The nepenthesin insert in the *Plasmodium falciparum* aspartic protease plasmepsin V is necessary for enzyme function

Received for publication, February 28, 2022, and in revised form, August 1, 2022. Published, Papers in Press, August 9, 2022. https://doi.org/10.1016/j.jbc.2022.102355

Alexander J. Polino¹, Justin J. Miller², Soumendranath Bhakat ¹, Sumit Mukherjee¹, Suhas Bobba¹, Gregory R. Bowman⁷, and Daniel E. Goldberg¹*

From the ¹Division of Infectious Diseases, Department of Medicine and Department of Molecular Microbiology, and ²Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri, USA

Edited by Chris Whitfield

Plasmepsin V (PM V) is a pepsin-like aspartic protease essential for growth of the malarial parasite *Plasmodium falciparum*. Previous work has shown PM V to be an endoplasmic reticulum–resident protease that processes parasite proteins destined for export into the host cell. Depletion or inhibition of the enzyme is lethal during asexual replication within red blood cells as well as during the formation of sexual stage gametocytes. The structure of the *Plasmodium vivax* PM V has been characterized by X-ray crystallography, revealing a canonical pepsin fold punctuated by structural features uncommon to secretory aspartic proteases; however, the function of this unique structure is unclear. Here, we used parasite genetics to probe these structural features by attempting to rescue lethal PM V depletion with various mutant enzymes. We found an unusual nepenthesin 1-type insert in the PM V gene to be essential for parasite growth and PM V activity. Mutagenesis of the nepenthesin insert suggests that both its amino acid sequence and one of the two disulfide bonds that undergird its structure are required for the insert’s role in PM V function. Furthermore, molecular dynamics simulations paired with Markov state modeling suggest that mutations to the nepenthesin insert may allosterically affect PM V catalysis through multiple mechanisms. Taken together, these data provide further insights into the structure of the *P. falciparum* PM V protease.

Malaria remains a substantial burden on the developing world, with over 240 million cases and 600,000 deaths annually (https://www.who.int/publications/i/item/9789240040496. Accessed August 26, 2022). While the parasite life cycle includes several host tissues, symptomatic human disease results entirely from infection of red blood cells (RBCs). After invasion of an RBC, the parasite dramatically modifies the host cell by exporting hundreds of effector proteins into the host cytosol. This program of protein export results in altered host cell rigidity, increased permeability to nutrients, and the constitution of a secretory apparatus that facilitates the display of parasite-encoded adhesins on the RBC surface (1, 2). These adhesins mediate binding of parasitized RBCs to host vascular endothelia, resulting in vascular disruption and, in extreme cases, death of the host.

Most known exported *Plasmodium falciparum* effectors bear a motif termed the *Plasmodium* export element (PEXEL) near the protein N terminus (3, 4). PEXEL—with consensus sequence RxLxE/Q/D—is cleaved in the parasite endoplasmic reticulum (ER) by the aspartic protease plasmepsin V (PM V) (5, 6). Inhibition of PM V activity disrupts protein export and kills parasites (7). PM V is essential at two distinct points of intraerythrocytic development (8), as well as during gametocytogenesis (9), raising the profile of this enzyme as a potential target for chemotherapeutic intervention.

A recent X-ray crystallographic structure of the *Plasmodium vivax* PM V bound to the peptidomimetic inhibitor WEHI-842 revealed features unusual for a secretory aspartic protease (10). A nepenthesin 1-type insert and a helix–turn–helix domain feature prominently at opposite surfaces of the protein. The nepenthesin insert—so-called as it was identified in the protease nepenthesin 1 from the digestive fluid of the pitcher plant *Nepenthesia*—is a trilobed structure held together by two pairs of disulfide-bonded cysteines (11). Its functional role is unknown. The helix–turn–helix domain is enigmatic in its own right as the motif canonically mediates protein–nucleic acid binding, which is not expected for PM V in the ER lumen. The structure in addition revealed a free cysteine at the flap over the active site, raising the possibility of redox regulation. Finally, not visible in the *P. vivax* PM V structure is a predicted flexible loop projecting away from the active site. In *P. falciparum* PM V, the corresponding sequence represents a substantially larger and highly charged insert, the role of which we were curious to determine.

Here, we utilized our previously described lethal depletion of PM V (8) to probe the aforementioned features. Using a genetic rescue method, we found the helix–turn–helix domain, free cysteine, and *P. falciparum*–specific charged insert to be dispensable for parasite growth. In contrast, PM V without a nepenthesin insert was unable to rescue PM V depletion. Further mutagenesis of the nepenthesin insert revealed that the amino acid sequence of the insert as well as some elements of its structure are required for its function. We deployed molecular dynamics simulations to understand how nepenthesin insert mutagenesis impacts the conformational
Plasmepsin V nepenthesin loop

dynamics of PM V. Our simulations show that the nepenthesin insert allosterically modulates positioning of the active site residues, and mutation leads to increased prevalence of catalytically nonfunctional states. Taken together, these data broaden our model of how PM V enacts its essential function(s) and provide new insight into the role of the nepenthesin insert in protease biochemistry.

Results

P. falciparum PM V has several unusual features

We used SWISS-MODEL to model PM V from the genetically tractable P. falciparum, using the recently published P. vivax PM V structure (Fig. 1)(12). The model closely resembles the P. vivax structure (Fig. S1), suggesting that the P. falciparum enzyme bears the same unusual motifs found in the P. vivax enzyme. The nepenthesin insert is present in PM V across the Plasmodium genus, yet is absent in other plasmepsins (Figs. 1A and S2). The spacing between the four key cysteines that comprise the insert’s structure is conserved among PM V sequences, whereas the intervening sequences contain some variation (Figs. 1A and S2). The unpaired cysteine over the active site is completely conserved in PM V sequences across the genus (Fig. 1B). The charged insertion projecting away from the active site is substantially expanded in P. falciparum relative to P. vivax and includes an unusually dense run of charged amino acids (Fig. 1C). Finally, the helix-turn-helix domain is conserved throughout the genus but absent in other P. falciparum plasmepsins (Figs. 1D and S3).

Regulatable system for PM V depletion and rescue

To probe the function of these structural motifs, we utilized a previously described post-transcriptional depletion of PM V (8). Briefly, we used CRISPR/Cas9 editing to replace the endogenous PM V sequence with a recodonized copy flanked by an N-terminal aptamer and a 10× array of C-terminal aptamers (Fig. 2A). In the mRNA, the aptamers fold into hairpins evolved to bind the Tet repressor (TetR). TetR is fused to the RNA helicase DOZI, and the binding of the TetR–DOZI fusion protein to the mRNA results in sequestration of the mRNA and depletion of protein levels. The system is controlled by the presence of anhydrotetracycline (aTc). In the presence of aTc, TetR–DOZI preferentially binds aTc, and translation continues unabated (13). In the absence of aTc, TetR–DOZI binds the aptamers, and translation is suppressed. Knockdown depletes PM V levels greater than 50-fold and is lethal to parasites (8). We attempted to complement the depletion with a second copy of PM V, tagged with C-terminal GFP, and integrated into the neutral cg6 locus of NF54attB.
parasites using the previously described attP/attB system (Fig. 2A) (14).

To probe the role of PM V structural motifs, we made six PM V-GFP versions: a full-length second copy (WT), a D118A mutant that alters one of the catalytic aspartates ablating catalytic activity ("catalytic dead" or CD) (10), a complete deletion of the nepenthesin insert from the first to fourth cysteine (ΔNep), replacement of the free flap cysteine with alanine (C178A), deletion of the charged insert (ΔCh. Ins.), and replacement of the helix–turn–helix motif with two alanines (ΔHTH) (Fig. 2B). Expression of each is driven by the PM V promoter, with the exception of ΔHTH, for which cloning issues forced us to use an alternative plasmid driven by the Hsp86 promoter. Parasite lines were selected in the presence of aTc, using blasticidin S (for the knockdown construct) and WR-99210 (for the second copies), and cloned by limiting dilution. The resulting lines each expressed the endogenous PM V as well as the second copy at approximately 2:1 M ratio (Figs. S4 and 2C); the ΔHTH line, being driven by a stronger promoter, expressed the second copy at an approximately equimolar ratio to the endogenous enzyme (Figs. S4 and 2C). Deletion of the charged insert resulted in an enzyme that was no longer recognized by our previously described PM V monoclonal antibody (15, 16) but was still recognized by anti-GFP (Fig. S4), suggesting that this region contains the binding site for the PM V monoclonal antibody.

**The nepenthesin insert is required for parasite growth**

To probe the role of each of these motifs in PM V function, we depleted endogenous PM V by washing aTc from parasite culture and followed parasite growth daily by flow cytometry. In each case, aTc washout caused substantial depletion of endogenous PM V levels (Figs. 2C and S8). The second copy of WT PM V rescued parasite growth whereas CD PM V did not (Figs. 2D and S5). To our surprise, parasite lines expressing PM V lacking the free cysteine, charged insert, or helix–turn–helix domain all grew indistinguishably with/without aTc, suggesting each of these motifs is unnecessary for the essential function of PM V in parasite growth. Only the ΔNep mutant failed to rescue growth in the absence of aTc, suggesting that the nepenthesin insert is required for proper PM V function in vivo.

To further dissect the functional contribution of the nepenthesin insert, we mutagenized the insert, attempting to separate its structure from the sequence of its loops. To this end, we made three additional mutants: replacing either the first or second pair of disulfide-bonded cysteines with serines (NepC1,3S and NepC2,4S, respectively), or replacing the entire nepenthesin insert with the orthologous sequence from the pitcher plant protease Nep1A (NepNep1A) (Fig. 3, A and B), to preserve the insert’s fold while largely altering the lobe sequences. As aforementioned, we transfected these as second copies into parasite lines with regulatable endogenous PM V. Washing aTc from the cultures depleted endogenous PM V...
levels but had no effect on levels of the second copy (Figs. 3 and S9). After washout, we followed growth daily by flow cytometry (Fig. 3D). PM V with the Nep1A sequence failed to rescue growth, suggesting that the general structure of the insert is not sufficient for its role in PM V function. Of the two disulfide-bonded cysteine pair mutants, replacement of the first had no effect on growth, whereas replacement of the second resulted in a partial growth phenotype—these parasites grew at approximately half the rate without aTc as with. We have previously shown expression of PM V at or above 3% of WT levels is adequate to achieve normal growth (8). Therefore, the fact that the NepC2,4S mutant has impaired growth suggests a substantial defect in enzyme function.

To rule out the possibility that PM V grossly mislocalizes when the nepenthesin insert is disrupted, we fixed ring-stage parasites and compared the GFP-tagged mutant enzyme with the FLAG-tagged endogenous copy. We found that in each case the mutant colocalized well with the ER-resident endogenous copy (Fig. S6), suggesting the nepenthesin insert is not required for proper localization of PM V.

The nepenthesin insert is required for PM V catalysis

We next queried whether the nepenthesin mutants failed to rescue growth because of an effect on PM V catalytic activity or whether they interfered with some other essential PM V role (e.g., a protein–protein interaction). Recombinant expression in mammalian cells afforded production of some but not all the enzyme mutants (Fig. S7). To employ a system where we could compare the full range of mutants, we immunoprecipitated PM V enzyme mutants using anti-GFP from the parasites described previously. The eluate contains each mutant PM V without any detectable signal from the endogenous enzyme (Fig. 4A). We then assessed the activity of each mutant enzyme using a previously described FRET peptide assay to assess their activity (5). Briefly, we incubated each enzyme preparation with either a “PEXEL” peptide derived from the recognition sequence of the PM V substrate HRP2 or a “mutant” peptide that has a single amino acid change, abrogating PM V recognition. Cleavage of the peptide substrate liberates the fluorophore EDANS from its quencher DABCYL, resulting in fluorescence (Fig. 4B). Incubation with the FRET peptides indicated that the enzymes that could rescue growth (WT, NepC1,3S) were active against PEXEL peptide but not the mutant peptide (Fig. 4C). The enzymes that could not rescue growth (CD, ΔNep, and NepNep1A) could cleave neither peptide. The NepC2,4S enzyme that partially rescues growth appears inactive in our enzyme assay, perhaps because of the in vitro activity buffer not fully recapitulating the environment of the parasite ER or because of the assay’s insensitivity given the small amounts of NepC2,4S we were able to purify (Fig. 4A).

Given the canonical role of pH in activating pepsin-like proteases, and the example of the Nep1A protease being active at low pHs (maximal activity at pH 2–3) (11), we sought to determine whether nepenthesin mutants might have activity at alternative pH ranges. To this end, we probed for PM V activity in buffers ranging from pH 3 to pH 7. We found that the two enzymes active at pH 6.5 (WT and NepC1,3S) were the only ones to display activity at any pH tested, and each displayed a similar pH profile to that previously described (Fig. 5) (5). Thus, perturbations to the nepenthesin insert do not influence the enzyme’s pH–activity relationship.

Mutations in the nepenthesin insert may impact flap conformations

The C2,4 disulfide-bonded cysteine pair tethers the nepenthesin insert directly to the N-terminal strand of the
flap over the active site (Fig. 6A). This tethering has been postulated as a mechanism for functional regulation of PM V and nepenthesin 1, both through protein–protein interactions and protein allostery (10, 11). To investigate the impact of nepenthesin insert mutations on the conformational preferences of the insert and its adjacent flap, we performed molecular dynamics simulations of WT and NepNep1A PM V. We generated initial structures of both P. falciparum PM V and NepNep1A PM V using the previously solved structure of P. vivax PM V (Protein Data Bank ID: 4ZL4) as a template and Swiss-Model (12, 17). For both constructs, we ran seven independent unbiased molecular dynamics simulations for 1 μs simulation time per replicate for an aggregate simulation time of 7 μs per mutant. To quantify our results and highlight conformational heterogeneity, we built Markov state models (MSMs), a statistical framework to analyze molecular dynamics simulations that provides atomistic insights into ensembles of conformations that proteins can adopt (18).

We first investigated how modifications affect the conformational preferences of the nepenthesin insert. We calculated the root mean square deviation per residue of the nepenthesin loop for 1 μs simulation time per replicate for an aggregate simulation time of 7 μs per mutant. To quantify our results and highlight conformational heterogeneity, we built Markov state models (MSMs), a statistical framework to analyze molecular dynamics simulations that provides atomistic insights into ensembles of conformations that proteins can adopt (18).

Figure 4. Nepenthesin mutants reduce catalytic activity. A, Western blot with anti-PM V showing immunoprecipitations of mutant PM V-GFPs. B, activity assay scheme showing the sequence of the PEXEL peptide (“WT peptide”) and the mutant peptide (“Mut peptide”). Immunoprecipitated PM V (blue) is incubated with a PEXEL peptide separating the FRET pair EDANS and DABCYL. When intact, DABCYL quenches EDANS fluorescence. When cleaved, EDANS is liberated, and blue fluorescence is measured. C, immunoprecipitated mutant enzymes were incubated with a PEXEL FRET peptide or a nonrecognized mutant peptide for 12 h at 37 °C, and then fluorescence intensity read as a measure of catalytic activity. Each sample was assayed in technical triplicate, mean fluorescence is shown, with error bars representing standard deviation of the measurements. The experiment was performed three times; a representative experiment is shown. PEXEL, Plasmodium export element; PM V, plasmepsin V.

Figure 5. Nepenthesin mutants do not influence pH profile. Activity of each enzyme was assayed as aforementioned (Fig. 4C) at buffers that differed only in pH, with pH ranging from 3 to 9 (shown above bars with red numbers). Each sample was incubated with either the PEXEL peptide (green) or the mutant (gray). Each sample was run in technical triplicate. The bar height represents the mean of those three measurements; error bars represent the standard deviation. The experiment was performed twice with similar results, one is shown. PEXEL, Plasmodium export element.
activity in NepNep1A is not a result of gross changes in nepenthesin insert conformational preferences.

Curious if mutations to the nepenthesin insert result in allosteric changes that may impact PM V activity, we next turned our attention to processes likely to disrupt catalysis. During PEXEL cleavage, PM V must adopt a fl open conformation more open than that seen in the crystal structure to accommodate PEXEL binding. Once PEXEL is bound, we expect the fl to close, trapping the PEXEL motif next to the catalytic aspartates for hydrolysis to occur. We reasoned that if the equilibrium distribution of fl conformations was disrupted in NepNep1A, this may lead to a loss of catalytic activity. Accordingly, we quantified the extent of fl opening by measuring the minimum distance between D118, one of the active site aspartates, and C178, the free cysteine on the tip of the fl (Fig. 6C). In WT PM V, the fl exists in two primary populations, one more closed in a pose similar to that seen in the crystal structure of P. vivax PM V, and another slightly expanded. While the NepNep1A fl explores these distances, it also explores conformations where the fl has significantly dissociated from the active site. We calculated the probability of the loop exploring a significantly dissociated state (D118–C178 distance >1.5 nm) and find that WT has an equilibrium probability of 0.3 ± 0.1% for these states, whereas NepNep1A has a probability of 17.8 ± 5.5%.

While characterizing the motion of the fl, we noticed extensive hydrogen bonding between the fl and the distal cavity of the active site. Previous studies have highlighted that loss of H-bonding between the fl and active site adjacent pockets of aspartic proteases dramatically impacts the catalytic activity of aspartic proteases (19). Notably, in WT PM V, S125 spends the majority of our simulations in a hydrogen bond with either Q175 or H212 (mimicking the H-bond network between Y77 and W41 in PM II) (20, 21). In NepNep1A, H-bonding to S125 is significantly disrupted (Fig. 7). In Nep-Nep1A, when S125 is not in an H-bond, the fl is more likely to explore conformations where the fl is skewed significantly to the side. This skewing allows the stable formation of a H-bond between Q175 on the fl and N236 in the distal loop (Fig. 6, A and C), a configuration not observed in our WT simulations.

We also wanted to understand why NepC1,3S maintains activity, whereas NepC2,4S results in a loss of activity. In the crystal structure of P. vivax PM V, the NepC1,3 disulfide bond holds the nepenthesin insert together, whereas NepC2,4

**Figure 6.** Computational modeling of *Plasmodium falciparum* PM V mutants suggests mechanisms for loss of activity. A, structural model of *P. falciparum* PM V adapted from Protein Data Bank (ID: 4ZL4). B, impact of Nep insert mutations on flexibility of the Nep insert. Shown is the equilibrium distribution of RMSD distances (in nanometers) for the Nep insert (L153–Y173), normalized per residue. C, impact of Nep insert mutations on the opening of the fl covering the active site. Shown are representative structures for different D118–C178 distances for WT PM V as well as the equilibrium distance distributions for each mutant. Quantification is the percent of states that occupy a distance larger than a fl opening cutoff of 1.5 nm ± SD. D, impact of Nep insert mutations on distance between Q175 (flap) and N236 (distal loop). Shown are representative structures for different channel widths and the equilibrium distributions for each PM V mutant. Quantification is the percent of states that occupy a Q175–N236 distance >0.5 nm ± SD. E, impact of Nep insert mutations on active site channel opening distance. Channel width defined as the minimum distance between any atom on side chains 176 to 181 and 485 to 490. Quantification is the percent of states with their channel more open than 0.4 nm ± SD. For all quantifications, the displayed value is the mean ± SD as determined by bootstrapping the MSM model by iteratively removing one of the seven trajectories for each mutant. MSM, Markov state model; PM V, plasmepsin V.
anchors the nepenthesin insert to the active site flap. As before, we simulated seven independent replicates for 1 μs each for NepC1,3S and NepC2,4S and built an MSM to quantify our results. While the NepC1,3S nepenthesin insert had similar dynamics to NepNep1A, NepC2,4S was much more flexible than all other mutants (Fig. 6B). Similarly, the NepC1,3S flap explores conformations similar to WT, whereas NepC2,4S explores conformations more similar to NepNep1A (Fig. 6C). We expected that the increased flexibility of the nepenthesin insert resulting from NepC2,4S would allow the flap to similarly interact with the distal loop, whereas NepC1,3S would, like WT PM V, never interact with the distal loop. Surprisingly, neither NepC1,3S nor NepC2,4S interact with the distal loop (Fig. 6D).

Having explored two mechanisms of catalytic prevention—active site closure by flap clamping/dissociation and loss of substrate interaction through flap pinning—we wanted to check the final mechanism by which the active site may be occluded: through closure of the active site. Accordingly, we measured the minimum distance between the active site flap and the loop opposite the active site, which forms the active site channel (Fig. 6E). In our simulations of WT PM V, the channel is “closed” (defined as roughly the width of a peptide backbone, minimum channel distance < 4 Å) 35.7 ± 7.3% of the time. NepNep1A and NepC1,3S have closed channels 17.4 ± 4.3% and 28.7 ± 4.3% of the time, whereas NepC2,4S has a stark increase in the population of closed channel conformations, occurring 55.1 ± 10.1% of the time.

Discussion

Here, we used parasite genetics and molecular dynamics simulations to investigate the relationship between the unusual structure of PM V and its function in parasite biology. We find that the PM V nepenthesin insert is required for parasite survival and for catalysis of a substrate peptide in vitro, whereas a helix–turn–helix motif, free cysteine near the active site, and large charged insert are dispensable for parasite survival (and presumably for PM V catalysis).

The most intriguing motif in our story is the nepenthesin insert. This motif was originally described by Athauda et al. (11) as a 22-residue insertion unique to nepenthesin-like aspartic proteases (and absent from their closest plant orthologs). The finding that PM V bears a similar insert was a surprise (16), given that it has little in common with the nepenthesin-type aspartic proteases. The nepenthesin proteases are digestive enzymes, secreted into the pitcher plant lumen, and activated at very acidic pH (optimal activity at pH ≈ 2.5) to participate in the digestion of insect prey material. Both nepenthesin I and nepenthesin II have broad substrate specificity and can cleave before or after many residues (22). In contrast, PM V is retained in the ER at much higher pH (optimal activity at pH ≈ 6.5) and has only been described to cleave C terminally to RxL (5, 23). The same motif in both sets of aspartic proteases, but not in others, defies obvious explanation.

Athauda et al. speculated that the nepenthesin insert may influence localization or regulation of the enzyme; the subsequent finding of a nepenthesin insert in PM V makes at least the former seem less likely. Hodder et al. (10) suggested that the insert may be involved in interaction with other proteins, perhaps the PEXEL substrates themselves. We have not tested this hypothesis here but can comment that the Toxoplasma gondii ortholog Asp5 cleaves an RRL motif that is very similar to the RxL recognized by PM V (17–19), yet bears a large insert between the third and fourth cysteine of its orthologous nepenthesin insert sequence, giving spacing of CX₅CX₃CXₓC compared with the CX₃CX₃CX₃C spacing in Plasmodium. One could imagine that this extra 15-amino acid run would disrupt the nepenthesin insert; however, without structural information on Asp5, this remains unclear. We have speculated that the insert may serve as a docking point for chaperones (24), yet to our surprise all nepenthesin insert mutations that disrupted parasite growth also blocked catalysis in vitro, suggesting that the insert contributes to substrate recognition, access, or cleavage. We cannot fully rule out that disrupting the nepenthesin loop results in a folding defect severe enough to inactive the enzyme; however, we think this is unlikely for the following reasons: (1) normal location and expression level of most mutants (we would expect some degradation and/or clustering with misfolding), (2) normal GFP fluorescence signal, suggesting proper folding of at least the C-terminal half of the molecule, (3) molecular dynamics simulations predict and explain the loss of activity without a folding defect. We therefore propose that the nepenthesin loop is likely to be important for enzyme function.
Our molecular dynamics simulations suggest that mutation of the nepenthesin insert results in a loss of catalytic activity for three reasons. First, Nep\textsuperscript{Nep1A} and Nep\textsuperscript{C2.4S} increase the likelihood that the active site flap has dissociated from the active site and is therefore unable to hold a substrate in the active site to enable hydrolysis. Second, Nep\textsuperscript{Nep1A} results in an extra flexible flap that becomes stably pinned to the distal loop via H-bonding. This pinning would likewise preclude the flap holding a substrate in the active site. Finally, Nep\textsuperscript{C2.4S} results in an increased probability of a closed active site channel, which may prevent substrate binding. The relative contributions of these mechanisms are yet unknown. In this regard, our work was limited by our inability to generate recombinant PM V with our full set of mutations of interest. Immunoprecipitation of tagged enzyme from parasites allowed us to probe activity in a relatively insensitive endpoint assay but prevented us from interrogating enzyme kinetics, form/structure, and binding characteristics to substrate peptides.

As for the other structural motifs studied here, the helix–turn–helix, free cysteine, and large charged insertion each appear dispensable for parasite growth in culture. Helix–turn–helix domains canonically mediate protein–DNA interactions but have also been described to have roles in protein–RNA interactions, protein–protein interactions, and structural support for enzymatic domains (25). In this case, the conservation of the domain in PM V sequences throughout the genus suggests a significant role for this motif. Our data are limited to blood-stage parasites in culture; perhaps studies in different Plasmodium life stages or animal models of infection would reveal a yet unappreciated role.

Similarly, the free cysteine appears dispensable for parasite growth. The enzyme domain of PM V is rich in cysteines, with 15 cysteines arranged into seven disulfide-bonded pairs—compared with plasmepsin II, which has five cysteines arranged into two pairs (10, 26). This disulfide-rich architecture has earned PM V classiﬁed into MEROPS peptidase subfamily A1B, with type member Nep1, rather than the canonical pepsin subfamily A1A (which includes all other plasmepsins) (27). In A1B subfamily members, the richer network of disulfide bonds has been hypothesized to bolster the enzyme structure against harsh conditions (11). This makes sense for the nepenthesins, which must act in acidic digestive conditions; perhaps for PM V, this enhanced structure protects against the oxidizing environment of the ER. The free cysteine C178 has been hypothesized to have a role in metal binding, and indeed recombinant PM V with the same C178A mutation we describe here was shown to have altered susceptibility to inhibition by Hg\textsuperscript{2+} ions (28). The physiological implications of this observation remain unresolved, as we see no impact of this mutation in parasite culture. C178 and its surrounding amino acids are perfectly conserved across the Plasmodium genus. Whether that conservation presages a functional contribution at another part of the parasite life cycle, we cannot yet say.

Finally, the P. falciparum genome is inexplicably full of insertions of charged and presumably poorly structured sequences in proteins that lack these insertions in the rest of the genus (29, 30). These insertions often form surface loops with no known function (30). Muralidharan et al. (31) showed that deletion of an asparagine-rich repeat from the proteasome subunit rpn6 had no discernible effect on the protein’s essential function and suggested that the P. falciparum Hsp110 may be so effective at preventing aggregation of repetitive prionogenic sequence as to nearly eliminate the fitness cost of maintaining these sequences (32). Here, we describe a similar example: a surface-exposed low-complexity loop, rich in charged amino acids, that has no apparent role in PM V function. Whether these sequences play roles in certain setting or are merely an evolutionary quirk merits further exploration.

In summary, our data provide insight into the essential function of PM V in the P. falciparum intraerythrocytic cycle. Given the unusual nature of this enzyme’s structure, this also provides atomistic insights into the contribution of the nepenthesin insert and its role in side-chain heterogeneity that modulates conformational dynamics of PM V, which we hope will support further investigation into this motif and A1B subfamily proteases more broadly. Future studies will also highlight role of side-chain heterogeneity in flap opening and designing novel “open flap” inhibitors targeting PM V (33, 34).

**Experimental procedures**

**Modeling and alignments**

The P. falciparum PM V structure was modeled using SWISS-MODEL (12, 17). Mutant sites were selected by aligning the P. falciparum PM V structural model with the published P. vivax PM V (10) one using PyMOL (version 2.3, Schrödinger). Sequence alignments were performed using reference genome sequences retrieved from PlasmoDB (35) and aligned with Clustal Omega (36). Structures for molecular dynamics simulations were generated using SWISS-MODEL (12, 17) using the published structure of P. vivax PM V (10) as a homology model.

**DNA constructs**

To modify the parasite genome, we used the previously described AttP/AttB recombinase system, which utilizes a recombinase-expressing plasmid pINT (14) along with a donor plasmid that gets integrated into the genome at the cg6 locus. Here, our donor plasmid was a derivative of pEOE (32, 37) modified to include the attP sequence using the QuikChange Lightning Multi Site Directed Mutagenesis kit and the primer 1 (see Table S1 for list of primers used in this study) as in (38) to create the plasmid pEOE-AttP. To make PM V complementation vectors, the full PM V sequence and its 5’ UTR were PCR amplified from the NF54 genome using primers 2 and 3, gel purified, and inserted by Infusion (Clontech) into pEOE-AttP cut with AatII and AvrII. The resulting plasmid, pEOE-AttP-PM V\textsubscript{prom}PM V-GFP, was then mutated using the QuikChange Lightning Multi Site Directed Mutagenesis kit and primers 5, 6, 7, 8, 10, 11, or 12 to make rescue plasmids with each PM V mutant in Figures 2B and 4A. The exception, as noted aforementioned, was the ΔHTH mutant. This plasmid was instead made by amplifying the PM V sequence...
with primers 2 and 4, and inserting into pEOE-AttP cut with XhoI and AvrII (maintaining the existing hsp86 promoter), then mutagenizing as aforementioned but with primer 9. Plasmid sequences were verified by Sanger sequencing (Genewiz) using primers 13, 14, 15, and 16.

For mammalian cell expression, we ordered a synthesized sequence for the soluble domain of PM V (Y36-S520) codon optimized for expression in human cells (Integrated DNA Technologies). This sequence was PCR amplified with primers 17 and 18, then inserted into KpnI-digested pHLSec (39) by Infusion. Plasmids for expressing mutant enzyme were made as aforementioned, using primers 19, 20, 21, 22, and 23. Completed plasmids were verified by Sanger sequencing using primers 24 and 25.

Parasite lines and culture

P. falciparum strain NF54attB (14) was cultured in RPMI1640 (Gibco) supplemented with 0.25% (w/v) Albumax (Gibco), 15 mg/l hypoxanthine, 110 mg/l sodium pyruvate, 1.19 g/l Hepes, 2.52 g/l sodium bicarbonate, 2 g/l glucose, and 10 mg/l gentamicin as well as human RBCs at 2% hematocrit. We obtain unidentified human RBCs from St Louis Children’s Hospital blood bank as well as American Red Cross Blood Services. Cultures were grown in the presence of 100 nM aTc (Gibco), 15 mg/l hypoxanthine, 110 mg/l sodium pyruvate, 0.035% saponin (10 ml PBS + saponin per 50 ml culture) for 15 min at 4 °C. Parasites were pelleted at 15,500g for 3 min, resuspended in PBS, pelleted again at 15,500g for 3 min, and then frozen at -80 °C for days to weeks until the day of the assay. Eventually, pellets were thawed and solubilized in PBS + 0.5% Triton X-100, sonicated to disrupt the pellet, and then incubated at 4 °C for 1 h. Lysates were then incubated for 1 h at 4 °C with Dynabeads-Protein A or Dynabeads-Protein G (Thermo Fisher Scientific) that had been bound to anti-GFP (clone 3E6; Thermo Fisher Scientific). Beads were collected on a magnet stand, washed five times in cold PBS + 0.02% Tween-20, and then incubated directly with either the fluorogenic PEXEL peptide DABCYL-LNKKLLHETQ-EDANS, or the mutant peptide DABCYL-LNKKLAHETQ-EDANS at 1 μM in PM V activity buffer (50 mM Tris–malate, 50 mM NaCl, 0.05% Triton X-100, pH 6.5 unless otherwise noted) for 12 h at 37 °C. Total fluorescence was read on a multimodal plate reader (PerkinElmer Envision). The pH profiling assay was performed as aforementioned, but with PM V activity buffer brought to pH 3 with hydrochloric acid, or brought up to pH 4, 5, 6, 7, 8, or 9 with sodium hydroxide.

Western blotting

Western blots were performed (8). Briefly, RBCs were lysed in PBS + 0.035% saponin, wash to remove hemoglobin, and the resulting parasite pellets solubilized with radioimmunoprecipitation assay (50 mM Tris, pH 7.4, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, and 0.5% dimethoxy-4-chloroanphethamine) plus HALT-Protease Inhibitor Cocktail (Thermo Fisher Scientific), cleaved by centrifugation to remove hemoglobin crystals, then diluted into 4× Protein Loading Buffer (Licor), and boiled at 99 °C. Prepared lysates were separated on Mini-PROTEAN TGX Gels (4–15%; Bio-Rad), transduced to 0.45 μm nitrocellulose, blocked in StartingBlock blocking buffer (Thermo Fisher Scientific), and probed with mouse anti-PM V (1:50 dilution) (15) and/or rabbit anti-GFP (1:100 dilution; Abcam; catalog no.: ab6556), followed by secondary antibodies goat antimouse IRDye 800CW 1:10,000 dilution (Licor), and donkey anti-rabbit IRDye 680RD 1:10,000 (Licor). Stained membranes were imaged on a Licor Odyssey platform, and brightness/contrast adjusted for display using ImageStudio Lite 5.2 (Licor).

Growth assays

To assess the ability of each mutant enzyme to rescue parasite growth, cultures were washed to remove aTc (three washes, 5 min each) and resuspended in medium containing either 100 nM aTc or an equal volume of dimethyl sulfoxide (to deplete endogenous PM V). Growth was monitored daily by flow cytometry using a BD FACS Canto and measuring fluorescence from acridine orange (1.5 μg/ml in PBS). Parasites were subcultured as needed to prevent overgrowth. Parasitemias were quantified using BD FACSDiva and are displayed as a ratio of the parasitemias of the dimethyl sulfoxide culture ("−aTc") to the aTc culture (+aTc). Graphs were made using GraphPad Prism (version 9; GraphPad Software, Inc).

Peptide assays

PM V activity assays were performed (5). Briefly, parasite cultures were pelleted and then RBCs were lysed in PBS + 0.035% saponin (10 ml PBS + saponin per 50 ml culture) for 15 min at 4 °C. Parasites were pelleted at 15,500g for 3 min, resuspended in PBS, pelleted again at 15,500g for 3 min, and then frozen at -80 °C for days to weeks until the day of the assay. Eventually, pellets were thawed and solubilized in PBS + 0.5% Triton X-100, sonicated to disrupt the pellet, and then incubated at 4 °C for 1 h. Lysates were then incubated for 1 h at 4 °C with Dynabeads-Protein A or Dynabeads-Protein G (Thermo Fisher Scientific) that had been bound to anti-GFP (clone 3E6; Thermo Fisher Scientific). Beads were collected on a magnet stand, washed five times in cold PBS + 0.02% Tween-20, and then incubated directly with either the fluorogenic PEXEL peptide DABCYL-LNKKLLHETQ-EDANS, or the mutant peptide DABCYL-LNKKLAHETQ-EDANS at 1 μM in PM V activity buffer (50 mM Tris–malate, 50 mM NaCl, 0.05% Triton X-100, pH 6.5 unless otherwise noted) for 12 h at 37 °C. Total fluorescence was read on a multimodal plate reader (PerkinElmer Envision). The pH profiling assay was performed as aforementioned, but with PM V activity buffer brought to pH 3 with hydrochloric acid, or brought up to pH 4, 5, 6, 7, 8, or 9 with sodium hydroxide.

Immunofluorescence

For immunofluorescence assays, ring-stage parasites (15–20 h postinvasion) were smeared on glass slides and fixed for 10 s in the methanol-based fixative from Hemacolor Stain Set (Harleco; catalog no.: 65044), then permeabilized for 10 min in PBS + 0.01% Triton X-100, blocked in PBS + 3% bovine serum albumin overnight at 4 °C, and stained with primary antibody and secondary antibody solutions for 1 h each at room temperature. Primary antibodies were rat anti-FLAG (clone L5; Novus Biologicals; catalog no.: 63730L, used at 1:500 dilution in blocking buffer) and rabbit anti-GFP (Abcam; catalog no.: ab6556; used at 1:500 dilution in blocking buffer); secondary antibodies were goat anti-rabbit-AlexaFluor488 and goat antirat-AlexaFluor555 diluted 1:2000.
Plasmin V nepenthesin loop

dilution in blocking buffer. Coverslips were mounted in Pro-Long Gold antifade with 4',6-diamidino-2-phenylindole (Life Technologies; catalog no.: P-36931). Images were taken on a Zeiss AxioImager.M1 epifluorescence microscope with a Hamamatsu ORCA-ER CCD camera and AxioVision, version 4.8.1 (Zeiss). Images were cropped and brightness/contrast adjusted for viewing in Zen Blue (Zeiss). Pearson’s correlation coefficient was calculated using the FIJI (version 1.53q) (40) plugin Coloc2 (version 3.0.5), and data were graphed using GraphPad Prism 9.

Mammalian cell expression

To express PM V in mammalian cells, we transfected human embryonic kidney 293-F cells (Thermo Fisher Scientific) using a modified version of the protocol (39). Briefly, human embryonic kidney 293-F cells were cultured in FreeStyle 293 Expression Medium (Thermo Fisher Scientific). A day before transfection, cells were seeded at 6 × 10^5 cells/ml, 100 ml/transfection. The next day, with cells ~10^6 cells/ml, transfection was performed by combining 300 μl polyethylamine (25 kDa branched) with 15 ml FreeStyle 293 Expression Medium, slowly adding 100 μg of the plasmid of interest, allowing to sit for 15 min at room temperature, then adding all to the flask of cells. Cells were incubated shaking at 37 °C for 48 h, and then supernatants were collected and frozen at −80 °C for up to a week.

On the day of purification, supernatants were thawed and bound to a column of nickel agarose beads (GoldBio), washed five times in wash buffer (20 mM imidazole, 20 mM H_2NaO_4P·2H_2O, 500 mM NaCl, and pH 7.4), and eluted three times in elution buffer (500 mM imidazole, 20 mM H_2NaO_4P·2H_2O, 500 mM NaCl, pH 7.4). Elutions were then prepared for Western blot as aforementioned (see “Western blotting”) or used directly for activity assay as aforementioned (see “Peptide assays”).

Molecular dynamics simulations

All simulations were prepared using Gromacs (https://github.com/gromacs/gromacs) 2020 (41). Initial structures were placed in a dodecahedral box extending 1.0 nm beyond the protein in all dimensions. These systems were subsequently explicitly solvated in water (TIP3P) along with 0.1 M NaCl (42). Systems were energy minimized with a steepest descents algorithm until the maximum force fell below 100 kJ mol^−1 nm^−1, using a step size of 0.01 nm and a cutoff distance of 1.2 nm for the neighbor list, Coulomb interactions, and van der Waals interactions. All simulations used the AMBER03 force field.

Systems were subsequently run for a total simulation time of 1 μs each using a 4 fs time step using virtual sites. Cutoffs of 1.1 nm were used for the neighbor list, with 0.9 for Coulomb and van der Waals interactions. The Verlet cutoff scheme for neighbor lists was used. Long-range interactions were calculated using the particle mesh Ewald method with a Fourier spacing of 0.12 nm. The system temperature was held constant at 300 K using the stochastic velocity rescaling (v-rescale) thermostat, and pressure was coupled with Parinello–Rahman barostat (43). Seven independent simulations were run for each mutant: WT PM V, Nep^NeplA, Nep^Cl3S, and Nep^C2,4S.

Each PM V mutant was clustered independently using RMSD. Clustering was performed using the k-hybrid clustering algorithm to 3000 cluster centers. MSMs (a network representation of the free-energy landscape (44, 45)) were built using our python package, Enspara (https://github.com/bowman-lab/enspara) (45). To estimate the errors associated with our model, we iteratively built MSMs, each time leaving one of our trajectories. Displayed quantifications are the mean of the seven MSMs and ± standard deviation of these MSMs. Transition probability matrices were produced by counting the transitions between each state of the cluster space, adding a prior count of 1/n where n is the total number of states, and row-normalizing as is common (45, 46). Equilibrium populations were calculated as the eigenvector of the transition probability matrix with an eigenvalue of 1. All subsequent analysis was performed using the built-in tools of MDTraj (https://github.com/mdtraj/mdtraj) (46), and graphs were produced in Matplotlib (https://matplotlib.org/) (47).

Data availability

The computational data generated for this article are publicly available via Open Science Framework: https://osf.io/5txud/. Scripts used for the analysis of these data have been made public on GitHub (https://github.com/shbhakat/PMV). GROMACS (https://github.com/gromacs/gromacs), MDTraj (https://github.com/mdtraj/mdtraj), and our Enspara code (https://github.com/bowman-lab/enspara) are all open source.

Supporting information—This article contains supporting information.

Acknowledgments—We thank members of the Goldberg laboratory for invaluable suggestions over the course of this project, particularly Eva Istvan for sharing expertise in protein biochemistry, and Barbara Vaupel for assistance making plasmids.

Author contributions—A. J. P. and D. E. G. conceptualization; A. J. P., J. J. M., Soumendranath Bhakat, S. M., and Suhas Bobba investigation; D. E. G. writing–review & editing; A. J. P., J. J. M., and Soumadranath Bhakat writing—original draft; G. R. B. and D. E. G. supervision; A. J. P. and D. E. G. funding acquisition.

Funding and additional information—This work was supported by the National Institute of Allergy and Infection Diseases (grant no.: RO1 AI047798; awarded to D. E. G.). A. J. P. was supported by an American Heart Association Predoctoral Fellowship (grant no.: 18PRE33960417). J. J. M. is supported by a National Institute on Aging training grant (grant no.: T32AG058518).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: aTc, anhydrotetracycline; CD, catalytic dead; ER, endoplasmic reticulum; MSM, Markov transition probability matrix; PM, plasminogen activator

Supporting information

This article contains supporting information.

Acknowledgments

We thank members of the Goldberg laboratory for invaluable suggestions over the course of this project, particularly Eva Istvan for sharing expertise in protein biochemistry, and Barbara Vaupel for assistance making plasmids.

Author contributions

A. J. P. and D. E. G. conceptualization; A. J. P., J. J. M., Soumendranath Bhakat, S. M., and Suhas Bobba investigation; D. E. G. writing–review & editing; A. J. P., J. J. M., and Soumadranath Bhakat writing—original draft; G. R. B. and D. E. G. supervision; A. J. P. and D. E. G. funding acquisition.

Funding and additional information

This work was supported by the National Institute of Allergy and Infection Diseases (grant no.: RO1 AI047798; awarded to D. E. G.). A. J. P. was supported by an American Heart Association Predoctoral Fellowship (grant no.: 18PRE33960417). J. J. M. is supported by a National Institute on Aging training grant (grant no.: T32AG058518).

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations

The abbreviations used are: aTc, anhydrotetracycline; CD, catalytic dead; ER, endoplasmic reticulum; MSM, Markov transition probability matrix; PM, plasminogen activator
state model; PEXEL, *Plasmodium* export element; PM V, plasmepsin V; RBC, red blood cell; TetR, Tet repressor.

References

1. Spillman, N. J., Beck, J. R., and Goldberg, D. E. (2015) Protein export into malaria parasite–infected erythrocytes: mechanisms and functional consequences. *Annu. Rev. Biochem.* 84, 813–841

2. Matthews, K. M., Pitman, E. L., and de Koning-Ward, T. F. (2019) Illuminating how malaria parasites export proteins into host erythrocytes. *Cell. Microbiol.* 21, e13009

3. Hiller, N. L., Bhattacharjee, S., van Ooj, C., Loliios, K., Harrison, T., Lopez-Estraño, C., et al. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 306, 1934–1937

4. Marti, M., Good, R. T., Rug, M., Knuepfer, E., and Cowman, A. F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* 306, 1930–1933

5. Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., and Goldberg, D. E. (2010) Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. *Nature* 463, 632–636

6. Boddey, J. A., Hodder, A. N., Günther, S., Gilson, P. R., Patsiouras, H., Kapp, E. A., et al. (2018) EXP2 is a nutrient-permeable channel in the malaria parasite. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10034–10039

7. Kojima, M., Kubomura-Yoshida, N., Marapana, D., M. T., Oksman, A., and Beck, J. R. (2018) EXP2 is a nutrient-permeable channel in the *Plasmodium* falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods* 3, 615–621

8. Bowman, G. R., Pande, V. S., Noé, F., eds. (2014) *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*. Advances in Experimental Medicine and Biology. Springer Netherlands, Dordrecht

9. Park, Y. N., Aikawa, J. I., Nishiyama, M., Horinouchi, S., and Beppu, T. (1997) Site-directed mutagenesis of conserved Trp39 in rhizomucor pellius pepsin: possible role of Trp39 in maintaining Tyr75 in the correct orientation for maximizing catalytic activity. *J. Biochem.* 121, 118–121

10. Andreeva, N. S., and Rumsch, L. D. (2001) Analysis of crystal structures of aspartic proteinases: on the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. *Protein Sci.* 10, 2439

11. Athauda, S. B. P., Matsumoto, K., Rajapakshe, S., Kuribayashi, M., Kojima, M., Kubomura-Yoshida, N., et al. (2015) An aspartyl protease directs malaria effector protein families from the Plasmepsin falciparum exportome. *Traf. J.* 14, 532–550

12. Goldberg, D. E. (2015) Plasmepsin V shows its carnivorous side. *Nat. Struct. Mol. Biol.* 22, 647–648

13. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M. (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEBS Microbiol. Rev.* 29, 231–262

14. Silva, A. M., Lee, A. Y., Gulkis, S. V., Maier, P., Collins, J., Bhat, T. N., et al. (1996) Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10034–10039

15. Rawlings, N. D., Waller, M., Barrett, A. J., and Bateman, A. (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 42, D503–D509

16. Loymundong, C., Sriraluk, P., Songtawee, N., Wongpanya, R., and Boonyala, N. (2019) Yield improvement and enzymatic dissection of Plasmodium falciparum plasmepsin V. *Mol. Biochem. Parasitol.* 231, 111188

17. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., et al. (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* 419, 498–511

18. Singh, G. P., Chandra, B. R., Bhattacharya, A., Akhoury, R. R., Singh, S. K., and Sharma, A. (2004) Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in Plasmodium falciparum. *Mol. Biochem. Parasitol.* 137, 307–319

19. Muralidharan, V., Oksman, A., Iwamoto, M., Wandless, T. J., and Goldberg, D. E. (2011) Asparagine repeat function in a Plasmodium falciparum protease assessed via a regulatable fluorescent affinity tag. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4411–4416

20. Muralidharan, V., Oksman, A., Pal, P., Lindquist, S., and Goldberg, D. E. (2012) Plasmodium falciparum heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fever. *Nat. Commun.* 3, 1310

21. Bobrows, R., Iaudzens, K., and Jirgensons, A. (2019) Exploiting structural dynamics to design open-flap inhibitors of malarial aspartic proteinases. *J. Med. Chem.* 62, 8931–8950

22. Bhakat, S. (2021) Pepsin-like aspartic proteinases (PAPs) as model systems for combining biomolecular simulation with biophysical experiments. *RSC Adv.* 11, 11026–11047

23. Aurrecoechea, C., Brestelli, J., Brunke, B. P., Dommer, J., Fischer, S., Gajria, B., et al. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res.* 37, D539–D543

24. Madeira, F., Park, Y. M., Lee, J., Buno, N., Gu, T., Madhusoodanan, N., et al. (2019) The EMBL–EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641

25. Spillman, N. J., Beck, J. R., Ganesan, S. N., Niles, J. C., and Goldberg, D. E. (2017) The chaperonin TrIC forms an oligomeric complex in the malaria parasite cytosol. *Cell. Microbiol.* 19, e12719

26. Saraf, M., Niles, J. C., Zingerberg, J., Goldberg, D. E., and Beck, J. R. (2018) EXP2 is a nutrient-permeable channel in the...
vacuolar membrane of Plasmodium and is essential for protein export via PTEX. *Nat. Microbiol.* 3, 1090–1098

39. Aricescu, A. R., Lu, W., and Jones, E. Y. (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1243–1250

40. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682

41. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25

42. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1998) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926

43. Russi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126, 014101

44. Wang, X., Unarta, I. C., Cheung, P. P. H., and Huang, X. (2021) Elucidating molecular mechanisms of functional conformational changes of proteins via Markov state models. *Curr. Opin. Struct. Biol.* 67, 69–77

45. Husic, B. E., and Pande, V. S. (2018) Markov state models: from an art to a science. *J. Am. Chem. Soc.* 140, 2386–2396

46. McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., *et al.* (2015) MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* 109, 1528–1532

47. Hunter, J. D. (2007) Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95