumy, B. (2022). *Nat. Rev. Gastroenterol. Hepatol.* **19**, 533–550.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Monet, D., Desdouits, N., Nilges, M. & Blondel, A. (2019). *J. Chem. Inf. Model.* **59**, 3506–3518.

Moraes, M., Jorge, S. A., Astray, R. M., Suazo, C. A., Calderón Riquelme, C. E., Augusto, E. F., Tonso, A., Pamboukian, M. M., Piccoli, R. A., Barral, M. F. & Pereira, C. A. (2012). *Biotechnol. Adv.* **30**, 613–628.

Mukherjee, S., Nguyen, S., Sharma, E. & Goldberg, D. E. (2022). *Nat. Commun.* **13**, 4537.

Nasamu, A. S., Glushakova, S., Russo, I., Vaupel, B., Oksman, A., Kim, A. S., Fremont, D. H., Tolia, N., Beck, J. R., Meyers, M. J., Niles, J. C., Zimmerberg, J. & Goldberg, D. E. (2017). *Science*, **358**, 518–522.

Pace, T., Grasso, F., Camarda, G., Suarez, C., Blackman, M. J., Ponzi, M. & Olivieri, A. (2019). *Cell. Microbiol.* **21**, e13028.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.

Pino, P., Caldelari, R., Mukherjee, B., Vahokoski, J., Klages, N., Maco, B., Collins, C. R., Blackman, M. J., Kursula, I., Heussler, V., Brochet, M. & Soldati-Favre, D. (2017). *Science*, **358**, 522–528.

Roberts, E., Eargle, J., Wright, D. & Luthey-Schulten, Z. (2006). *BMC Bioinformatics*, **7**, 382.

Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.
Seidah, N. G. & Chrétien, M. (1999). *Brain Res.* **848**, 45–62.
Seidah, N. G. & Prat, A. (2012). *Nat. Rev. Drug Discov.* **11**, 367–383.
Siezen, R. J. & Leunissen, J. A. M. (1997). *Protein Sci.* **6**, 501–523.
Tan, M. S. Y., Koussis, K., Withers-Martinez, C., Howell, S. A., Thomas, J. A., Hackett, F., Knuepfer, E., Shen, M., Hall, M. D., Snijders, A. P. & Blackman, M. J. (2021). *EMBO J.* **40**, e107226.
Tawk, L., Lacroix, C., Gueirard, P., Kent, R., Gorgette, O., Thiberge, S., Mercereau-Puijalon, O., Ménard, R. & Barale, J.-C. (2013). *J. Biol. Chem.* **288**, 33336–33346.
Thomas, J. A., Tan, M. S. Y., Bisson, C., Borg, A., Umrekar, T. R., Hackett, F., Hale, V. L., Vizcay-Barrena, G., Fleck, R. A., Snijders, A. P., Saibil, H. R. & Blackman, M. J. (2018). *Nat. Microbiol.* **3**, 447–455.
Uwimana, A., Legrand, E., Stokes, B. H., Ndikumana, J. M., Warsame, M., Umulisa, N., Ngamije, D., Munyaneza, T., Mazarati, J. B., Munguti, K., Campagne, P., Criscuolo, A., Ariey, F., Murindahabi, M., Ringwald, P., Fidock, D. A., Mbituyumuremyi, A. & Menard, D. (2020). *Nat. Med.* **26**, 1602–1608.
Weber, P., Pissis, C., Navaza, R., Mechaly, A. E., Saul, F., Alzari, P. M. & Haouz, A. (2019). *Molecules*, **24**, 4451.
Withers-Martinez, C., Saldanha, J. W., Ely, B., Hackett, F., O'Connor, T. & Blackman, M. J. (2002). *J. Biol. Chem.* **277**, 29698–29709.
Withers-Martinez, C., Strath, M., Hackett, F., Haire, L. F., Howell, S. A., Walker, P. A., Christodoulou, E., Dodson, G. G. & Blackman, M. J. (2014). *Nat. Commun.* **5**, 3726.
Withers-Martinez, C., Suarez, C., Fulle, S., Kher, S., Penzo, M., Ebejer, J. P., Koussis, K., Hackett, F., Jirgensons, A., Finn, P. & Blackman, M. J. (2012). *Int. J. Parasitol.* **42**, 597–612.
World Health Organization (2021). *World Malaria Report*. Geneva: World Health Organization. https://www.who.int/publications/i/item/9789240040496.
Yeoh, S., O'Donnell, R. A., Koussis, K., Dluzewski, A. R., Ansell, K. H., Osborne, S. A., Hackett, F., Withers-Martinez, C., Mitchell, G. H., Bannister, L. H., Bryans, J. S., Kettleborough, C. A. & Blackman, M. J. (2007). *Cell*, **131**, 1072–1083.