Role of dynamic cooperativity in the mechanism of hexameric M17 aminopeptidases

Nyssa Drinkwater [a,1,2], Wei Yang [a,1], Blake T. Riley [b], Komagal Kannan Sivaraman [a], Itamar Kass [b,c], Ashley M. Buckle [b], Sheena McGowan [a, 2]

[a] Biomedicine Discovery Institute, Department of Microbiology, Monash University, Clayton Melbourne, VIC 3800, Australia

[b] Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Clayton Melbourne, VIC 3800, Australia

[c] Victorian Life Sciences Computation Centre, Monash University, Clayton 3800, Victoria, Australia

[1] N.D. and W.Y. contributed equally to this work.

[2] To whom correspondence should be addressed: nyssa.drinkwater@monash.edu; +613 99029371 sheena.mcgowan@monash.edu; +613 990293009

Abstract

M17 aminopeptidases possess a conserved hexameric arrangement throughout all kingdoms of life. Of particular interest is the M17 from *Plasmodium falciparum* (*Pf*A-M17), which is a validated antimalarial drug target. Herein we have examined *Pf*A-M17 using an integrated structural biology and biochemical approach to provide the first description of the fundamental role of oligomerisation. We found that, rather than operating as discrete units, the active sites of the *Pf*A-M17 hexamer are linked by a dynamic loop, which operates cooperatively to regulate activity. Further, we characterised motions in key surface loops that moderate access to the central catalytic cavity. Based on our new understanding of the dynamics inherent to *Pf*A-M17, we propose a novel mechanism that would allow exquisite control of enzyme function in response to cellular signals, and go on to discuss how, through divergent evolution, this mechanism might have developed to moderate key differences in M17 function across species.

Introduction:

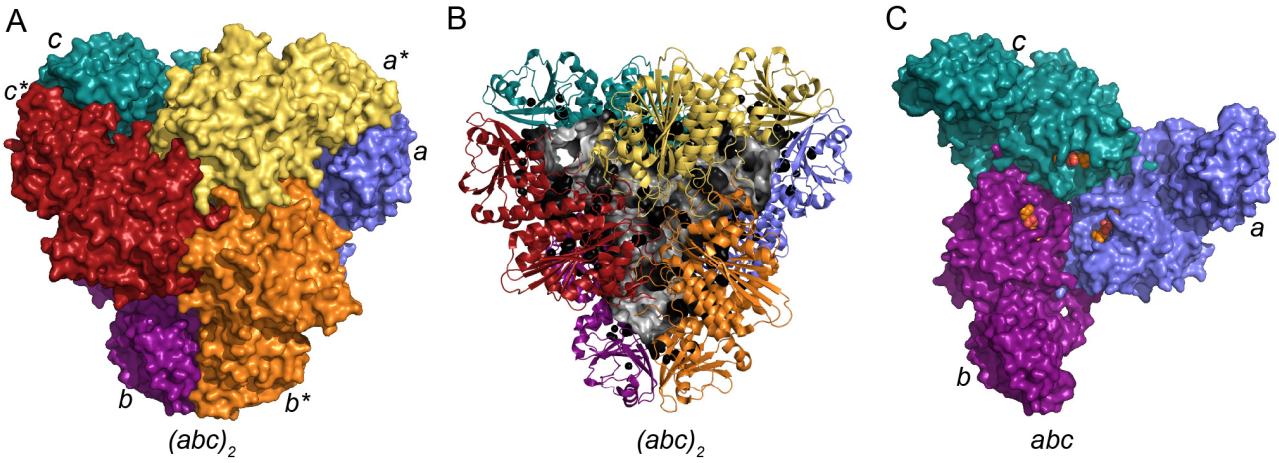
Our appreciation of protein flexibility and dynamics has grown remarkably from the early static lock-and-key model, to our current understanding that proteins are dynamic entities capable of extreme flexibility. Dynamics are now recognised as integral to protein function and regulation, and therefore fundamental to life. For enzymes, flexibility influences substrate binding (Vögeli, Bibow, & Chi, 2016; Wurm, Holdermann, Overbeck, Mayer, & Sprangers, 2017) and reaction mechanism (Bhabha et al., 2011; McElheny, Schnell, Lansing, Dyson, & Wright, 2005), allowing precise control of reactions that if left unregulated, could be harmful to the cell. Despite the clear importance of enzyme flexibility, there are few cases where dynamics, and how those dynamics link to function, are truly understood on an atomic level.

Intracellular proteolysis requires precise spatial and temporal control to prevent cleavage of proteins not destined for destruction. For this purpose, high-molecular weight protease enzymes are self-compartmentalised, whereby the active sites are enclosed in inner cavities isolated from the

cellular environment. Such an arrangement is often mediated by multimeric self-association, such as that seen for the family of M17 aminopeptidases, otherwise known as PepA or LAP (leucyl aminopeptidases; Clan MF; M17 family) (MEROPS (Rawlings, Barrett, & Finn, 2016)), which are present in all kingdoms of life. Although overall sequence conservation is low, M17 aminopeptidases possess a conserved homohexameric structure wherein a dimer of trimers encloses an inner cavity harbouring the six active sites (Lowther & Matthews, 2002). The sequence and structure of the proteolytic sites themselves, as well as the reaction they catalyse, are highly conserved, all utilising two divalent metal ion cofactors to catalyse the removal of select N-terminal amino acids from short peptide chains (Lowther & Matthews, 2002). This reaction contributes to intracellular protein turnover, a fundamental housekeeping process across all living organisms (Matsui, Fowler, & Walling, 2006). However, a wide range of additional functions beyond aminopeptidase activity have also been attributed to M17 family members. M17 proteases have been shown in plants to function as molecular chaperones to control stress-induced damage (Kumar, Kaur, Chattopadhyay, & Bachhawat, 2015; Scranton, Yee, Park, & Walling, 2012), while in bacteria they form hetero-oligomers (Minh, Devroede, Massant, Maes, & Charlier, 2009; Sträter, Sherratt, & Colloms, 1999) and contribute to site-specific DNA recombination (Alén, Sherratt, & Colloms, 1997; Stirling, Colloms, Collins, Szatmari, & Sherratt, 1989) and transcriptional control (Charlier et al., 1995). Therefore, although the family of M17 aminopeptidases have a highly conserved structure across different organisms, they are multifunctional, capable of performing diverse organism-specific functions far beyond peptide hydrolysis. Our understanding of mammalian M17 aminopeptidases is based on early investigations into the structure and catalytic mechanism of the bovine lens M17 aminopeptidase, which while considerable, consider the six chains within the hexamer as discrete entities. Further, although some flexibility within the active site has been considered (Schurer, Horn, Gedeck, & Clark, 2002), the role/s of the conserved hexameric arrangement of M17 aminopeptidases, as well as motions within this arrangement, have thus far not been investigated.

70 Of particular interest is the M17 aminopeptidase from *Plasmodium falciparum* (*Pf*A-M17),
71 the major causative agent of malaria in humans. *Pf*A-M17 has a proposed role in haemoglobin
72 digestion, the essential process by which parasites break down human haemoglobin to yield the free
73 amino acids necessary for growth and development (Rosenthal, 2002; Stack et al., 2007). *Pf*A-M17
74 is essential for the blood stage of the parasite life cycle (Dalal & Klemba, 2007) and inhibition of
75 *Pf*A-M17 results in parasite death both *in vitro* (Harbut et al., 2011) and *in vivo* (Skinner-Adams et
76 al., 2007). However, this effect occurs early in the life cycle before haemoglobin digestion is
77 thought to have initiated, which has led to speculation that *Pf*A-M17 possesses additional unknown
78 function/s (Harbut et al., 2011). Irrespectively, *Pf*A-M17 is an exciting target for the development
79 of novel antimalarial therapeutics (Drinkwater, Bamert, Sivaraman, Paiardini, & McGowan, 2015;
80 Drinkwater et al., 2016; Mistry et al., 2014; Skinner-Adams et al., 2007; Skinner-Adams et al.,
81 2012). The crystal structure of *Pf*A-M17 shows the homohexameric arrangement characteristic of
82 M17 aminopeptidase enzymes (Fig. 1A), with the six active sites orientated inwards and accessible
83 to the central cavity (Fig. 1B and C) (McGowan et al., 2010). We were interested to determine the
84 mechanistic basis of the conserved hexameric assembly by investigating the dynamics inherent to
85 *Pf*A-M17 and probing how they contribute to function.
86

Figure 1. *Pf*A-M17 is a homo hexamer composed of a dimer of trimers. (A) Three chains, *a* (blue), *b* (purple), and *c* (teal) form a 'trimer' (*abc*), which interacts with an identical trimer made up of *a** (yellow), *b** (orange), and *c** (red) to form the hexamer $(abc)_2$. (B) *Pf*A-M17 (cartoon) possesses a large internal cavity (grey surface representation) that contains the six active sites. (C) A single trimeric face (*abc*) shown in the same orientation as A (front trimer hidden). Orange spheres substrate-analogue bestatin bound in the active sites, which are exposed to the inner cavity.



96 **Results and Discussion:**

97 **Does hexamerisation play a functional role?**

98 The hexameric arrangement of M17 aminopeptidases is highly conserved, however, the
99 contribution of this assembly to enzyme function is unknown. The arrangement is characterised as a
100 dimer of trimers (Burley, David, Taylor, & Lipscomb, 1990); three chains (*a*, *b*, and *c*, Fig. 1A)
101 interact via their C-terminal catalytic domains to form a 'trimer' (*abc*), two of which (*abc* and *abc**)
102 then associate to form the hexamer, $(abc)_2$ (Fig. 1A). The result is a large central cavity containing
103 the six active sites arranged in two symmetrical disk-like arrangements (Fig. 1B and 1C). Despite
104 the proximity of the active sites, both unliganded and liganded crystal structures of *Pf*A-M17 show
105 they exist as six discrete units (McGowan et al., 2010). Each of the active sites is composed entirely
106 of residues from a single chain and contains its own catalytic machinery that includes two zinc ions
107 and a carbonate ion. Based on the current evidence, hexamerisation does not have a clear role in the
108 catalysis. However, the M17 aminopeptidase from *H. pylori* (*Hp*-M17) exhibits positive
109 cooperativity, which suggests that communication between the *Hp*-M17 active sites does occur
110 (Dong et al., 2005). Therefore, to probe the role of *Pf*A-M17 hexamerisation in catalysis, we first
111 attempted to identify whether any cooperativity between the *Pf*A-M17 active sites exists, or if the
112 sites operate as discrete units. We employed a fluorescence-based aminopeptidase activity assay in
113 substrate saturation experiments, wherein we assessed the relationship between substrate
114 concentration and reaction velocity and analysed the Hill coefficient (SI 1). In contrast to the results
115 of *Hp*-M17, we were unable to detect any evidence of cooperativity during *Pf*A-M17 catalysis (Hill
116 coefficient, $n_H = 1.0$), which suggests that although both *Hp*-M17 and *Pf*A-M17 share a conserved
117 overall assembly and highly conserved active site structures, they process substrates differently.
118 This result is in contrast to the major assumption currently made within the M17 aminopeptidase
119 literature, that the catalytic mechanism of the enzymes are conserved, and operate as has been
120 described for the M17 enzyme from bovine lens (Lowther & Matthews, 2002).

121

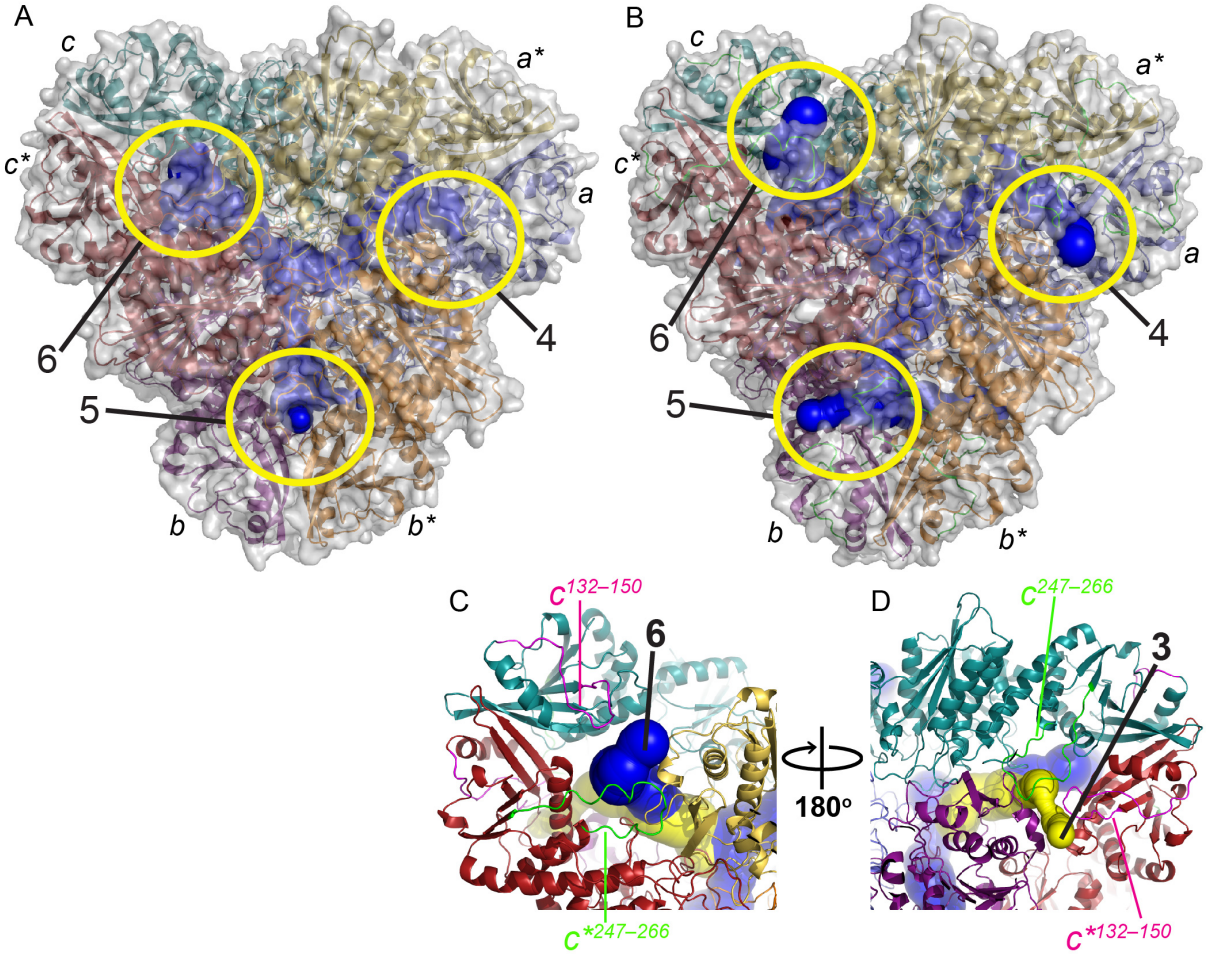
122 **Gatekeeper loops mediate access to central catalytic cavity**

123 In the absence of evidence showing catalytic cooperativity between the active sites of the
124 *Pf*A-M17 hexamer, we hypothesised that oligomerisation plays a regulatory role. To examine this
125 theory, we turned to all-atom molecular dynamics (MD) simulations to determine the mechanism
126 by which protein dynamics might be moderating *Pf*A-M17 activity. We performed all-atom MD
127 simulations of hexameric *Pf*A-M17 (3 × 400 ns). Analysis of the root mean square deviation
128 (RMSD) over the course of the simulation indicated that the hexamer did not undergo any large
129 conformational changes, nor rigid body movements (average RMSD of Cα atoms = 2.5 ± 0.04 Å,
130 SI 2A). We performed a principal component analysis (PCA) of the simulations on the backbone
131 and zinc atoms of hexameric *Pf*A-M17, and observed that the top PC, PC1, accounts for 62 % of the
132 total variance, while PC2 accounts for only 10 %. Projecting the trajectories onto the top two PCs
133 showed the major motion, described by PC1, is an expansion of the hexamer from 120 Å to 127 Å
134 (average of three measurements between Cα of Asn181^{a}, Asn181^{b} and Asn181^{c}) (SI 2B). The
135 expansion is ~ 6%, and likely results primarily from the release of crystal constraints (Gerstein &
136 Chothia, 1996).

137 The crystal structure of *Pf*A-M17 showed six channels connecting the central catalytic
138 cavity to the protein surface, and identified a ~20 Å flexible loop (residues 246–265) that sits at the
139 entrance to the channels (McGowan et al., 2010). The complete loop was able to be modelled in
140 only one of the six chains, where it occluded the channel entrance, suggesting that the loops may
141 regulate substrate ingress and/or product egress (McGowan et al., 2010). The putative access
142 channels, flexible loops, and any motions they might undergo, were therefore of great interest to us.
143 Initial modelling and energy minimisation of the starting loop conformations resulted in almost
144 complete occlusion of the channel entrances (diameters ~4–5 Å across the bottleneck, Fig 2A, SI 3),
145 in line with the conformation modelled for the single complete chain in the original crystal
146 structure. Over the course of the MD simulation, the diameter of the pores increased, opening some
147 of the channels to external solvent (Fig. 2B). Pore increases were partially mediated by the flexible

148 loop identified in the crystal structure (residues 246–265), but also by a second loop on the other
149 side of the channel, which is contributed from the opposing trimer in the hexamer (residues 132–
150 150). The 'gatekeeper loops' underwent a range of motions, affecting the size of the channels to
151 varying degrees. While one channel increased in diameter throughout the simulation (channel 6,
152 increased from 4.9 Å to 8.9 Å, Δ4.0 Å, Fig. 2C, SI 3), another remained occluded throughout the
153 simulation (channel 3 starting diameter was 4 Å compared to 4.5 Å at completion, Δ0.5 Å, Fig. 2D,
154 SI 3). Intriguingly, these two pores that sampled the most extreme changes (Δ0.5 Å in channel 3
155 versus Δ4.0 Å in channel 6) are in close proximity, on opposite sides of the same dimer pair
156 (channel 3 is formed by $c^{247\text{-}266}$ and $c*^{132\text{-}150}$, channel 6 by $c^{132\text{-}150}$ and $c*^{247\text{-}266}$, Fig. 2C and 2D),
157 which suggests that the channels are linked and operating in opposing concert. Overall, our analysis
158 of the putative access channels by MD demonstrates that channel opening, and consequently access
159 to the central catalytic cavity, may be mediated by motions in these key gatekeeper loops. Further,
160 each channel entrance is controlled by the loops of neighbouring chains of the hexamer, which
161 presents the first evidence that hexamerisation plays a role in moderating access to the catalytic
162 core. By regulating the size and/or nature of the access channels to the central cavity,
163 hexamerisation might also play an important role in substrate discrimination by mediating the type
164 of substrate allowed access to the inner cavity. Such a role has key implications for understanding
165 the wide range of different functionalities observed throughout the family of M17 aminopeptidases.
166
167

Figure 2. Flexible loops moderate channel size and access to catalytic cavity of hexameric *Pf*A-M17. (A) *Pf*A-M17 conformation at the start of the MD simulation shown in surface representation (transparent grey) and cartoon (*a* in blue, *b* in purple, *c* in teal, *a** in yellow, *b** in orange, *c** in red). Access of the six channels (blue) to external solvent is blocked by gatekeeper loops at the entrance to the channels. Yellow circles indicate the occluded channels (4, 5, and 6) on the front face of the hexamer. (B) During the simulation, the six access channels become accessible, creating a range of pore access sizes from 'open' (8.9 Å, channel 6, $c^{132–150}$:$c*^{247–266}$) shown in (C) to 'closed' (3.9 Å, channel 3, $c^{247–266}$:$c*^{132–150}$) shown in (D).



180 **Hexamerisation is essential for catalytic function**

181 Our MD analysis suggests that hexamerisation serves a regulatory role by mediating ingress
182 and/or egress routes to the central catalytic cavity via the channels described above. If this is the
183 sole purpose of the hexameric arrangement, then oligomerisation would be expected to control
184 reaction rate or specificity, not the ability of the enzyme to perform the hydrolysis reaction itself.
185 We therefore questioned whether trimeric, or indeed monomeric, *Pf*A-M17 would possess
186 aminopeptidase activity. Examination of the *Pf*A-M17 crystal structure identified Trp525 and
187 Tyr533 as key residues that mediate association of *abc* and *abc** into $(abc)_2$ and from this, we
188 hypothesised that their elimination would ablate hexamerisation. *Pf*A-M17(W525A+Y533A)
189 expressed and purified similarly to the wild type enzyme, however, analysis by analytical gel
190 filtration chromatography showed that the mutated enzyme was monomeric in solution (SI 4A).
191 Examination of *Pf*A-M17(W525A+Y533A) activity showed that the monomeric enzyme is
192 inactive, even at high concentrations (SI 4B). It is therefore evident that hexamer formation is
193 essential for *Pf*A-M17 proteolytic activity, and that oligomerisation plays an important functional
194 role beyond regulation of substrate ingress and/or product egress.

195

196 **Hexameric assembly stabilises and links the *Pf*A-M17 active sites**

197 Having established that monomeric *Pf*A-M7 is unable to function independently, we
198 continued to search for the key functional role of hexamerisation. We performed MD simulations of
199 the monomeric enzyme and compared the output to the simulation of hexameric *Pf*A-M17. Over the
200 time course sampled, the monomer trajectories underwent movements equivalent to an average
201 RMSD of 4.1 ± 0.5 Å, demonstrating that the monomer underwent greater structural changes than
202 the hexameric enzyme (RMSD of 2.5 ± 0.04 Å). Further, in the simulation of monomeric *Pf*A-M17,
203 variation was observed between triplicate runs (SI 5A). This is in contrast to the hexamer
204 simulation where all three runs showed a similar RMSD profile (SI 2A). This difference in the
205 dynamics of hexameric versus monomeric *Pf*A-M17 supports an assertion that the hexameric

206 assembly stabilises the monomer to enable catalysis. To determine how this stabilisation translates
207 to the catalytic machinery, we examined the local environment of the active site throughout the
208 simulation. The catalytic mechanism of M17 aminopeptidases relies on the deprotonation of a
209 catalytic water for nucleophilic attack on the peptide carbonyl (Schurer et al., 2002; Sträter &
210 Lipscomb, 1995). In the simulation of monomeric *Pf*A-M17, we did not observe a stable position
211 for a water molecule within the active site (SI 5D). In contrast, in the hexameric *Pf*A-M17
212 simulation, we observed a stable water conformation wherein it associates with the site 1 zinc ion
213 (SI 5E), consistent with the position of a nucleophilic water (Schurer et al., 2002; Yang et al.,
214 2017). An active site with increased mobility may compromise catalysis, which relies on precise
215 placement of chemical moieties. We therefore assessed the mobility of individual residues
216 throughout the simulations via calculation of the root mean square fluctuations (RMSF) of the Cα
217 atoms (SI 5B and C). The RMSF analysis showed clear differences between the two simulations,
218 particularly residues 385–391, which lie on a loop flanking the active site on the interior of the
219 central cavity. In the simulation of monomeric *Pf*A-M17, this loop exhibited a high degree of
220 flexibility (average RMSF over three runs for residues 385–391 = 4.5 ± 0.4 Å). However, in the
221 simulation of hexameric *Pf*A-M17, the loop was relatively stable in four chains (chain *a*, *a**, b* and
222 *c**, average RMSF = 1.6 – 2.1 Å) and only moderately flexible in two chains (chain *b* and *c*,
223 average RMSF *b* = 2.8 ± 0.4 and *c* = 3.1 ± 0.4 Å). These key differences between the dynamics of
224 monomeric and hexameric *Pf*A-M17 demonstrate that hexamerisation enables catalysis by shielding
225 the essential nucleophilic water, and further, stabilises key active site loop structures. Such a
226 mechanism of preserving active site stability by oligomerisation has been previously described, for
227 example, DHDPS, which utilises a tetrameric arrangement to stabilise the catalytic dimerization
228 interface for optimal catalytic efficiency (Reboul et al., 2012).

229 Early work based on static structures of M17 aminopeptidases, primarily the M17 from
230 bovine lens, suggested that the nucleophilic water is likely deprotonated by either a bound
231 carbonate molecule or a conserved active site lysine residue (Sträter & Lipscomb, 1995). Density

function theory was applied, which suggested that the energy barrier for the lysine to act as the catalytic base is prohibitively high, and therefore implicated the carbonate ion for this role (Zhu et al., 2012). Conversely, use of a semiempirical quantum mechanical/molecular mechanical hybrid approach supported the model in which the lysine residue is the most likely candidate as a catalytic base (Schurer et al., 2002; Zhu et al., 2012). Sequence alignment of bovine M17 and *Pf*A-M17 determined the equivalent lysine to be residue 386 (Lys386), which lies on the active site loop shown to fluctuate in the *Pf*A-M17 MD. In the crystal structure the loop lines the active site and extends into the solvent of the inner cavity. PCA showed that in PC2 of the *Pf*A-M17 hexamer simulation, the loop from one chain of each trimer b and c*($b^{383\text{-}397}$ and $c^{*383\text{-}397}$) moves away from its own active site, extends across the entrance of the pocket and stretches toward the active site of the neighbouring chain (Movie 1). In the most extreme conformation, observed in *c** (and *b* to a slightly lesser extent), the loop occludes the entrance to the pocket and makes contact with the equivalent loop of the neighbouring chain, which has not undergone the movement. Although the loop does show some movement in the remaining chains, the extreme motion is observed for one site per trimeric unit only (*b* and *c**). This analysis suggests communication between active sites within each of the trimeric faces of *Pf*A-M17, and prompted us to employ further methods to experimentally examine the atomic detail of alternative *Pf*A-M17 conformations.

Movie 1: Movie showing displacement along PC2 of *Pf*A-M17 hexamer simulation. *Pf*A-M17 trimer with greatest level of movement shown only (*abc**), with chain *a** in yellow, *b** in orange, *c** in red.

https://www.dropbox.com/s/ir8pzku1he4khwq/Movie1_2880x2160.10Mbps.mp4?dl=0

255 **Novel *Pf*A-M17 conformation captured by crystallography**

256 The MD simulations identified a flexible loop capable of linking one *Pf*A-M17 active site to
257 the neighbouring unit of the trimer. This is the first evidence that shows the active sites within the
258 trimer might be linked, and raises the possibility that *Pf*A-M17 might utilise cooperativity during
259 catalysis. If different conformations of *Pf*A-M17 do indeed exist throughout the catalytic reaction,
260 we rationalised that we might be able to capture a previously unobserved state in a novel crystal
261 form and determine its structure by X-ray crystallography. We therefore screened for different *Pf*A-
262 M17 crystallisation conditions and solved a novel structure to 2.3 Å by molecular replacement (SI
263 6). The structure consisted of two copies of the hexamer in the asymmetric unit, with the overall
264 quaternary structure similar to previous *Pf*A-M17 structures. However, the new structure shows a
265 vastly different active site arrangement to any previously observed M17 structure. In the previously
266 determined *Pf*A-M17 structure, herein referred to as the 'active' conformation, the flexible active
267 site loop (residues 379–391) lines the active site and extends into the solvent of the inner cavity. In
268 the new conformation, this loop is observed to cross the active site, completely occluding its
269 entrance (Fig 3). Further, the loop extends to the active site of the neighbouring chain in the trimer,
270 where it occupies the binding pocket with a key lysine residue (Lys386) (Fig 4, Movie 2). The same
271 ~13 Å loop rearrangement is observed for all binding pockets, resulting in a direct link between the
272 three active sites of each trimer: a^{386} inserts into the binding pocket of *b*, b^{386} inserts into the
273 binding pocket of *c*, c^{386} inserts into the binding pocket of *a* (equivalent movements repeated in
274 *abc**, Fig. 4C–D, Movie 2). To obtain the slack within the loop to adopt this extended, occluded
275 conformation, the secondary structure at both ends of the loop has been disrupted (Fig 3C). This
276 includes disruption of the α-helical structure of nine residues (392–401) and complete disruption of
277 a short beta strand (372–379) (Fig 3D, Fig 4).

278

Figure 3. Novel *Pf*A-M17 conformation shows the active site occluded by loop (A and B) Overlay of a single monomer of the 'occluded' *Pf*A-M17 (yellow cartoon) with bestatin-bound 'active' *Pf*A-M17. Bestatin binding mode (orange) indicates the active site. Yellow cartoon protruding through the surface shows the position of the loop in the 'occluded' conformation. The loop moves ~13 Å (yellow dash) to occlude the active site. (C) Overlay of 'active' (blue) and 'occluded' (grey and yellow) crystal structures of *Pf*A-M17. Orientation equivalent to B. Black dash indicates the ~13 Å movement of the loop (yellow), which disrupts an alpha helix and beta sheet. (D) inverted representation of panel C, wherein 'occluded' *Pf*A-M17 (grey) is shown in surface representation, with the loop (yellow) completely occluding the active site.

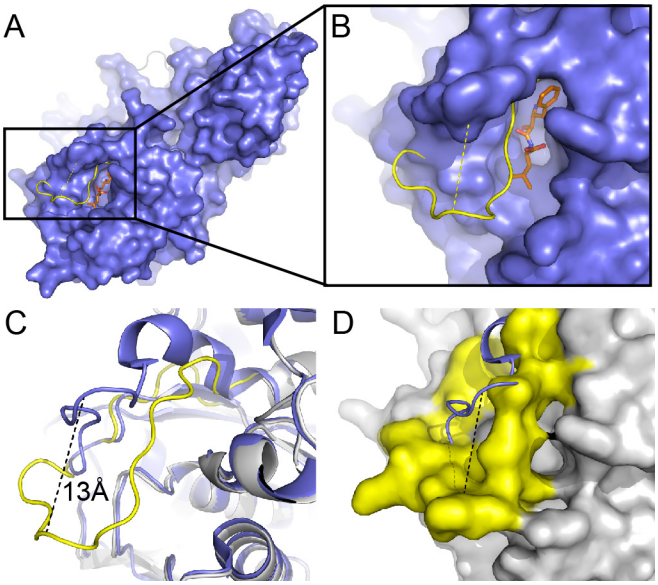
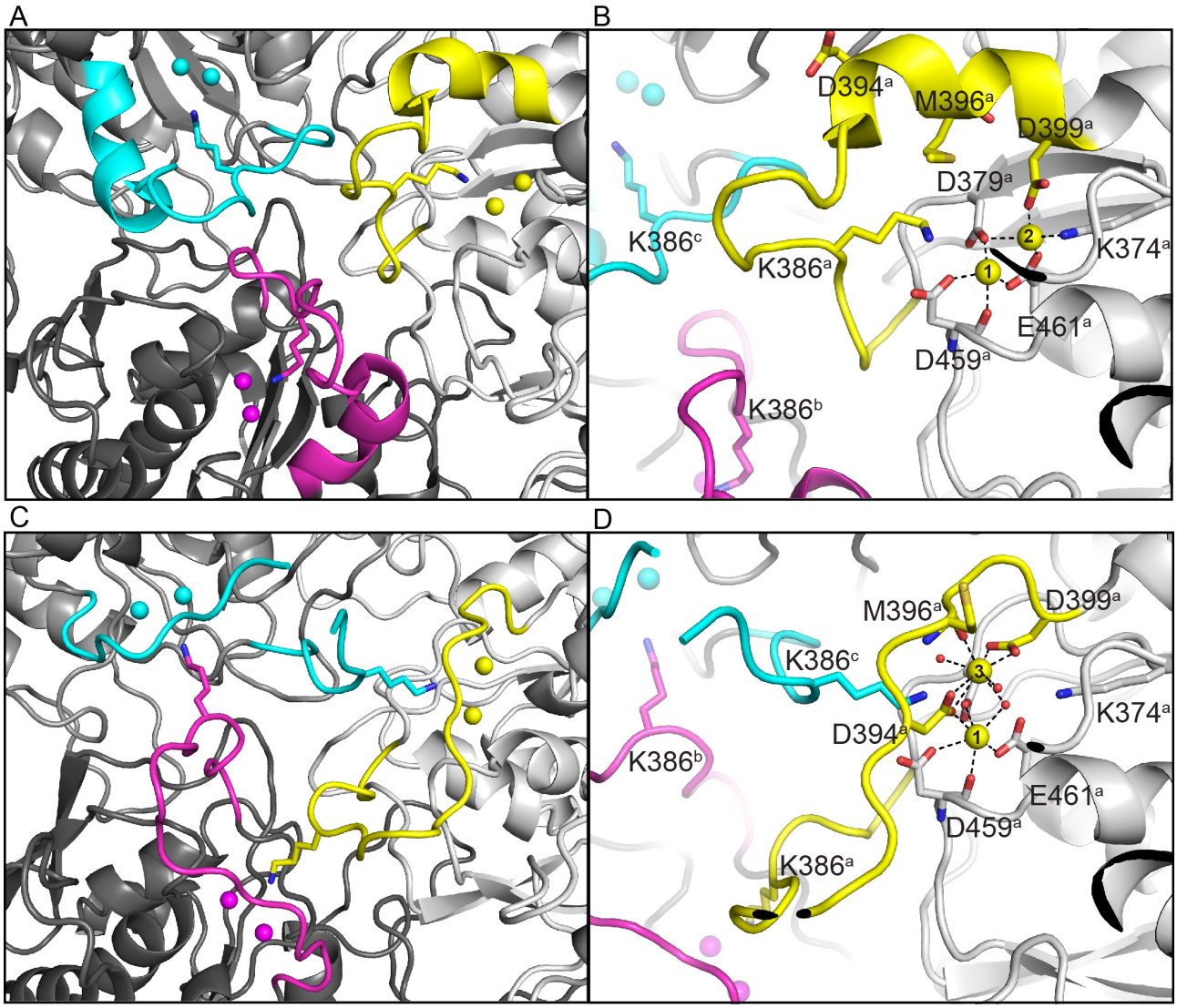


Figure 4. Conformational change from active to inactive *Pf*A-M17 mediated by a flexible regulatory loop. (A) Active site arrangement of active *Pf*A-M17 at the junction of the *abc* trimer; chain *a* (light grey, regulatory loop and zincs yellow), chain *b* (dark grey, regulatory loop and zincs magenta), chain *c* (medium grey, regulatory loop and zincs cyan). Lys386 of all chains shown in stick representation. (B) Zoom of chain *a* active site in active *Pf*A-M17. Orientation equivalent to panel A. Zinc positions 1 and 2 are and active site residues are indicated (subscript denotes chain identifier). (C) Active site arrangement of inactive *Pf*A-M17, showing that the active sites are linked by the regulatory loop. Coloring and orientation consistent with panel A. Lys386 sits in the active site of the neighbouring chain, thereby linking all three active sites within the trimer. (D) Zoom of chain *b* active site in inactive. Disruption of active site α-helix (residues 394–400) is observed, and re-organisation of the active site architecture, including zinc binding positions has occurred.



306 **Movie 2:** Movie showing morph between active and inactive *Pf*A-M17. Chain *a* in blue, *b* in

307 purple, *c* in teal, *a** in yellow, *b** in orange, and *c** in red.

308 https://www.dropbox.com/s/sl82io7vwur62bv/Movie2_2880x2160.10Mbps.mp4?dl=0
309

310 The active site rearrangement extends beyond protein backbone changes, and includes the
311 catalytic zinc ions absolutely required for enzyme activity. Based on kinetic and biophysical
312 characterisation, the two zinc sites of the M17 aminopeptidases have previously been termed site 1
313 and site 2, whereby site 1 is that closest to the mouth of the active site (Fig. 4B) (Lowther &
314 Matthews, 2002). Site 1 is readily lost from the active site, as observed by kinetic and
315 crystallographic studies, while site 2 is always occupied in structures, and its removal has been
316 reported to result in irreversible ablation of catalytic activity (Maric et al., 2009). Therefore, site 2 is
317 generally considered the 'tight' binding, catalytic site (Allen, Yamada, & Carpenter, 1983; Maric et
318 al., 2009; McGowan et al., 2010). In the novel conformation described here, we observed re-
319 arrangement of the zinc binding positions (Fig. 4B and 4D). While site 1 is occupied with a zinc
320 ion, the 'catalytic' zinc in site 2 is absent. Further, a zinc ion is bound in a third, previously
321 uncharacterised site, coordinated by the side chains of Asp394 and Asp399, and the main chain
322 oxygen of Met362, as well as two ordered water molecules, changing the coordination from
323 tetrahedral (Fig 4B) to octahedral (Fig 4D). Dialysis in mixed metal buffers (Zn^{2+} and Co^{2+}) has
324 previously shown that metal ion exchange between site 1 and 2 is also possible (Allen et al., 1983),
325 therefore re-arrangement of the active site zinc ions is not unprecedented, though the existence of a
326 third site was not considered. In the new conformation characterised here, the zinc coordination by
327 Asp394 is significant since this residue lies within the flexible loop (residues 375–401), and it is the
328 rotation of Asp394 from the external solvent into the active site that resulted in disruption of the
329 active site alpha helix. There is therefore a direct link between the third zinc binding site and the
330 flexible loop, which has substantial implications for catalysis. To investigate the potential impact of
331 the loop rearrangement on the catalytic activity of *Pf*A-M17, we superposed the previously
332 observed active *Pf*A-M17 conformation with the occluded conformation characterised here. The
333 overlay shows that in the occluded conformation, Lys386 occupies approximately the same space in
334 the neighbouring binding pocket that was previously occupied by the original Lys386 conformation
335 (Fig. 4B and 4D). Therefore, within the trimer, the two conformations are incompatible, and one

336 chain would not be expected to adopt the active conformation whilst its neighbour is in the
337 occluded conformation. Further, overlay of *Pf*A-M17 in complex with substrate-analog bestatin,
338 shows that the occluded loop conformation occupies the bestatin-binding space. Based on this, we
339 would not expect the occluded conformation to be capable of binding substrate. To confirm this
340 hypothesis, we attempted to soak bestatin into the occluded crystal form; bestatin was never
341 observed to bind. We therefore propose that the occluded conformation of *Pf*A-M17 is unable to
342 bind substrate, and conclude that we have captured an inactive conformation of *Pf*A-M17.

343 Importantly, the crystal structure of occluded *Pf*A-M17 is consistent with our MD
344 simulations, although the MD simulation showed the loop rearrangement in only two active sites
345 (one of each trimer). The major difference between the MD and crystallographic 'inactive'
346 conformations is the rearrangement of the catalytic zinc ions, which is a key component of the
347 crystallographic conformational change. However, the hybrid 'bonded / non-bonded approach' that
348 was used to model the metal centre in our simulations, required us to fix the positions of the zinc
349 ions (Yang et al., 2017). Since the novel zinc coordination site is directly linked to the flexible loop
350 through Asp394, and the MD simulations, which do not allow zinc rearrangement, show only a
351 partial loop conformational change, we suggest that zinc movement is necessary for the complete
352 rearrangement of the active conformation into the inactive conformation.

353

354 **The dynamic loop is key to *Pf*A-M17 catalysis**

355 Our MD and crystallographic studies identified a flexible loop in *Pf*A-M17, which has
356 important mechanistic implications. Therefore, to confirm that the characterised conformational
357 changes occur as part of the catalytic mechanism, we sought to (1) probe the role of key loop
358 residue, Lys386, (2) examine how loop flexibility affects catalysis, and (3), determine the functional
359 role of the third zinc binding site. For this purpose, we used site directed mutagenesis to alter key
360 residues, and measured their activity using our fluorescence-based activity assay.

361 Lys386 has a potential role in the *Pf*A-M17 catalytic mechanism, but is also a key player in
362 the identified loop re-arrangement. To probe the importance of Lys386 to catalysis, we generated
363 *Pf*A-M17(K386A). *Pf*A-M17(K386A) showed substantially reduced activity compared to the wild
364 type enzyme (Table 1). This retardation of activity is due to a decrease in the ability of the enzyme
365 to accelerate the reaction (decrease in k_{cat}), as opposed to reduced substrate binding (K_M is largely
366 unchanged). Although this result demonstrates that Lys386 is key to *Pf*A-M17 function, it does not
367 discriminate between a role in the catalytic reaction versus the loop rearrangement. To specifically
368 examine the role of loop conformational dynamics in catalysis, we introduced a proline in place of
369 Ala387, a loop residue that does not form interactions in either the active or occluded structures.
370 *Pf*A-M17(A387P), similarly to *Pf*A-M17(K386A), showed substantially reduced catalytic activity
371 (Table 1). Large concentrations of enzyme were required to measure enzyme activity, which
372 manifested as reduced product turnover (decreased k_{cat}). Therefore, flexibility of the loop that links
373 the active sites of *Pf*A-M17 is clearly important to catalytic function.
374 The presence of a third zinc binding site in *Pf*A-M17 was surprising since, to our
375 knowledge, an equivalent site has never been observed in any M17 aminopeptidase. We were
376 therefore curious to examine potential functional roles for the site. In the active conformation of
377 *Pf*A-M17, Asp394 has no clear role. While close to the active site, the side chain is directed to
378 solvent and makes no direct interactions. Therefore, we disrupted the ability of site 3 to coordinate
379 zinc with *Pf*A-M17(D394A). The effect of this mutation on enzyme activity was profound and
380 unexpected. Rather than a reduction in activity observed for the previous mutations, *Pf*A-
381 M17(D394A) showed greatly increased catalytic ability. Again, this resulted from an altered rate
382 (increased k_{cat}) rather than substrate binding affinity (K_M unchanged, Table 1). Therefore,
383 preventing the rearrangement of the catalytic zinc ions by removal of zinc binding site 3 greatly
384 increases the catalytic efficiency of *Pf*A-M17.
385

386 **Table 1.** Kinetic analysis of mutant *Pf*A-M17 enzymes.
387

	[E] (nM)	K_M ± S.E.M (μM)	k_{cat} ± S.E.M. (FU/min)
Wild type	150	22.8 ± 1.6	27 ± 1
*Pf*A-M17(K386A)	1000	21.6 ± 4.2	4.9 ± 0.3
*Pf*A-M17(A387P)	1000	34.8 ± 4.4	7.7 ± 0.3
*Pf*A-M17(D394A)	20	17.3 ± 0.2	164 ± 6

388

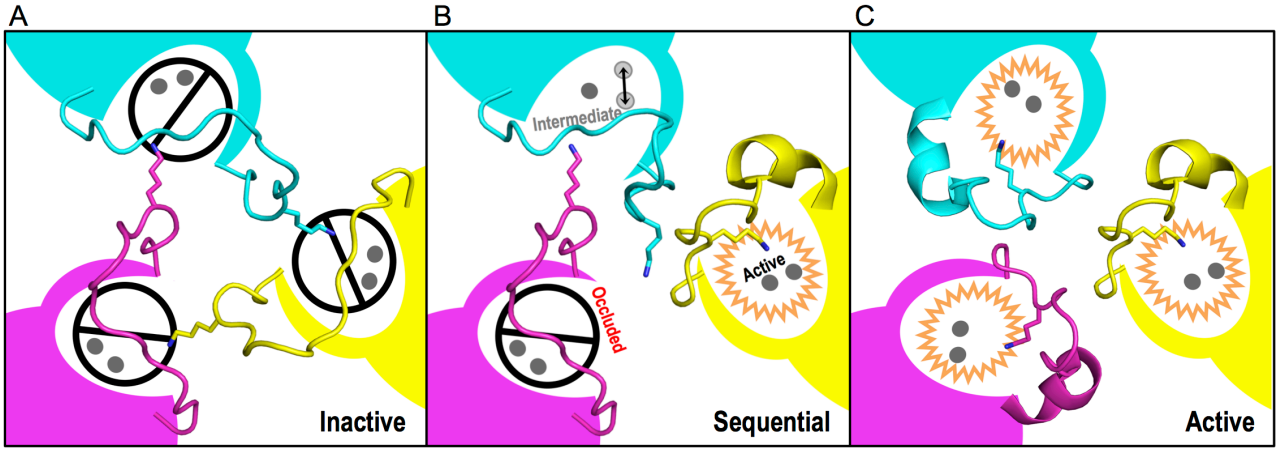
389 **A dynamic, cooperative mechanism for regulation of *Pf*A-M17 catalysis**

390 We have demonstrated that the characterised loop motion linking the *Pf*A-M17 active sites
391 has an important role in catalysis, but what is that role? For proteases, regulation of the catalytic
392 turnover rate is crucial, particularly under conditions in which substrate and/or product
393 concentrations are variable, or aberrant proteolytic activity is undesirable. The loop movement is
394 coupled with the rearrangement of the active site zinc ions, which we propose is key to the
395 mechanism. We therefore suggest that the inactive conformation we have characterised is the 'off'
396 switch in a dynamic regulatory mechanism. Based on current evidence, there are two potential
397 mechanisms by which the loop motion might regulate activity: (1) concerted activation, wherein all
398 chains transition between the active and inactive states concurrently, or (2) sequential activation,
399 where only one active site of each trimer is active at any one time, and turnover in one site triggers
400 the neighbouring site in a continual cycle. A mechanism of concerted activation is supported by our
401 crystal structures, which show that all six chains can adopt the active or inactive conformations
402 concurrently. Our MD simulations however, show only a single chain of each trimer initiating the
403 loop motion. Further, in a concerted activation mechanism, we would expect kinetic assays to show
404 cooperativity similarly to *Hp*-M17, which possesses a Hill coefficient of 2.3 (Dong et al., 2005).
405 Detection of cooperativity by the Hill coefficient relies on the assumption that ligand molecules
406 bind simultaneously; therefore, a Hill coefficient of 1.0 for *Pf*A-M17 supports a sequential
407 mechanism.

408 Our current data lend support to both concurrent and sequential activation mechanisms.
409 Although it is possible that we have not yet identified the specific conditions to differentiate
410 between the two mechanisms, we propose the more likely scenario, is that *Pf*A-M17 functions by a
411 combination of the two (Fig. 5). In this mechanism, the enzyme is capable of sampling three major
412 states: (1) an inactive state wherein all chains adopt the occluded conformation and no catalysis
413 occurs (Fig. 5A), (2) an active state which allows catalysis to occur in all active sites concurrently
414 (Fig. 5C), and (3) an intermediate, or sequential activation state, which is moderated by the

415 regulatory loop to allow proteolysis in only one active site per trimeric unit (Fig. 5B). This
416 'combined' mechanism would allow exceptionally fine control of catalytic rate when fast (active
417 state), moderate (sequential activation state), or no (inactive state) proteolysis is required. Dynamic
418 ensembles are known to play an important role in enzyme catalysis (Boehr, Nussinov, & Wright,
419 2009; Nagel & Klinman, 2009). We propose a dynamic system in which the conformational space
420 available to *Pf*A-M17 is influenced by the presence of specific environmental signals that alter the
421 conformational energy landscape to favour certain activation states depending on the rate of
422 proteolysis required. Candidates for the signal molecules that regulate the rate of catalysis include:
423 substrate/product, co-factor, or external signal molecule. While *Hp*-M17 identified an allosteric link
424 between substrate concentration and activity, we were unable to find a substrate-based link for *Pf*A-
425 M17 using a similar substrate. However, M17 aminopeptidases from different organisms show
426 altered activity when different metal ions are bound (Allen et al., 1983; Cappiello et al., 2006;
427 Carroll et al., 2013; Maric et al., 2009; Zhu et al., 2012). For this reason, many studies have focused
428 on examination of the identity and role of the two metal ions; however, have never considered that
429 the positions themselves might be dynamic and playing a regulatory role. Our studies described
430 herein have shown that re-arrangement of active site metal ions play a key role in the mechanism of
431 *Pf*A-M17. The metal co-factor is therefore a prime candidate to act as a signalling factor in the
432 regulation of *Pf*A-M17 activity. The concentrations of different divalent metal cations are known to
433 fluctuate throughout the parasite life cycle (Marvin et al., 2012). Therefore, environmental metal
434 ion concentration could be serving to moderate *Pf*A-M17 activity according to different metabolic
435 demands throughout the complex parasite life cycle.
436

Figure 5. Schematic representation of the proposed *Pf*A-M17 regulatory mechanism. In the proposed regulatory mechanism, each trimer operates independently (trimer *abc* is shown with *a* in yellow, *b* in pink, and *c* in cyan). Environmental conditions alter the conformational space to favour (A) an inactive state in which proteolysis does not occur, (B) an intermediate state that is capable of hydrolysing substrate in one active site only, and (C) an active state wherein all sites are active concurrently.



446 **Divergence of M17 function between species**

447 M17 aminopeptidases have vital roles in a wide range of physiological processes throughout
448 all kingdoms of life. We therefore sought to determine whether the functionally important
449 communication between the active sites is a conserved feature of the conserved hexameric assembly
450 of M17 aminopeptidases. We performed a sequence alignment of M17 aminopeptidases from a
451 range of different organisms and compared the region of the flexible loop (SI 7). The alignment
452 showed that while Lys386 is highly conserved, neither the flexible loop nor the aspartic acid residue
453 that defines the third zinc binding site are conserved. Closer examination showed that the loop
454 regions could be classified into two groups; those with loops of similar length and nature to *Pf*A-
455 M17 as well as zinc site 3 (Asp or Glu), and those that have a shortened loop, often coupled with
456 increased proline content and absence of zinc site 3 (though Asp394 is still present in some of these
457 enzymes). Due to the predicted increase in rigidity and decreased length of the loop in this latter
458 group, we speculate that these enzymes are incapable of the loop dynamics described for *Pf*A-M17
459 here. Interestingly *Hp*-M17, the only M17 aminopeptidase of our knowledge to have identified
460 cooperativity of substrate binding (Dong et al., 2005), fits into the latter class. The flexible loop,
461 and the dynamic regulatory mechanism that it moderates, might therefore be key to explaining the
462 differences between the mechanisms of *Hp*-M17 and *Pf*A-M17. It has been suggested that flexible
463 loops in proteins can facilitate the emergence of novel functions, which results in divergent
464 evolution (Bhabha et al., 2013; Campbell et al., 2017; Tokuriki & Tawfik, 2009). The dynamic
465 regulatory mechanism that we have described is therefore a prime candidate for role of mediating
466 different functionalities throughout the M17 enzyme family.

467

468 **Conclusion:**

469 The M17 aminopeptidases have been of interest as key players in a wide range of
470 physiological processes for over 50 years. From a protein mechanics perspective, the conserved
471 hexameric assembly containing six discrete active sites, represents an exciting system in which to

472 investigate the role of dynamic cooperativity within a large oligomer. Herein, we have shown that
473 the hexameric assembly is absolutely essential for the proteolytic activity of *Pf*A-M17. The
474 arrangement stabilises and preserves key catalytic machinery, and further, regulates access routes to
475 the catalytic cavity through movements in surface loops that moderate the pore size of key
476 channels. Within the hexamer, each of the two disc-like trimers operate independently; a flexible
477 loop links each of the three active sites to its neighbour, thereby operating cooperatively to convert
478 the enzyme between inactive and active states. Further, movement of the regulatory loop is coupled
479 with a rearrangement of active site metal cofactors. Our studies show that not only do metal ion
480 dynamics exist, but that the positions of the metals are manipulated to control the activity of the
481 enzyme. We therefore propose a dynamic regulatory mechanism for *Pf*A-M17 that is moderated by
482 changes in the physiological metal environment. Characterisation of the function, and regulation of
483 that function, described herein for *Pf*A-M17 thereby provides insight into how the M17
484 aminopeptidases can operate in such varying capacities throughout all kingdoms of life.
485
486 **Materials and Methods:**
487 Bacterial strains and plasmids.
488 The cloning of the truncated *Pf*A-M17 gene (encoding amino acids 85 – 605) into
489 pTrcHis2B has been previously reported (McGowan et al., 2010). Site directed mutagenesis was
490 performed by PCR and confirmed by DNA sequencing. *Escherichia coli* strains DH5α and K-12
491 were used for DNA manipulation, and BL21(DE3) used for protein expression. Recombinant His_6-
492 tagged wild type and mutant *Pf*A-M17 were expressed using an autoinduction method as previously
493 described for *wild type Pf*A-M17 (McGowan et al., 2010).
494
495 Protein purification and analysis by gel filtration chromatography.
496 Proteins were purified as has previously been described for wild type *Pf*A-M17 (McGowan
497 et al., 2010) using a two-step purification procedure of Ni–NTA agarose column followed by gel

498 filtration chromatography on a Superdex S200 10/300 column in 50 mM HEPES pH 8.0, 150 mM
499 NaCl buffer. Analytical gel filtration chromatography was performed on a Superdex S200 Increase
500 10/300 in 50 mM HEPES pH 8.0, 150 mM NaCl. Approximate molecular weight of eluate was
501 calculated by interpolation of a standard curve constructed with appropriate molecular weight
502 standards.
503
504 Aminopeptidase assays and analysis.
505 Aminopeptidase activity was determined by measuring the release of the fluorogenic-
506 leaving group, *L*-Leucine-7-amido-4-methylcoumarin hydrochloride (Sigma-Aldrich, L2145)
507 (NHMec), from the fluorogenic peptide substrates H-Leu-NHMec as described previously (Stack et
508 al., 2007). Briefly, assays were performed in triplicate and carried out in 50 µL total volume in 100
509 mM Tris-HCl, pH 8.0, 1 mM $MnCl_2$ at 37 °C and activity was monitored until steady-state was
510 achieved. Fluorescence was measured using a FluoroStar Optima plate reader (BMG Labtech), with
511 excitation and emission wavelengths of 355 nm and 460 nm, respectively. Activity of the
512 association mutant *Pf*A-M17(W525A,Y533A) was assessed with 10 µM H-Leu-NH-Mec, up to an
513 enzyme concentration of 1000 nM with no detectable activity.
514 For active enzymes, the Michaelis constant, K_M, was calculated from the initial rates over a
515 range of substrate concentrations (0.5 – 500 µM) with enzyme concentrations fixed at 150 nM for
516 the *wild type Pf*A-M17, 20 nM for *Pf*A-M17(D394A), and 1000 nM for *Pf*A-M17(A387P) and *PfA*-
517 M17(K386A). Gain was fixed at 800 for all Michaelis-Menten assays. Kinetic parameters
518 including, K_M, k_{cat}, and n_H were calculated with non-linear regression protocols by using GraphPad
519 Prism 7.
520
521 Crystallisation, data collection, structure determination and refinement.
522 *Pf*A-M17 was concentrated to 10 mg/mL in 50 mM HEPES pH 8.0, 150 mM NaCl for
523 crystallisation. Crystals were grown by hanging drop vapour diffusion, in 20 % PEG3350, 0.2 M

524 calcium acetate, with drops composed of 2 μL protein plus 1 μL precipitant. Crystals grew to large
525 plates in 7 days. For soaking experiments, crystals were transferred to fresh drop composed of
526 crystallisation solution supplemented with 1 mM of bestatin (Sigma-Aldrich, B8385) for 24 hours
527 prior to cryoprotection. Crystals were cryoprotected in mother liquor supplemented with 15 % 2-
528 Methyl-2,4-pentanediol for 30 s before flash cooling in liquid nitrogen.

529 Data were collected at 100 K using synchrotron radiation at the Australian Synchrotron
530 using the micro crystallography MX2 beamline 3ID1. Data were collected from two wedges of the
531 same crystal, which were merged after integration. Data were processed using iMosflm (Battye,
532 Kontogiannis, Johnson, Powell, & Leslie, 2011) and Aimless (Evans & Murshudov, 2013) as part
533 of the CCP4i suite (Winn et al., 2011). The structure was solved by molecular replacement in
534 Phaser using the structure of unliganded *Pf*A-M17 (RCSB ID 3KQZ) as the search model
535 (McGowan et al., 2010). Refinement was carried out with iterative rounds of model building in
536 Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010) and refinement using Phenix (Adams et al.,
537 2010). Water molecules and metal ions were placed manually based on the presence of F_o–F_c and
538 $2F_o$–F_c electron density of appropriate signal. The structure was validated with MolProbity (Chen et
539 al., 2010) and figures generated using PyMOL version 1.8.23. Final structure coordinates were
540 deposited in the protein databank (RCSB ID 6BVO), and a summary of data collection and
541 refinement statistics is provided in Supplementary Information 8. For clarity and consistency, the
542 chains in the deposited structure (PDB ID 6VBO) are numbered according to the model of
543 unliganded *Pf*A-M17 (PDB ID 3KQZ), whereby chain *a*, *b*, and *c* are equivalent to A, B, and C
544 respectively, and *a**, *b**, *c** are equivalent to D, F, and E respectively.

545

546 MD system setup and simulation protocol.

547 The starting *Pf*A-M17 model for MD simulations was based on the unliganded crystal
548 structure, 3KQZ. To prepare the model, missing atoms and residues ($a^{84,\ 257\text{-}261}$, $b^{84\text{-}85,\ 255\text{-}262}$, $c^{84\text{-}85,}$
549 $^{255\text{-}259}$, $a^{*84,\ 255\text{-}259}$, $b^{*84\text{-}85,\ 136,\ 255\text{-}261}$, $c^{*84\text{-}85,\ 152}$) were rebuilt using Modeller v9.11 (Fiser & Sali,

2003), protonated according to their states at pH 7.0 using the PDB2PQR server (Dolinsky et al., 2007) and subjected to energy minimisation using Modeller (Fiser & Sali, 2003). The hexameric M17 system consisted of ~ 294 000 atoms with a periodic box of 167 Å × 166 Å × 121 Å and the monomeric M17 system consisted of ~ 81,000 atoms with a periodic box of 120Å × 85 Å × 94 Å (after solvation). Periodic boundary conditions (pbc) were used at all simulations. System charges were neutralised with sodium counter ions. Proteins and ions were modelled using the AMBER force field FF12SB (Case et al., 2005) , the metal centre was defined as described previously (Yang et al., 2017) and waters represented using the 3-particle TIP3P model (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983). All atom MD simulations were performed using NAMD 2.9 on an IBM Blue Gene/Q cluster (monomer simulations) or x86 (hexamer simulations). Equilibration was performed in three stages. First, potential steric clashes in the initial configuration were relieved with 50000 steps of energy minimization. Initial velocities for each system were then assigned randomly according to a Maxwell–Boltzmann distribution at 100 K. Each system was then heated to 300 K over 0.1 ns, under the isothermal-isometric ensemble (NVT) conditions, with the protein atoms (excluding hydrogens) harmonically restrained (with a force constant of 10 kcal mol^{-1} A^{-2}). Following this, each system was simulated for 100 ps under the isothermal-isobaric ensemble (NPT) with heavy atoms restrained. The harmonic restrained used were reduced from 10 to 2 kcal mol^{-1} A^{-2} during the simulations. The above equilibration process was performed three times from the same starting structure in order to initiate three production simulations with different initial velocities. For production simulations, the time step was set to 2 fs and the SHAKE algorithm was used to constrain all bonds involving hydrogen atoms. All simulations were run at constant temperature (300 K) and pressure (1 atm), using a Langevin damping coefficient of 0.5 fs^{-1}, and a Berendsen thermostat relaxation time of τ_P = 0.1 ps. The Particle-Mesh Ewald (PME) method was used to set the periodic boundary conditions (PBC) that were used for long-range electrostatic interactions and a real space cut-off of 10 Å was used. Conformations were sampled every 10 ps for subsequent analysis. All frames with time interval of 10 ps were saved to disk.

576

577 MD Analysis.

578 Simulation trajectories were analysed using the GROMACS 5.14 simulation package. For
579 principle component analysis (PCA), 3N*3N atom covariance matrices of the protein displacement
580 in simulations were generated based on backbone atoms (N, Cα, C, O) of the *Pf*A-M17 crystal
581 structure. Principle Components (PCs), that taken together accounted for more than 50% of the
582 overall covariance, were chosen for essential dynamics analysis. The GROMACS 5.14 simulation
583 package was used to project the trajectory onto the top PCs. Graphs and plots were produced with
584 Xmgrace and GraphPad Prism7. Molecular graphics were prepared with PyMOL 1.8.23 and
585 VMD1.9.3.

586

587 Channel analysis.

588 The channels to the interior of the *Pf*A-M17 hexamer were examined using Caver
589 (Chovancova et al., 2012). Channel dimensions were assessed at the start of the simulation, which
590 represented the crystallographic conformation of *Pf*A-M17 after the missing loop residues were
591 modelled and subjected to energy minimisation, and the end of the simulation. Values reported are
592 for the channel 'bottleneck', which represent the diameter of each channel at the narrowest point. If
593 multiple forks off a single channel were observed, the dimensions for the larger channel (which
594 represent the most likely access points) were reported. Figures of channels were created using
595 Caver (Chovancova et al., 2012) and visualised using the PyMOL plugin.

596

597 Sequence alignment and analysis.

598 ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) was used to align sequences of
599 M17 aminopeptidases (*C. elegans* P34629.1; *B. taurus* AAB28170.1; *H. sapiens* AAD17527.1, *M.*
600 *musculus* AAK13495.1; *R. norvegicus* AAH79381.1; *D. melanogaster* AAF50390.2; *A. gambiae*
601 XP_321111.3; *A. thaliana* Q944P7; *S. Lycopersicum* (LAP2) XP_006350101.1; *S. Lycopersicum*

602 (LAP1) NP_001233862.2; *S. tuberosum* CAA48038.1; *T. annulata* CAI76586.1; *T. gondii*
603 XP_018636441.1; *P. falciparum* XP_001348613.1; *C. hominis* XP_667960.1; *C. parvum*
604 XP_626197.1; *L. major* AAL16097.1; *H. pylori* WP_064816828.1; *S. aureus* WP_075108680.1; *E.*
605 *coli* P68767; *R. typhi* AAU03616.1; *C. tetani* WP_011100044.1; *B. cereus* AAP11794.1).
606
607 **Abbreviations:**
608 *Pf*A-M17, *Plasmodium falciparum* M17 aminopeptidase*; Hp*-M17, *Helicobacter pylori*
609 M17 aminopeptidase; MD, molecular dynamics; PCA, principle component analysis; RDF, radial
610 distribution function.
611
612 **Author Contributions:**
613 S.M. and N.D. designed the research; N.D., W.Y., K.K.S., B.T.R., and I.K. performed the
614 research and analysed the data; A.M.B. and S.M. supervised the research; N.D. and S.M. wrote the
615 manuscript with contributions from all other authors.
616
617 **Acknowledgements:**
618 We thank the National Health and Medical Research Council (Project Grant 1063786 to S.M. and
619 P.J.S.; Fellowship 1022688 to A.M.B.) and the Australian Research Council (Fellowship
620 FT100100690 to S.M.) for funding support. This work was supported by the Victorian Life
621 Sciences Computation Initiative (VLSCI) and National Computational Infrastructure (NCI). This
622 research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of
623 ANSTO.
624

References:

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D., 66*, 213-221. doi: 10.1107/s0907444909052925

Alén, C., Sherratt, D. J., & Colloms, S. D. (1997). Direct interaction of aminopeptidase A with recombination site DNA in Xer site-specific recombination. *EMBO J., 16*(17), 5188-5197. doi: 10.1093/emboj/16.17.5188

Allen, M. P., Yamada, A. H., & Carpenter, F. H. (1983). Kinetic-parameters of metal-substituted leucine aminopeptidase from bovine lens. *Biochemistry, 22*(16), 3778-3783. doi: 10.1021/bi00285a010

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D., 67*, 271-281. doi: 10.1107/s0907444910048675

Bhabha, G., Ekiert, D. C., Jennewein, M., Zmasek, C. M., Tuttle, L. M., Kroon, G., Dyson, H. J., Godzik, A., Wilson, I. A., & Wright, P. E. (2013). Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nature Structural & Molecular Biology, 20*(11), 1243-U1262. doi: 10.1038/nsmb.2676

Bhabha, G., Lee, J., Ekiert, D. C., Gam, J., Wilson, I. A., Dyson, H. J., Benkovic, S. J., & Wright, P. E. (2011). A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science, 332*(6026), 234-238. doi: 10.1126/science.1198542

Boehr, D. D., Nussinov, R., & Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology, 5*(11), 789-796. doi: 10.1038/nchembio.232

Burley, S. K., David, P. R., Taylor, A., & Lipscomb, W. N. (1990). Molecular structure of leucine aminopeptidase at 2.7 Å resolution *Proc. Natl. Acad. Sci., 87*(17), 6878-6882. doi: 10.1073/pnas.87.17.6878

Campbell, E. C., Correy, G. J., Mabbitt, P. D., Buckle, A. M., Tokuriki, N., & Jackson, C. J. (2017). Laboratory evolution of protein conformational dynamics. *Curr Opin Struct Biol, 50*, 49-57. doi: 10.1016/j.sbi.2017.09.005

Cappiello, M., Alterio, V., Amodeo, P., Del Corso, A., Scaloni, A., Pedone, C., Moschini, R., De Donatis, G. M., De Simone, G., & Mura, U. (2006). Metal ion substitution in the catalytic site greatly affects the binding of sulfhydryl-containing compounds to leucyl aminopeptidase. *Biochemistry, 45*(10), 3226-3234. doi: 10.1021/bi052069v

Carroll, R. K., Veillard, F., Gagne, D. T., Lindenmuth, J. M., Poreba, M., Drag, M., Potempa, J., & Shaw, L. N. (2013). The Staphylococcus aureus leucine aminopeptidase is localized to the bacterial cytosol and demonstrates a broad substrate range that extends beyond leucine. *Biol. Chem., 394*(6), 791-803. doi: 10.1515/hsz-2012-0308

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem., 26*(16), 1668-1688. doi: 10.1002/jcc.20290

Charlier, D., Hassanzadeh, G., Kholti, A., Gigot, D., Pierard, A., & Glansdorff, N. (1995). *carP*, Involved in Pyrimidine Regulation of the *Escherichia coli* Carbamoylphosphate Synthetase Operon Encodes a Sequence-specific DNA-binding Protein Identical to XerB and PepA, also Required for Resolution of ColEl Multimers. *Journal of Molecular Biology, 250*(4), 392-406. doi: 10.1006/jmbi.1995.0385

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2010). MolProbity: all-atom structure

validation for macromolecular crystallography. *Acta Crystallogr. D., 66*, 12-21. doi: 10.1107/s0907444909042073

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., & Damborsky, J. (2012). CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *Plos Comput. Biol., 8*(10). doi: 10.1371/journal.pcbi.1002708

Dalal, S., & Klemba, M. (2007). Roles for two aminopeptidases in vacuolar hemoglobin catabolism in Plasmodium falciparum. *J. Biol. Chem., 282*(49), 35978-35987. doi: 10.1074/jbc.M703643200

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res., 35*, W522-W525. doi: 10.1093/nar/gkm276

Dong, L., Cheng, N., Wang, M. W., Zhang, J. F., Shu, C., & Zhu, D. X. (2005). The leucyl aminopeptidase from Helicobacter pylori is an allosteric enzyme. *Microbiology, 151*, 2017-2023. doi: 10.1099/mic.0.27767-0

Drinkwater, N., Bamert, R. S., Sivaraman, K. K., Paiardini, A., & McGowan, S. (2015). X-ray crystal structures of an orally available aminopeptidase inhibitor, Tosedostat, bound to anti-malarial drug targets PfA-M1 and PfA-M17. *Proteins, 83*(4), 789-795. doi: 10.1002/prot.24771

Drinkwater, N., Vinh, N. B., Mistry, S. N., Bamert, R. S., Ruggeri, C., Holleran, J. P., Loganathan, S., Paiardini, A., Charman, S. A., Powell, A. K., Avery, V. M., McGowan, S., & Scammells, P. J. (2016). Potent dual inhibitors of Plasmodium falciparum M1 and M17 aminopeptidases through optimization of S1 pocket interactions. *Eur. J. Med. Chem., 110*, 43-64. doi: 10.1016/j.ejmech.2016.01.015

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D., 66*, 486-501. doi: 10.1107/s0907444910007493

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D., 69*, 1204-1214. doi: 10.1107/s0907444913000061

Fiser, A., & Sali, A. (2003). MODELLER: Generation and refinement of homology-based protein structure models. In C. W. Carter & R. M. Sweet (Eds.), *Methods Enzymol.* (Vol. 374, pp. 461-491).

Gerstein, M., & Chothia, C. (1996). Packing at the protein-water interface. *Proceedings of the National Academy of Sciences of the United States of America, 93*(19), 10167-10172. doi: 10.1073/pnas.93.19.10167

Harbut, M. B., Velmourougane, G., Dalal, S., Reiss, G., Whisstock, J. C., Onder, O., Brisson, D., McGowan, S., Klemba, M., & Greenbaum, D. C. (2011). Bestatin-based chemical biology strategy reveals distinct roles for malaria M1-and M17-family aminopeptidases. *Proc. Natl. Acad. Sci., 108*(34), E526-E534. doi: 10.1073/pnas.1105601108

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys., 79*(2), 926-935. doi: 10.1063/1.445869

Kumar, S., Kaur, A., Chattopadhyay, B., & Bachhawat, A. K. (2015). Defining the cytosolic pathway of glutathione degradation in Arabidopsis thaliana: role of the ChaC/GCG family of gamma-glutamyl cyclotransferases as glutathione-degrading enzymes and AtLAP1 as the Cys-Gly peptidase. *Biochemical Journal, 468*, 73-85. doi: 10.1042/bj20141154

Lowther, W. T., & Matthews, B. W. (2002). Metalloaminopeptidases: Common functional themes in disparate structural surroundings. *Chem. Rev., 102*(12), 4581-4607. doi: 10.1021/cr0101757

Maric, S., Donnelly, S. M., Robinsin, M. W., Skinner-Adams, T., Trenholme, K. R., Gardiner, D. L., Dalton, J. P., Stack, C. M., & Lowther, J. (2009). The M17 Leucine Aminopeptidase of the Malaria Parasite Plasmodium falciparum: Importance of Active Site Metal Ions in the Binding of Substrates and Inhibitors. *Biochemistry, 48*(23), 5435-5439. doi: 10.1021/bi9003638

Marvin, R. G., Wolford, J. L., Kidd, M. J., Murphy, S., Ward, J., Que, E. L., Mayer, M. L., Penner-Hahn, J. E., Haldar, K., & O'Halloran, T. V. (2012). Fluxes in "Free" and Total Zinc Are Essential for Progression of Intraerythrocytic Stages of Plasmodium falciparum. *Chem. Biol., 19*(6), 731-741. doi: 10.1016/j.chembiol.2012.04.013

Matsui, M., Fowler, J. H., & Walling, L. L. (2006). Leucine aminopeptidases: diversity in structure and function. *Biol. Chem., 387*(12), 1535-1544. doi: 10.1515/bc.2006.191

McElheny, D., Schnell, J. R., Lansing, J. C., Dyson, H. J., & Wright, P. E. (2005). Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc. Natl. Acad. Sci., 102*(14), 5032-5037. doi: 10.1073/pnas.0500699102

McGowan, S., Oellig, C. A., Birru, W. A., Caradoc-Davies, T. T., Stack, C. M., Lowther, J., Skinner-Adams, T., Mucha, A., Kafarski, P., Grembecka, J., et al. (2010). Structure of the Plasmodium falciparum M17 aminopeptidase and significance for the design of drugs targeting the neutral exopeptidases. *Proc. Natl. Acad. Sci., 107*(6), 2449-2454. doi: 10.1073/pnas.0911813107

Minh, P. N. L., Devroede, N., Massant, J., Maes, D., & Charlier, D. (2009). Insights into the architecture and stoichiometry of Escherichia coli PepADNA complexes involved in transcriptional control and site-specific DNA recombination by atomic force microscopy. *Nucleic Acids Research, 37*(5), 1463-1476. doi: 10.1093/nar/gkn1078

Mistry, S. N., Drinkwater, N., Ruggeri, C., Sivaraman, K. K., Loganathan, S., Fletcher, S., Drag, M., Paiardini, A., Avery, V. M., Scammells, P. J., & McGowan, S. (2014). Two-Pronged Attack: Dual Inhibition of Plasmodium falciparum M1 and M17 Metalloaminopeptidases by a Novel Series of Hydroxamic Acid-Based Inhibitors. *J. Med. Chem., 57*(21), 9168-9183. doi: 10.1021/jm501323a

Nagel, Z. D., & Klinman, J. P. (2009). A 21(st) century revisionist's view at a turning point in enzymology. *Nat. Chem. Biol., 5*(8), 543-550. doi: 10.1038/nchembio.204

Rawlings, N. D., Barrett, A. J., & Finn, R. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res., 44*(D1), D343-D350. doi: 10.1093/nar/gkv1118

Reboul, C. F., Porebski, B. T., Griffin, M. D. W., Dobson, R. C. J., Perugini, M. A., Gerrard, J. A., & Buckle, A. M. (2012). Structural and Dynamic Requirements for Optimal Activity of the Essential Bacterial Enzyme Dihydrodipicolinate Synthase. *Plos Comput. Biol., 8*(6). doi: 10.1371/journal.pcbi.1002537

Rosenthal, P. J. (2002). Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr. Opin. Hematol., 9*(2), 140-145. doi: 10.1097/00062752-200203000-00010

Schurer, G., Horn, A. H. C., Gedeck, P., & Clark, T. (2002). The reaction mechanism of bovine lens leucine aminopeptidase. *J. Phys. Chem. B, 106*(34), 8815-8830. doi: 10.1021/jp025575s

Scranton, M. A., Yee, A., Park, S. Y., & Walling, L. L. (2012). Plant Leucine Aminopeptidases Moonlight as Molecular Chaperones to Alleviate Stress-induced Damage. *J. Biol. Chem., 287*(22), 18408-18417. doi: 10.1074/jbc.M111.309500

Skinner-Adams, T. S., Lowther, J., Teuscher, F., Stack, C. M., Grembecka, J., Mucha, A., Kafarski, P., Trenholme, K. R., Dalton, J. P., & Gardiner, D. L. (2007). Identification of phosphinate dipeptide analog inhibitors directed against the Plasmodium falciparum M17 leucine aminopeptidase as lead antimalarial compounds. *J. Med. Chem., 50*(24), 6024-6031. doi: 10.1021/jm070733v

774 Skinner-Adams, T. S., Peatey, C. L., Anderson, K., Trenholme, K. R., Krige, D., Brown, C. L., Stack,
775 C., Nsangou, D. M. M., Mathews, R. T., Thivierge, K., Dalton, J. P., & Gardinerd, D. L.
776 (2012). The Aminopeptidase Inhibitor CHR-2863 Is an Orally Bioavailable Inhibitor of
777 Murine Malaria. *Antimicrob. Agents Ch., 56*(6), 3244-3249. doi: 10.1128/aac.06245-11
778 Stack, C. M., Lowther, J., Cunningham, E., Donnelly, S., Gardiner, D. L., Trenholme, K. R.,
779 Skinner-Adams, T. S., Teuscher, F., Grembecka, J., Mucha, A., Kafarski, P., Lua, L., Bell, A.,
780 & Dalton, J. P. (2007). Characterization of the Plasmodium falciparum M17 leucyl
781 aminopeptidase - A protease involved in amino acid regulation with potential for
782 antimalarial drug development. *J. Biol. Chem., 282*(3), 2069-2080. doi:
783 10.1074/jbc.M609251200
784 Stirling, C. J., Colloms, S. D., Collins, J. F., Szatmari, G., & Sherratt, D. J. (1989). xerB, an
785 *Escherichia coli* gene required for plasmid ColE1 site-specific recombination, is
786 identical to pepA, encoding aminopeptidase A, a protein with substantial similarity to
787 bovine lens leucine aminopeptidase. *Embo Journal, 8*(5), 1623-1627.
788 Sträter, N., & Lipscomb, W. N. (1995). 2-Metal ion mechanism of bovine lens leucine
789 aminopeptidase - active site solvent structure and binding mode of L-leucinal, a gem-
790 diolate transition-state analog, by X-ray crystallography. *Biochemistry, 34*(45), 14792-
791 14800. doi: 10.1021/bi00045a021
792 Sträter, N., Sherratt, D. J., & Colloms, S. D. (1999). X-ray structure of aminopeptidase A from
793 *Escherichia coli* and a model for the nucleoprotein complex in Xer site-specific
794 recombination. *EMBO J., 18*(16), 4513-4522. doi: 10.1093/emboj/18.16.4513
795 Tokuriki, N., & Tawfik, D. S. (2009). Protein Dynamism and Evolvability. *Science, 324*(5924),
796 203-207. doi: 10.1126/science.1169375
797 Vögeli, B., Bibow, S., & Chi, C. N. (2016). Enzyme Selectivity Fine-Tuned through Dynamic
798 Control of a Loop. *Angewandte Chemie-International Edition, 55*(9), 3096-3100. doi:
799 10.1002/anie.201511476
800 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M.,
801 Krissinel, E. B., Leslie, A. G. W., McCoy, A., et al. (2011). Overview of the CCP4 suite and
802 current developments. *Acta Crystallogr. D., 67*, 235-242. doi:
803 10.1107/s0907444910045749
804 Wurm, J. P., Holdermann, I., Overbeck, J. H., Mayer, P. H. O., & Sprangers, R. (2017). Changes in
805 conformational equilibria regulate the activity of the Dcp2 decapping enzyme. *Proc.*
806 *Natl. Acad. Sci., 114*(23), 6034-6039. doi: 10.1073/pnas.1704496114
807 Yang, W., Riley, B. T., Lei, X., Porebski, B. T., Kass, I., Buckle, A. M., & McGowan, S. (2017).
808 Generation of AMBER force field parameters for zinc centres of M1 and M17 family
809 aminopeptidases. *J. Biomol. Struct. Dyn.*, 1-10. doi: 10.1080/07391102.2017.1364669
810 Zhu, X. X., Barman, A., Ozbil, M., Zhang, T. T., Li, S. H., & Prabhakar, R. (2012). Mechanism of
811 peptide hydrolysis by co-catalytic metal centers containing leucine aminopeptidase
812 enzyme: a DFT approach. *J. Biol. Inorg. Chem., 17*(2), 209-222. doi: 10.1007/s00775-
813 011-0843-2
814

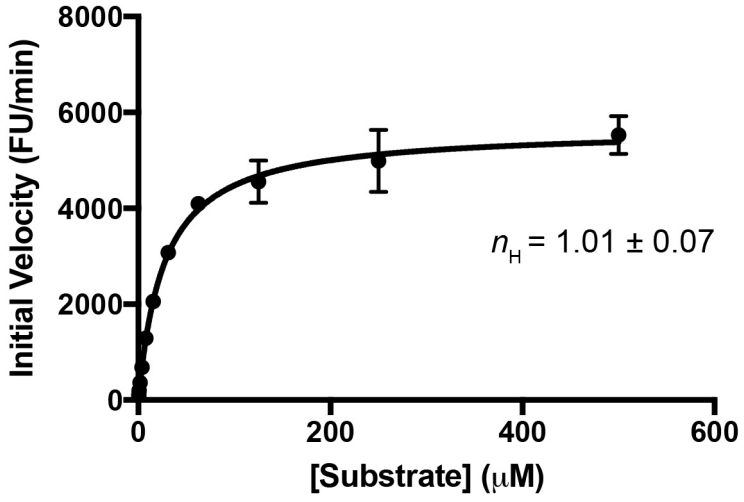
815

816 **Supplementary Information 1:** Analysis of the Hill coefficient (n_H) by substrate-saturation

817 experiment showed no evidence of cooperativity in *Pf*A-M17 catalysis. Data shown are means ± SD

818 (n = 3).

819



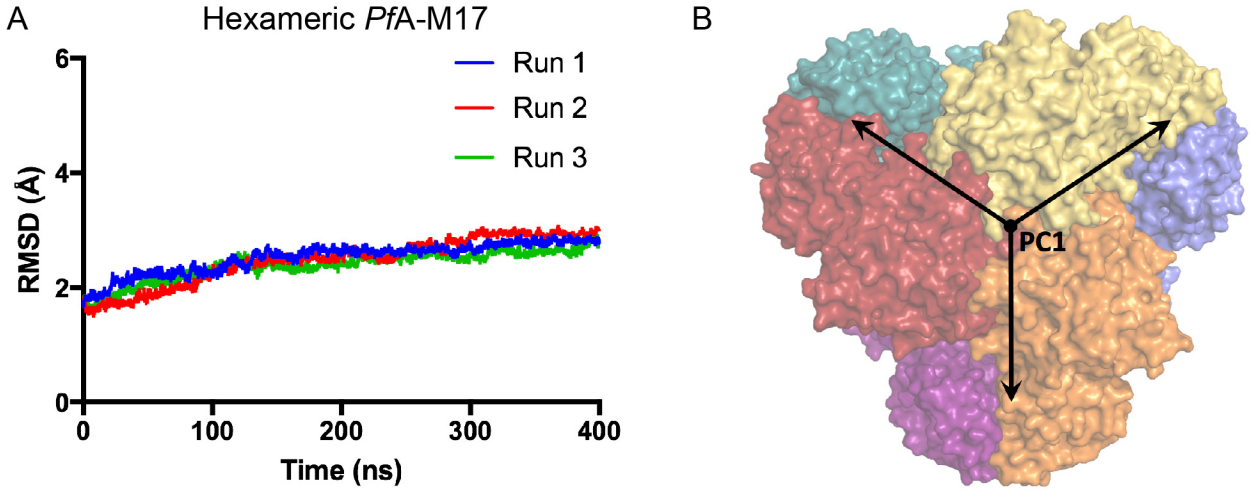
820

821

822 **Supplementary Information 2:** (A) RMSD plot of hexameric *Pf*A-M17 over course of simulation.

823 (B) Graphical representation of PC1 (black arrows, stretching motion).

824



825

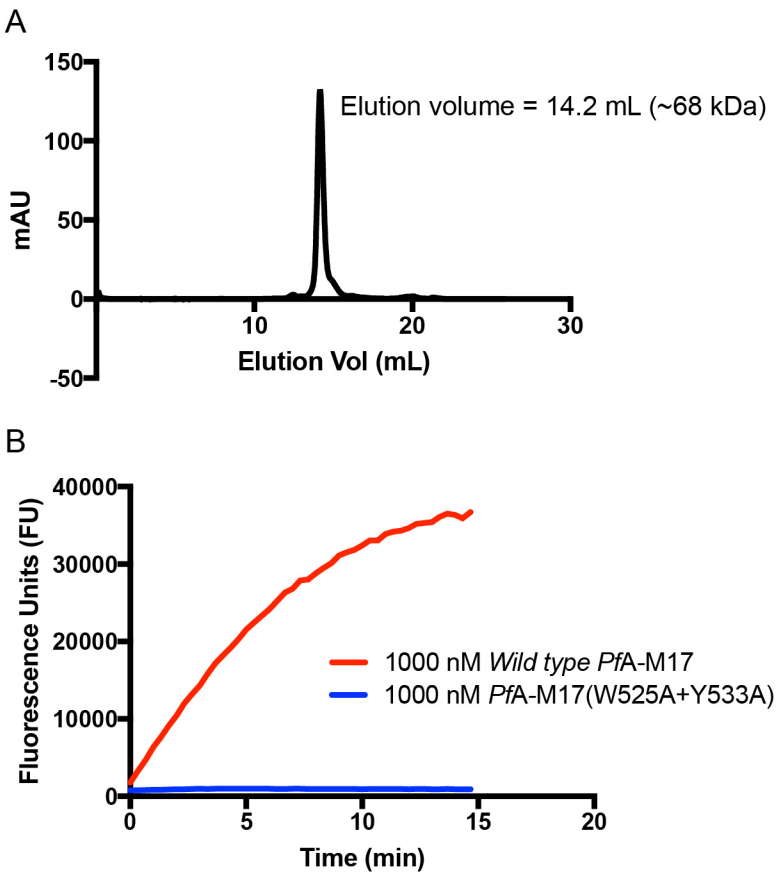
826

Supplementary Information 3: Table of change in pore dimensions over course of the simulation (measurement is of the bottleneck of each of the channels as calculated by Caver).

Channel	Defined by loops	Channel diameter at bottleneck (Å)		
		Start	*End*	*Change*
1	$a^{247–266}$:$a*^{132–150}$	3.8	6.5	2.7
2	$b^{247–266}$:$b*^{132–150}$	4.0	5.9	1.9
3	$c^{247–266}$:$c*^{132–150}$	3.9	4.5	0.6
4	$a^{132–150}$:$a*^{247–266}$	4.6	6.5	1.9
5	$b^{132–150}$:$b*^{247–266}$	4.4	7.1	2.7
6	$c^{132–150}$:$c*^{247–266}$	4.9	8.9	4.0

831 **Supplementary Information 4:** *Pf*A-M17(Y525A+Y533A) is monomeric in solution as
832 determined by analytical size exclusion chromatography. Absorbance trace (A) shows that the
833 association mutant elutes at 14.2 mL, which corresponds to an approximate molecular weight of
834 68 kDa (*Pf*A-M17 molecular weight = 59 kDa). The mutant therefore represents monomeric *Pf*A-
835 M17. Assessment of the catalytic activity of *Pf*A-M17(W525A+Y533A) under *wild type*
836 conditions, showed that the monomeric enzyme is inactive (B).

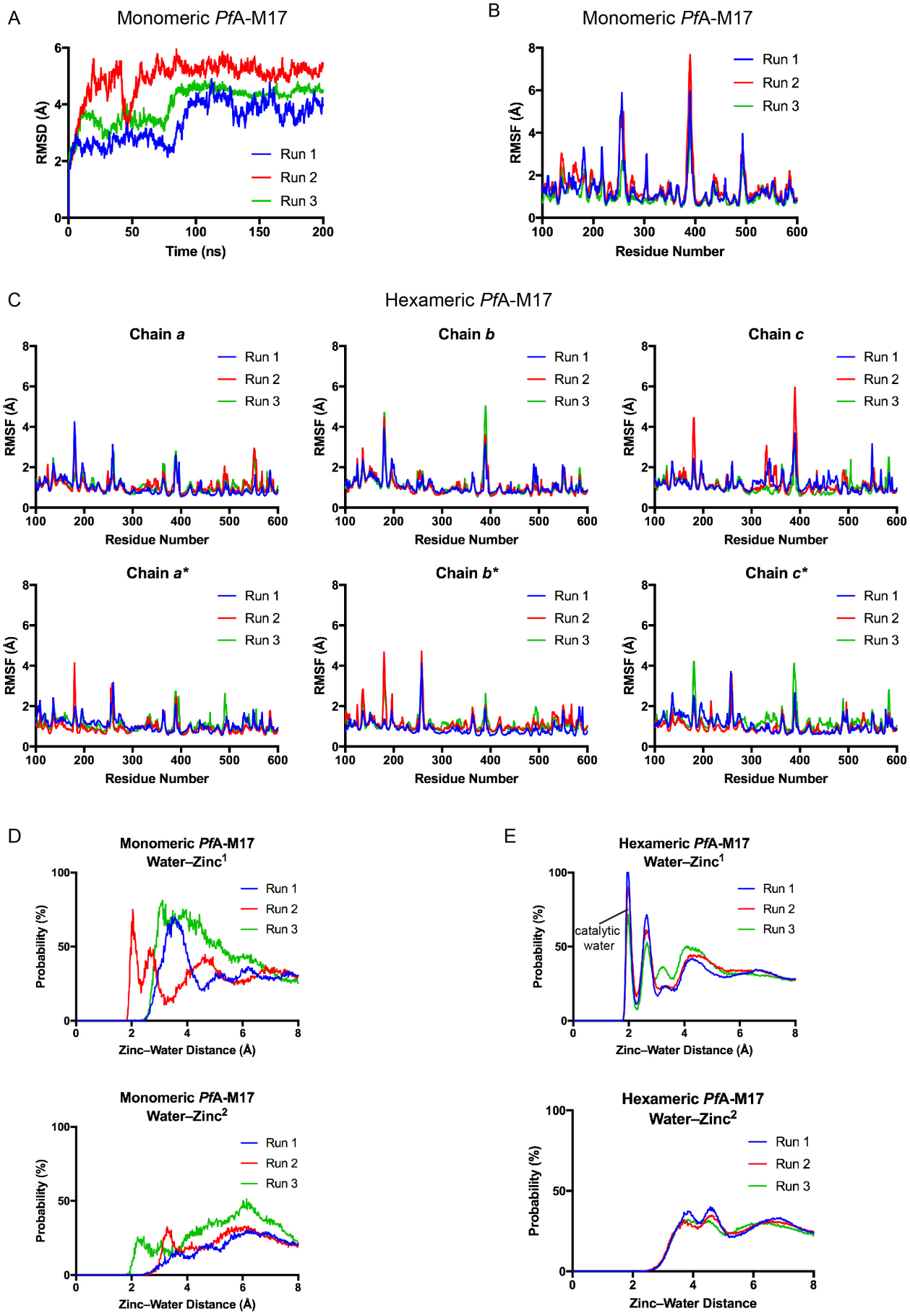
837



838

839

840 **Supplementary Information 5:** MD simulations of monomer versus hexamer (A) RMSD analysis
841 of monomer throughout the simulation. (B) RMSF analysis of monomer throughout simulation. (C)
842 RMSF analysis of individual six chains in hexamer throughout simulation. Radial distribution
843 function (RDF) plot of water molecules in the zinc environment from the MD simulations of (D)
844 the monomer and (E) the hexamer. The RDF was calculated by the average of all six sites in the
845 hexamer. Probability percentage was calculated by the occupancy of the water divided by the
846 number of waters (as measured by oxygen atom of the waters) at any given distance to the zinc
847 ions. The plots show that the catalytic water in simulation of monomer is unstable, however, the
848 hexamer has a single water molecule in the catalytic position throughout the simulation.


A
Monomeric PfA-M17
RMSD (Å)
Time (ns)
Run 1
Run 2
Run 3
B
Monomeric PfA-M17
RMSF (Å)
Residue Number
C
Hexameric PfA-M17
Chain a
Chain b
Chain c
Chain a*
Chain b*
Chain c*
D
Monomeric PfA-M17
Water–Zinc1
Probability (%)
Zinc–Water Distance (Å)
Monomeric PfA-M17
Water–Zinc2
E
Hexameric PfA-M17
Water–Zinc1
catalytic water
Hexameric PfA-M17
Water–Zinc2
Zinc–Water Distance

851 **Supplementary Information 6: Crystallographic Data Collection and Refinement Statistics**

Data collection	
Space Group	P 21 21 21
Cell dimensions:	
a, b, c (Å)	172.73, 178.53, 222.79
α = β = γ (°)	90.0
Resolution range (Å)	48.38 – 2.30 (2.382 – 2.30)
Total Observations	2,540,053 (123,225)
Unique Observations	301,951 (14,522)
Multiplicity	8.4 (8.5)
Completeness (%)	99.4 (97.4)
<I/σI>	3.6 (0.9)
CC(1/2)	0.987 (0.451)
R_{pim} (%)	14.9 (100)
PDB	6VBO
Structure refinement	
Non hydrogen atoms	
Protein	47,419
Water	3235
Metal (Zn)	12
Ligand	111
R_{free} *(%)*	23.9
R_{cryst} *(%)*	28.9
RMS deviations	
Bond lengths (Å)	0.002
Angles (°)	0.57
Ramachandran plot	
Favoured(%)	96.3
Outliers (%)	0.2
B factors (Å^2)	
Protein	37.1
Solvent	40.5
Ligand	52.3
Molprobity	
Clashscore	3.2 (100th percentile)
MolProbity score	1.4 (99th percentile)

852

Supplementary Information 7: Sequence alignment of verified M17 aminopeptidases from a range of organisms performed with Clustal Omega. Conserved lysine residue (Lys386 in *Pf*A-M17) is highlighted in green and loop residues are highlighted in pink. Yellow highlight shows that Asp394 of *Pf*A-M17, representative of zinc binding site 3, is not conserved throughout the enzyme family. Blue box indicates the M17 aminopeptidases that possess a loop of similar length and nature to *Pf*A-M17.

```
C. elegans            IYHVGK-AG--PTPPAFVVLSHEVP-G-STEHIALVGKGVVYDTGGLQIKT--KTGMPNM   272
B. taurus             FLSVAK-GS--EEPPVFLEIHYKGSPNASEPPLVFVGKGITFDSGGISIKA--AANMDLM   296
H. sapiens            FLSVAK-GS--DEPPVFLEIHYKGSPNANEPPLVFVGKGITFDSGGISIKA--SANMDLM   302
M. musculus           FLSVAK-GS--EEPPVFLEIHYMGSPNATEAPLVFVGKGITFDSGGISIKA--SANMDLM   302
R. norvegicus         FLSVAK-GS--EEPPVFLEIHYTGSPNATEAPLVFVGKGITFDSGGISIKA--SANMDLM   302
D. melanogaster       FLAVGK-AS--CEPPIFLELSYYGT-CAEERPIVLVGQGITYDAGGLCLKK--KKELFNM   331
A. gambiae            FLTVAK-GS--CEPPIFLELSYYGT-NTKERPIVLIGQGNTFDSGGLCLKT--LEALTDM   315
A. thaliana           YLAVAA-AS--ANPPHFIHLIYKPSSGPVKTKLALVGKGLTFDSGGYNIKTGPGCLIELM   373
S. Lycopersicum LAP2  YLRVAA-AS--ANPAHFIHLCYKPSSGEIKKKIALVGKGLTFDSGGYNIKTGPGCSIELM   361
S. tuberosum          YLGVAAAAT--ENPPYFIHLCFKTNSRERKTKIALVGKGLTFDSGGYNLKTGAGSKIELM   345
S. Lycopersicum LAP1  YLAVAAAAT--ENPPYFIHLCFKTPTKERKTKLALVGKGLTFDSGGYNLKVGAGSRIELM   364
T. annulata           YLSVCQ-GS--KFPPKFLHAVYKP-DGEVKKTLAFVGKGITFDAGGYNVKN-SSTMIHYM   318
T. gondii             YLAVAK-GS--LFPPQFIHLTYKS-SN-PKRRVAFVGKGICFDAGGYNLKT-GGAQIELM   566
P. falciparum         YLSVGK-GS--MYPNKFIHLTYKS-KGDVKKKIALVGKGITFDSGGYNLKAAPGSMIDLM   396
C. hominis            FLAVAQ-GS--KSPAQFVHLTYKP-KGEIKKRIALVGKGITMDTGGYNIKH---QMIHFM   315
C. parvum             FLAVAQ-GS--KSPAQFVHLTYKP-KGEIKKRIALVGKGITMDTGGYNIKH---QMIHFM   341
L. major              MYNVGR-GS--RYEPYLVVLEYIGNPR-SSATTAIVGKGVTFDCGGLNIKP--YTSMETM   345
H. pylori             FLAVNK-ASLSVNPPRLIHLVYKPKKA--KKKIALVGKGLTYDCGGLSLKP--ADYMVTM   284
S. aureus             IHAVGK-GS--EHPPVVITMTYNGGDS-NDAPIALVGKGITYDSGGYSIKS--KIGMQTM   280
E. coli               YLAVGQ-GS--QNESLMSVIEYKGNASEDARPIVLVGKGLTFDSGGISIKP--SEGMDEM   290
R. typhi              LLGVGQ-GS--QNESKLVVMEYKGGSR-DDSTIALVGKGVIFDTGGISLKP--SSNMHLM   285
C. tetani             FLAVSK-GS--FEESQLIVMNYKGNSN-SDKKLALVGKGLTYDSGGYSIKP--TSSMINM   267
B. cereus             LLAVNQ-GS--VEPPKMIALIYKGKEE-WKDVIGFVGKGITYDTGGYSLKP--REGMVGM   280
                         *   .        .    .           ::*:*   * **  :*      :  *
```