

## STRUCTURAL BIOLOGY

# Asymmetry in catalysis by *Thermotoga maritima* membrane-bound pyrophosphatase demonstrated by a nonphosphorus allosteric inhibitor

Keni Vidilaseris<sup>1</sup>, Alexandros Kiriazis<sup>2</sup>, Ainoleena Turku<sup>2</sup>, Ayman Khattab<sup>3</sup>, Niklas G. Johansson<sup>2</sup>, Teppo O. Leino<sup>2</sup>, Paula S. Kiuru<sup>2</sup>, Gustav Boije af Gennäs<sup>2</sup>, Seppo Meri<sup>3</sup>, Jari Yli-Kauhaluoma<sup>2</sup>, Henri Xhaard<sup>2</sup>, Adrian Goldman<sup>1,4\*</sup>

Copyright © 2019  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
License 4.0 (CC BY).

Membrane-bound pyrophosphatases are homodimeric integral membrane proteins that hydrolyze pyrophosphate into orthophosphates, coupled to the active transport of protons or sodium ions across membranes. They are important in the life cycle of bacteria, archaea, plants, and parasitic protists, but no homologous proteins exist in vertebrates, making them a promising drug target. Here, we report the first nonphosphorus allosteric inhibitor of the thermophilic bacterium *Thermotoga maritima* membrane-bound pyrophosphatase and its bound structure together with the substrate analog imidodiphosphate. The unit cell contains two protein homodimers, each binding a single inhibitor dimer near the exit channel, creating a hydrophobic clamp that inhibits the movement of  $\beta$ -strand 1–2 during pumping, and thus prevents the hydrophobic gate from opening. This asymmetry of inhibitor binding with respect to each homodimer provides the first clear structural demonstration of asymmetry in the catalytic cycle of membrane-bound pyrophosphatases.

## INTRODUCTION

Membrane-bound pyrophosphatases (mPPases) are a family of enzymes that hydrolyze pyrophosphate into two phosphates and couple this reaction with proton and/or sodium transport across the membrane, creating an electrochemical gradient. These enzymes, initially discovered in photosynthetic bacteria and plants (1–3) and later found in parasitic protists, archaea, and certain species of bacteria, do not occur in animals and humans (4, 5). In various organisms, mPPases are important for survival under diverse stress situations due to energy limitation, such as osmotic stress, anoxia, mineral deficiency, low temperature, and intense light (6). In bacteria and archaea, mPPases reside in the cell membrane (4), while in protists, algae, and plants, they can also be located in acidocalcisomes, the vacuole, and/or Golgi apparatus (7, 8).

mPPases are homodimeric. Each monomer contains 15 to 17 transmembrane helices (TMHs) with a molecular weight of 70 to 81 kDa. So far, there are structures from just two mPPases: four of *Thermotoga maritima* (TmPPase) and two of mung bean (*Vigna radiata*: VrPPase). TmPPase structures are available in the resting state (TmPPase:Ca:Mg) (9), with two phosphates bound (TmPPase:2P<sub>i</sub>) (9), with the substrate-analog imidodiphosphate (IDP) bound (TmPPase:IDP) (10), and with the phosphate analog (WO<sub>4</sub>) bound (TmPPase:WO<sub>4</sub>) (10). VrPPase has been solved in the IDP-bound (VrPPase:IDP) (11) and single phosphate-bound states (VrPPase:P<sub>i</sub>) (10). In all these structures, mPPase is a symmetric homodimer, with each monomer consisting of 16 TMHs. These helices form two concentric layers, with 6 helices (TMH5, TMH6, TMH11, TMH12, TMH15, and TMH16) forming the inner layer and the other 10 (TMH1 to TMH4, TMH7 to TMH10, TMH13, and

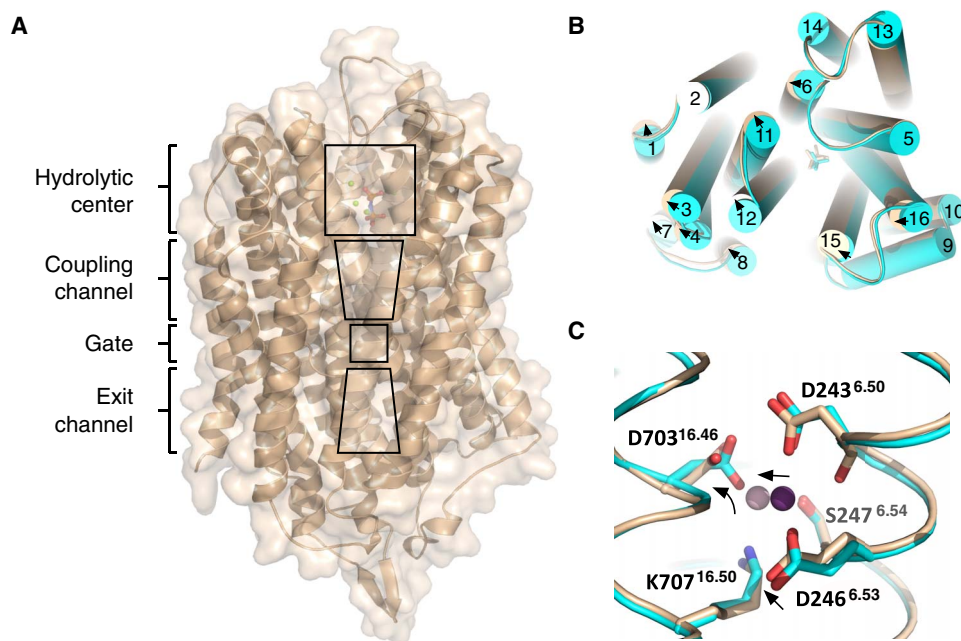
TMH14) forming the outer layer. Each monomer consists of four regions: a hydrolytic center, a coupling funnel, an ion gate, and an exit channel (Fig. 1A) (12). The hydrolytic center and coupling funnel are composed of TMH5, TMH6, TMH11, TMH12, TMH15, and TMH16 (Fig. 1B). Upon substrate binding, the active site is closed by a long loop between TMH5 and TMH6 and opened again after hydrolysis and ion pumping (9–11). In the structure of TmPPase:IDP, Na<sup>+</sup> binds within the membrane plane at the ionic gate between TMH6 and TMH16. Na<sup>+</sup> is pentacoordinated by the carboxylate groups of D243<sup>6,50</sup>, E246<sup>6,53</sup>, and D703<sup>16,46</sup>, by the O<sup>T</sup> of S247<sup>6,54</sup>, and with the main-chain carbonyl of D243<sup>6,50</sup> (Fig. 1C) (10). [Residues are numbered as in (12), shown as X $\Sigma^{a,b}$ , where X is the amino acid,  $\Sigma$  is amino acid position in TmPPase, *a* is the helix number, and *b* is the amino acid position according to a central conserved residue in each helix.]

Parasitic protists such as *Plasmodium falciparum*, *Toxoplasma gondii*, *Trypanosoma brucei*, and *Leishmania donovani* all have H<sup>+</sup>-pumping mPPases (13) of two different types (K<sup>+</sup> dependent and K<sup>+</sup> independent) in *Plasmodium* (14) and of just one type (K<sup>+</sup> dependent) in the others (15–17). The mPPases, along with a V-adenosine triphosphatase, maintain the ionic gradient across the acidocalcisome membrane, which is necessary for acidocalcisome function, such as osmotic homeostasis upon passage of the parasite from the insect vector into the mammalian bloodstream (8). Knockdown/knockout of mPPases causes severe reduction in polyphosphate and the loss of acidocalcisome acidity, leading to the failure of the parasites to stabilize their intracellular pH upon exposure to external basic pH (18). In addition, it has been shown that mPPase is essential for virulence of *T. gondii* in a mouse model (19). mPPases are thus promising targets for structure-based drug design because they are essential and do not exist in multicellular animals and structures in different conformations are known.

Nonhydrolyzable PP<sub>i</sub> analogs are not viable mPPase drug candidates, as they will inhibit the many human enzymes that produce or hydrolyze PP<sub>i</sub>, from inorganic pyrophosphatases to polymerases and ectonucleotide pyrophosphatases/phosphodiesterases (20). Our goal

<sup>1</sup>Research Program in Molecular and Integrative Biosciences, University of Helsinki, Helsinki, Finland. <sup>2</sup>Drug Research Program, Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland. <sup>3</sup>Malaria Research Laboratory, Immunobiology Research Program, Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki, Helsinki, Finland. <sup>4</sup>School of Biomedical Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK.

\*Corresponding author. Email: a.goldman@leeds.ac.uk



**Fig. 1. Overview of the TmPPase structure.** (A) Monomer showing the location of the hydrolytic center, coupling channel, ion gate, and exit channel. (B) Top view of the superposition of the TmPPase:IDP:ATC (wheat) and TmPPase:IDP complex (cyan) structure showing relative TMH movements (arrow) upon binding of ATC. (C) Superposition of the gate region between two structures [TmPPase:IDP:ATC (wheat) and TmPPase:IDP complex (cyan)]. D246<sup>6.53</sup>, D703<sup>16.46</sup>, and Na<sup>+</sup> slightly moving away (arrow) relative to their positions in the TmPPase:IDP structure. Violet-purple and pink spheres are for Na<sup>+</sup> of TmPPase:IDP and TmPPase:IDP:ATC, respectively.

is thus the discovery of drug-like small molecular weight compounds that would specifically inhibit mPPases from parasitic protists (13). Here, we report the discovery and synthesis of the first nonphosphorus, non-substrate-analog inhibitor of TmPPase, *N*-[(2-amino-6-benzothiazolyl)methyl]-1*H*-indole-2-carboxamide (ATC), its preliminary structure-activity relationships (SARs), and a 3.4 to 4.0 Å resolution structure of ATC bound to TmPPase. It inhibits TmPPase with low micromolar affinity. The binding site of ATC is allosteric and furthermore asymmetric with respect to the structural homodimer, providing strong evidence that the substrate binding and catalysis events proceed asymmetrically.

## RESULTS

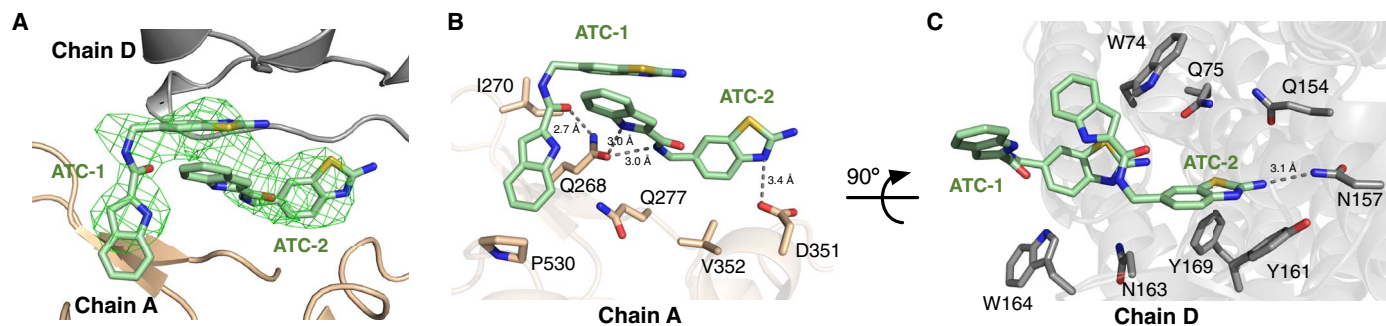
### Structure overview

We solved the 3.4 to 4.0 Å resolution of the TmPPase structure in complex with IDP and ATC (TmPPase:IDP:ATC) by molecular replacement using the TmPPase:IDP structure [Protein Data Bank (PDB) ID: 5LZQ] (10) as the search model but with the IDP removed. [We quote two numbers for the resolution, as the data are quite anisotropic, as shown by the STARANISO server (<http://staraniso.globalphasing.org>) with a maximum resolution of 3.6 Å along the *a*\* axis, 3.4 along *c*\*, but slightly worse than 4.0 Å along *b*\* (table S1); we followed current best practice and used all of the data above a local  $I/\sigma(I)$  cutoff of 3 (21).] Molecular replacement identified four TmPPase molecules (two dimers) in the asymmetric unit. After the first round of refinement, we observed positive ( $F_o - F_c$ ) density at 3.0 σ at the hydrolytic center (fig. S1A) in all four chains and at the ionic gate (fig. S1B) in chains A, B, and D. We built the Mg<sub>5</sub>IDP complex into the former and a sodium ion into the latter. There was positive ( $F_o - F_c$ ) density between atom O3 of IDP and A495 in chains A, B, and D after placement of the Mg<sub>5</sub>IDP

complex. We fit K<sup>+</sup> into the density (fig. S1A) as this is its known binding site (11).

We also observed positive 3.0 σ ( $F_o - F_c$ ) density that was shaped like two connected pancakes near the exit channels of chains A and C. Fitting the density near chain A with ATC in combination with different ions and water molecules (fig. S1, E to J) showed that it was best described as two ATC molecules in a head-to-tail orientation (Fig. 2A and fig. S1E). These ATCs (ATC-1 and ATC-2) are located at the interface between chains A and D (Fig. 2A). However, only one ATC (ATC-3) could be located at the interface between chains B and C as the density in these chains was weaker. After refinement, manual rebuilding, and manual placement of the IDP and ATC into the electron density, we were able to refine it to an  $R_{\text{work}}/R_{\text{free}}$  of 22.8%/28.4% with acceptable stereochemistry (table S1). Unlike previously solved TmPPase structures (9, 10), there are four molecules (two dimers, AB and CD) in the asymmetric unit (fig. S2A). The root mean square deviation (RMSD) per Cα is 0.31 Å<sup>2</sup> for the AB dimer versus the CD dimer and 0.26 to 0.33 Å<sup>2</sup> when comparing individual monomers (table S2).

There are thus two dimer interfaces in the crystal structure: the AB interface described before (9, 10) and ATC-mediated AD and BC dimers (fig. S2). Are these latter dimers physiologically relevant? It is fairly clear that the answer is no. First, the interactions are weak for a protein interface (249 Å<sup>2</sup> buried) and the proteins interact tail to tail (fig. S2B), placing the extracellular loops together in a manner that would not allow a bilayer to form. Second, size exclusion chromatography–multiangle laser light scattering (SEC-MALLS) analysis showed that the addition of ATC to the TmPPase:IDP complex (158.9 ± 0.1 kDa) did not change the oligomerization state of the protein (162.8 ± 3.7 kDa for the TmPPase:IDP:ATC complex) (fig. S3). The ATC pair binding near A interacts structurally as a dimer. In the “Kinetics of ATC



**Fig. 2. ATC interactions with TmPPase domains.** (A) ATC [positive ( $F_o-F_c$ ) electron density map at  $3\sigma$ ] located on the interface of chains A (wheat) and D (gray) in a hydrophobic cleft. Pale green, carbon; blue, nitrogen; red, oxygen; yellow, sulfur. (B) Residues in chain A that are important for the interactions with the ATC dimer. The closest approach of the two ATC molecules is 3.3 Å, consistent with  $\pi$ - $\pi$  stacking. (C) Residues on chain D that are important for the interactions with the ATC dimer. Dashed lines represent hydrogen bonds.

binding” section, we show that ATC binds functionally as a dimer to both the  $E_2S$  and  $E_2S_2$  complexes in solution.

The overall structure of the TmPPase:IDP:ATC complex is very similar to that of the TmPPase:IDP complex, with an RMSD/ $C\alpha$  of 0.50 and 0.43 Å to dimers AB and CD, respectively. (In what follows, all structural alignments were made with respect to monomer A unless otherwise stated.) Although the volume of the hydrolytic pocket in TmPPase:IDP:ATC is very similar to that of TmPPase:IDP (707 Å<sup>3</sup> versus 720 Å<sup>3</sup>), on the basis of the structural alignment on TMH13 and TMH14 (residues 542 to 629), there are slight clockwise movements on the cytoplasmic side of the protein helices relative to the TmPPase:IDP structure (Fig. 1B). This causes a slight movement of the  $Mg_5$ :IDP complex in the same direction as the helices. In addition, at the gate of the TmPPase:IDP:ATC structure,  $Na^+$  is pentacoordinated by D243<sup>6.50</sup>, E246<sup>6.53</sup>, S247<sup>6.54</sup>, and D703<sup>16.46</sup> (bidentate to both carboxylate oxygens) (Fig. 1C). In TmPPase:IDP, the coordinating residues are D243<sup>6.50</sup> (side chain and main chain), E246<sup>6.53</sup>, S247<sup>6.54</sup>, and D703<sup>16.46</sup>, and the D243<sup>6.50</sup> C=O— $Na^+$  distance is 2.8 Å versus the considerably longer 4.0 Å in TmPPase:IDP:ATC. This is mostly because the sodium ion appears to be translated by about  $\sim 1.7$  Å. We believe that these structural changes are functionally significant and related to the binding of ATC (see the “Asymmetry” section below).

### ATC binds in a cleft beside the exit channel

During refinement, we observed extra density near chains A and D (Fig. 2A), where we were able to model ATC. ATC is a rigid multiring system with only two free torsion angles, which we modeled in trans geometry, the most prevalent in the solution [Supplementary Materials, nuclear magnetic resonance (NMR) spectra]. We could thus (Fig. 2A) place two molecules of ATC into the ( $F_o-F_c$ ) electron density map. There is ( $F_o-F_c$ ) density in an equivalent region near chains B and C, where a single ATC could be fitted, but the density there is too poor to fit an ATC dimer, presumably due to molecular disorder in the crystal. We thus focus on ATC-1 and ATC-2, which bind to chains A and D, as representing the TmPPase<sub>2</sub>(Mg<sub>5</sub>PPi)<sub>2</sub>(ATC<sub>2</sub>)<sub>1</sub> complex, where the ATC dimer binds asymmetrically to TmPPase.

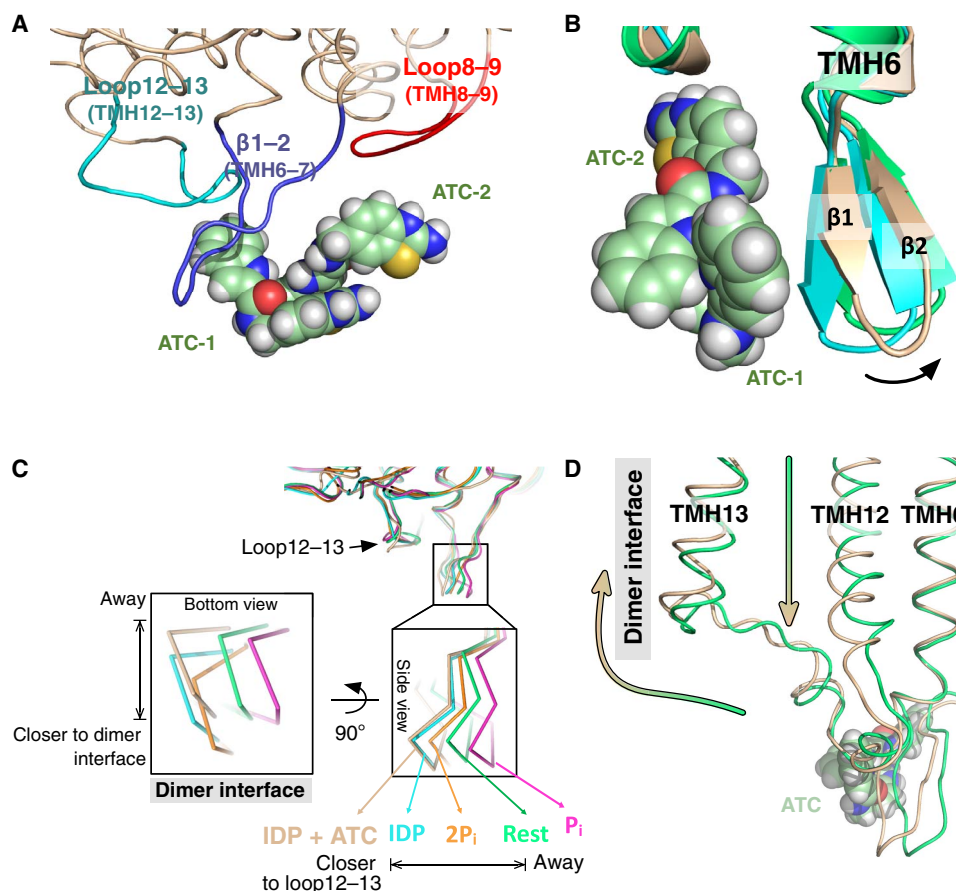
Two molecules of ATC  $\pi$ -stack head to tail (Fig. 2 and fig. S1D) and bind in a hydrophobic cleft. The cleft is formed by strands  $\beta$ 1–2 (loop6–7), loop8–9, and loop12–13, near the exit channel of chain A, with loop2–3 and loop4–5 on chain D (Fig. 3A). (A similar cleft forms between chains B and C.) In terms of buried surface area, ATC-1 buries 112 Å<sup>2</sup> on chain A and 85 Å<sup>2</sup> on chain D, while ATC-2 buries 85 Å<sup>2</sup> on chain A and 92 Å<sup>2</sup> on chain D (table S3).

ATC has a 2-aminothiazole group, which is a weak base [ $pK_a$  (where  $K_a$  is the acid dissociation constant),  $4.23 \pm 0.03$ ; (22)] reported in both neutral and protonated forms in the PDB and CSD (Cambridge Structural Database). We modeled the 2-aminothiazole groups in the protonated state because both ATCs in the AD interface form ion pairs/hydrogen bonds with chain A (Fig. 2B and table S3). Chain A Q268 seems to be the key interacting residue. It forms H bonds to both N11 and N12 of ATC-2, and to O01 of ATC-1 (see fig. S1D for atom numbering of ATC), suggesting how the ATC dimer is stabilized on the enzyme. Chain A D351 (O $\delta$ 1) forms a salt bridge with 2-aminothiazole on ATC-2 (Fig. 2B and table S3). In chain D, on the other hand, the chief interactions do not seem as strong: A single hydrogen bond is formed between ATC-2 N19 and N157 O $\delta$ 1, and the remaining interactions are the hydrophobic stacking of W164 on ATC-1, and the indole of ATC-2 sandwiched between W74 and W164 and stacking on Y161 (Fig. 2C and table S3). These interactions with chain D are probably due to crystal packing, as the binding of ATC to the protein did not change the protein oligomerization in the solution (fig. S3). We therefore focus on the interactions of the ATC dimer with chain A.

### Asymmetry

Artukka and co-workers (23) recently demonstrated functional asymmetry in  $K^+$ -dependent mPPases in the presence of excess potassium ion. Here, we show the first structural evidence of asymmetry. The IDP-bound state has the most symmetrical structure (the lowest RMSD values) between each monomer compared to others (table S4 and fig. S4A). However, the binding of ATC makes the IDP-bound enzyme asymmetric, especially in the loops that interact with the inhibitor (loop8–9,  $\beta$ 1–2, and loop12–13): The RMSD between the  $\beta$ 1–2 loop in chain A (ATC bound) and chain B (no ATC) is 1.25 Å (table S4). Structural alignment between monomers of other states (the resting state, 2P<sub>i</sub>-bound state, and WO<sub>4</sub><sup>2-</sup> bound state) shows that these loops differ far less (table S4). In addition, there is a change in the binding mode of  $Na^+$  at the ionic gate (see above).

ATC binds as a dimer in chain A, holding the  $\beta$ 1–2 loop (Q268, K269, I270, and Q277), loop8–9 (D351 and V352), and loop12–13 (P530) together (Fig. 3A and fig. S4, B and C). Comparison of the structure of the loops interacting with ATC with the equivalent loops in TmPPase:IDP showed that the loops  $\beta$ 1–2 and 12–13 in chain A move the most, with an RMSD per  $C\alpha$  of 1.13 and 1.09 Å, respectively (table S5), because these loops interact with the ATC dimer (Fig. 3B and table S3). In contrast, the RMSDs/ $C\alpha$  for the other loops that interact with ATC is smaller (table S5). Comparing all the TmPPase



**Fig. 3. Loop movement caused by the interactions with ATC.** (A) ATC dimer interaction with  $\beta 1-2$  strand (deep blue), loop8-9 (red), and loop12-13 (cyan). ATC is depicted in space-filling model (pale green, carbon; blue, nitrogen; red, oxygen; yellow, sulfur). (B) Superposition of TmPPase:IDP (cyan) with the TmPPase:IDP:ATC structure (wheat) and TmPPase:Ca:Mg (green) showing the large side movement of  $\beta 1-2$  strand (arrow). ATC is depicted as in (A). (C) Superposition of TmPPase structures in all different states showing the closer/away movement of  $\beta 1-2$  strand relative to the loop12-13 and the dimer interface of the protein. Coloring scheme from left to right: pink, phosphate analog ( $\text{WO}_4$ )-bound state (TmPPase:  $\text{WO}_4$ ); green, resting state (TmPPase:Ca:Mg); orange, two-product-bound state (TmPPase:2P<sub>i</sub>); cyan, IDP-bound state (TmPPase:IDP); wheat, IDP:ATC-bound state (TmPPase:IDP:ATC). (D) Superposition of TmPPase structures showing the downward movement of TMH12 and away movement of loop12-13 from the resting state (green) to the IDP:ATC-bound form (wheat) of TmPPase. The green-to-wheat colored arrow shows the movement.

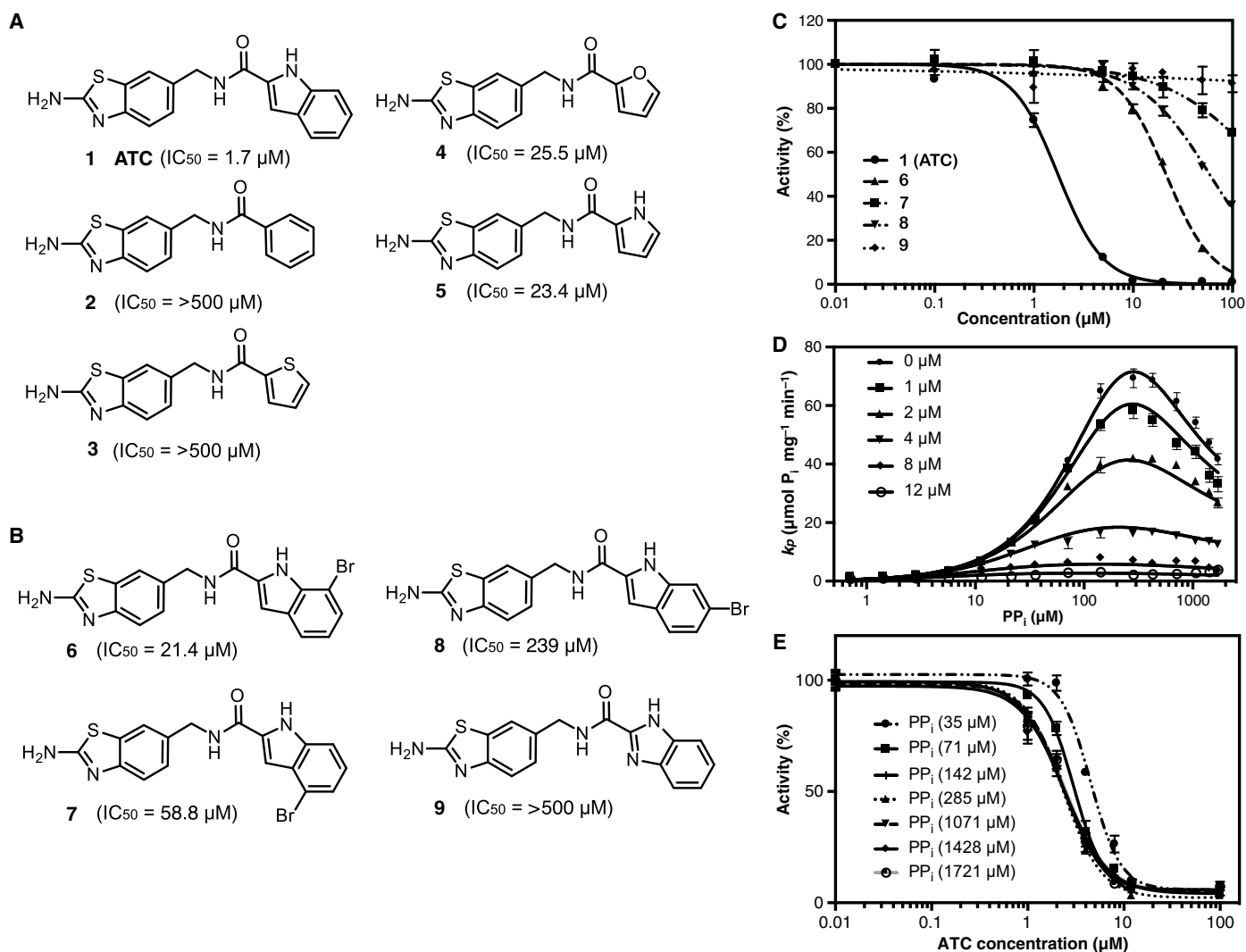
states shows that the binding of ATC causes the  $\beta 1-2$  strand of chain A to move away from the dimer interface relative to its position in TmPPase:IDP and TmPPase:2P<sub>i</sub> to roughly the same angle as in the resting and P<sub>i</sub> structures (Fig. 3C, inset bottom view; periplasmic side)—when the exit channel must remain closed at all times. The movement of chain A  $\beta 1-2$  due to the binding of the ATC dimer thus creates a hydrophobic clamp that locks the TmPPase exit channel in a closed state after substrate binding. In contrast, it appears that ATC does not bind to chain B because the  $\beta 1-2$  strand has moved away from loop12-13, closer to its position in the resting state (table S4 and fig. S5), and has become disordered. ATC appears not to bind because Q268 and Q277, key elements in the binding site, have become disordered. The same disordering, and lack of binding of ATC, is seen in monomer D.

### Pharmacology of ATC and analogs at TmPPase

We synthesized not only ATC but also four analogs (compounds 2, 3, 4, and 5) (Fig. 4A) with different ring structures. Four other analogs (compounds 6, 7, 8, and 9), three of them brominated, were also designed for structural studies. The median inhibitory concentra-

tions ( $\text{IC}_{50}$ s) of all the compounds were at least 20 times worse than the original hit ATC (Fig. 4, B and C), and they did not help us solve the structure. They do, however, provide initial SARs for this family of compounds on TmPPase.

The suggested binding mode explains the SARs of the ATC analogs (Fig. 4). First, they show that the hydrogen bonding functionality of the indole ring is important; compounds 2 and 3 that lack this functionality are inactive. In the TmPPase:IDP:ATC structure, Q268 near the indole nitrogen of ATC explains not only the lack of activity of compounds 2 and 3 but also the tolerance of both hydrogen bond donor (nitrogen; compound 5) and acceptor (oxygen; compound 4) functionalities in this position. Second, the aromatic nature of the indole ring seems important for activity: Compounds 4 and 5 that include suitable hydrogen bonding functional groups but not a bulky ring structure are approximately 10-fold less active than ATC. The indole rings of ATC-1 and ATC-2 form  $\pi-\pi$  stacking interactions with each other: Removing the benzene ring weakens this interaction. Third, compounds 6, 7, and 8 with bromine substitutions are 10- to 100-fold weaker binders than ATC, suggesting the importance of the unsubstituted indole ring. The bromine substitutions may



**Fig. 4. Compound library and inhibition activity against TmPPase.** (A) Screening hits. (B) Analogs developed for structural studies. (C) Inhibition curve of ATC, compounds 6 to 9. (D) Kinetic inhibition plot for the inactivation of TmPPase at six different concentrations of ATC. The solid lines represent nonlinear regression fits with a residual SD ( $Sy_x$ ) of 1.8. (E) Concentration-response curve of TmPPase at different ATC and substrate concentrations. The curves are the same for all concentrations of PP<sub>i</sub> greater than 100  $\mu\text{M}$  (i.e., with significant amounts of the substrate bound) but differ at very low concentrations of the substrate. The enzymatic activity was related to the activity in the absence of the inhibitor (100%), and the background was measured from the signal in the absence of the enzyme. All data are shown as mean  $\pm$  SD with  $n = 3$  replicates.

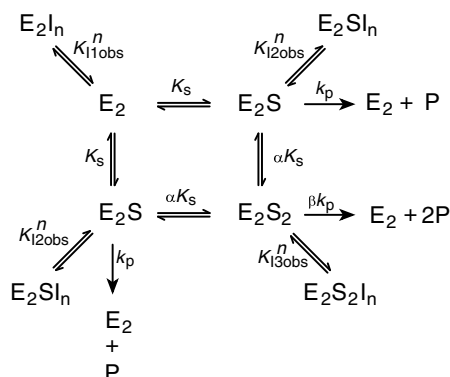
weaken the  $\pi$ - $\pi$  stacking interactions by altering the shape and location of the  $\pi$ -electron cloud. However, there are also direct clashes with loop6-7 and loop12-13, which are key sites of interaction (see above). In particular, the weakest brominated compound 8 would clash with P530 in loop12-13 (ATC C08-P530: 3.1 to 3.3 Å), while brominated compound 7 would clash with the K269 side chain (ATC C06-K269: 3.8 Å). Last, benzimidazole substitution instead of an indole yields fully inactive compound 9. This is probably due to the loss of the  $\pi$ - $\pi$  head-to-tail stacking: Both the 2-aminothiazole and the benzimidazole groups are protonated at physiological pH, so they would repel each other.

### Kinetics of ATC binding

As ATC is a potent inhibitor, we further characterized its effect on the rate of substrate (PP<sub>i</sub>) hydrolysis using a range of ATC concentrations (0.0 to 12.0  $\mu\text{M}$ ). We performed the kinetic assay using PP<sub>i</sub> concentra-

tions from 0 to 1714  $\mu\text{M}$  at 71°C with a single-point measurement at 2 min (Fig. 4D), having shown that this is within the linear range for initial rates (fig. S7). The shape of these curves was unexpected, so we performed a Hill analysis (Fig. 4E and table S6), which showed that, at substrate concentrations greater than 100  $\mu\text{M}$  (i.e., when the substrate is bound; see below), the Hill coefficient for ATC binding is 2.

We therefore performed a simultaneous analysis of the data for all inhibitor and substrate concentrations using the model in Fig. 5 with Eq. 1 and  $n = 2$  (Table 1), with the exception that  $K_{11}^2$  is not fit (when we included it, the residuals did not improve, and its mean value is  $\geq 0.25$  mM; data not shown) (Fig. 4D). The model is a variant on that used by Artukka *et al.* (23) and is required as there is half-of-the-sites reactivity. The sum of squares when fitting the inhibition curves with an ATC monomer (i.e.,  $n = 1$  in Fig. 5) is 5268, and the SD of the residuals ( $Sy_x$ ) is 4.3; when fitting an ATC dimer, the values are 855 and 1.8, respectively. The model shown (Figs. 4D and 5) is the simplest that fits



**Fig. 5. Substrate binding, hydrolysis, and inhibition at two active sites of a dimeric TmPPase.**

**Table 1. Kinetic parameters of TmPPase activity and ATC inhibition constant.**

Kinetic parameter	Value
$k_p$	$235 \pm 28 \mu\text{mol P}_i \text{ mg}^{-1} \text{ min}^{-1}$
$\beta$	$0.05 \pm 0.005$
$\beta k_p$	$12.2 \pm 1.03 \mu\text{mol P}_i \text{ mg}^{-1} \text{ min}^{-1}$
$K_s$	$666 \pm 110 \mu\text{M}$
$\alpha$	$0.12 \pm 0.05$
$\alpha K_s$	$80.7 \pm 19.3 \mu\text{M}$
$K_{i2\text{obs}}$	$1.37 \pm 0.09 \mu\text{M}$
$K_{i3\text{obs}}$	$3.18 \pm 0.15 \mu\text{M}$

the data: Other models do not fit (data not shown). We did not endeavor to fit sequential binding of ATC to the TmPPase dimer as the data are insufficient for the purpose. This means that ATC binds as a dimer to TmPPase. ATC binds tightest to the  $E_2S$  complex ( $K_{i2\text{obs}} = 1.37 \pm 0.09 \mu\text{M}$ ): Its binding to  $E_2S_2$  is about 2.5-fold weaker ( $K_{i3\text{obs}} = 3.18 \pm 0.15 \mu\text{M}$ ), and it does not bind to  $E_2$  at concentrations we are able to achieve (Table 1). ATC is thus an uncompetitive dimeric inhibitor with a clear preference for binding to the single-substrate bound form.

Our data also show that, after one monomer binds the substrate to form the  $E_2S$  complex, the other monomer binds the substrate much more tightly: The value of  $\alpha$  is very low (0.12) (Table 1). However, the hydrolysis rate of the  $E_2S_2$  complex is 20-fold slower ( $\beta = 0.05$ ) compared to the  $E_2S$  complex ( $k_p = 234.5 \pm 27.1 \mu\text{mol P}_i \text{ mg}^{-1} \text{ min}^{-1}$ ) (Table 1). In comparison with earlier reports (23), the binding behavior is different but the ratio of hydrolysis rates is similar (see Discussion). In the crystal structure we solved, ATC is bound to the  $E_2S_2$  form, not to the  $E_2S$  form. The concentration of ATC added for crystallization (1 mM) was higher than the binding constant of ATC to the  $E_2S_2$  form (Table 1), and ATC binding to the  $E_2S$  form may induce a conformational change in the other monomer that allows further substrate binding (see Discussion). This is consistent with the fact that we could only crystallize the TmPPase:ATC

complex in the presence of IDP. ATC is thus the first nonphosphorus inhibitor and the first allosteric inhibitor identified for mPPase.

### Is ATC a viable lead against parasitic mPPases?

ATC would be a potential lead as an antiparasitic agent if it were active against parasitic protists such as *Plasmodium* spp. Preliminary studies with isolated *P. falciparum* membranes showed that ATC had an  $\text{IC}_{50}$  of around  $80 \mu\text{M}$  (fig. S6A), i.e., at least 40-fold lower than in our test system *T. maritima* (Fig. 4A). In the *P. falciparum* survival assays, ATC did not show any anti-plasmodial activity (fig. S6B). On the other hand, knockout of mPPases in *T. gondii* (19) and *T. brucei* (24) demonstrate that functional mPPase is required for survival and virulence of these parasites.

To explain these findings and study the potential specificity determinants at the binding sites of ATC-1 and ATC-2 across mPPases, we retrieved the sequences of 16 mPPases from parasitic protists as well as representatives of all the other mPPase families (fig. S8). The areas corresponding to loop6-7 (i.e.,  $\beta$ 1-2 strand), loop8-9, and loop12-13 showed a high degree of difference between TmPPase and the other mPPases (Fig. 6A and fig. S8). In particular,  $\beta$ 1-2 (loop6-7) is missing in parasite mPPases. As the binding site of ATC-1 and ATC-2 is located on these loops (6-7, 8-9, and 12-13 of chain A), ATC could not bind the mPPases of pathogenic parasites as it does TmPPase. To validate this finding, we built a homology model of *P. falciparum* mPPase (PfPPase; Fig. 6B). Loop6-7 and loop12-13 of the modeled structure are notably shorter than those of TmPPase: The ATC binding site is not present, as loop6-7 contributes about two-thirds of the interface (table S3). This not only explains why ATC does not inhibit PfPPase but also supports the notion that it functions asymmetrically, preventing a full catalytic cycle from occurring by binding to one chain and preventing the exit channel from opening.

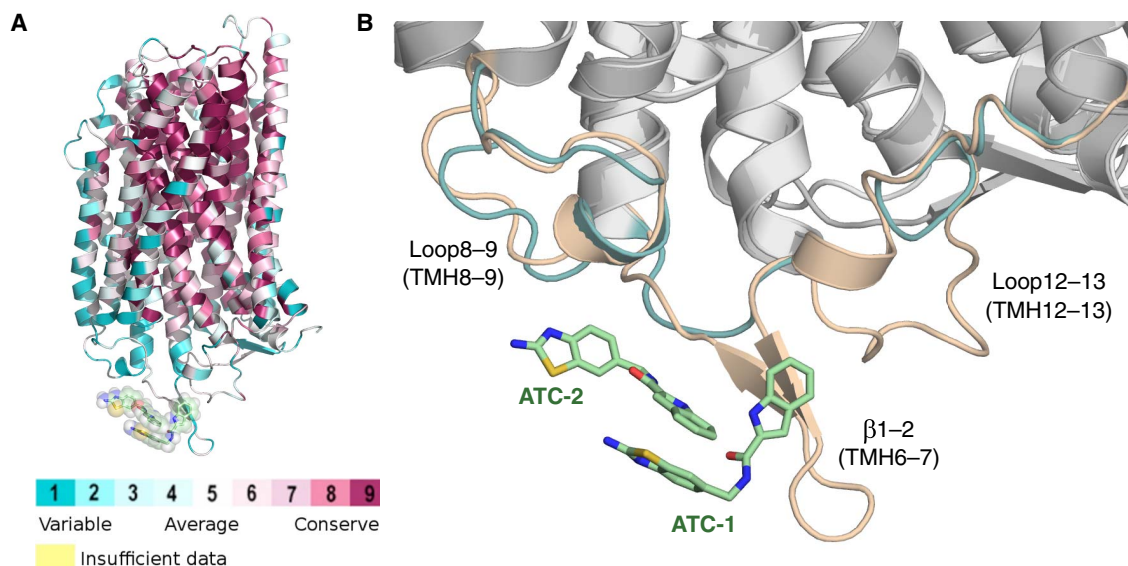
## DISCUSSION

### The first novel inhibitors of integral membrane pyrophosphatases

mPPases are potential targets for anti-protozoan drug design as they do not occur in humans and other mammals, are essential in *T. brucei* (24), and render *T. gondii* noninfectious (19). Bisphosphonate derivatives (PP<sub>i</sub> analogs) can inhibit mung bean mPPase (25) and prevent the proliferation of *T. gondii* (26) and other parasitic protists (27, 28). However, they also inhibit human enzymes such as farnesyl pyrophosphate synthase, so they are not viable as anti-parasitic drug candidates [for a review, see (13)]. This study reports the first nonphosphorus inhibitors of mPPase, which were identified through a screening process. The best compound (ATC) binds and inhibits both the  $E_2S$  and  $E_2S_2$  complexes, with inhibition constants  $K_{i2\text{obs}}$  and  $K_{i3\text{obs}}$  of 1.37 and 3.18  $\mu\text{M}$ , respectively (Table 1). The overall  $\text{IC}_{50}$  is 1.7  $\mu\text{M}$ . It appears to act in an allosteric manner. This mode of inhibition has been observed for an inhibitor of soluble PPase from *Mycobacterium tuberculosis* (MtsPPase) (29), but to our knowledge, this is the first one observed in mPPase.

### Identification of alternate inhibitory mechanisms and evidence for allostery in mPPases

ATC binds beside the exit channel (Figs. 2 and 3) and acts in allosteric manner. Similar types of inhibition occur in GPCRs (G protein-coupled receptors) and many other membrane proteins. For instance, the voltage-dependent  $K_v$ 11.1 (hERG) potassium channel that mediates the repolarizing current in the cardiac action potential is easily



**Fig. 6. The conservation of ATC binding site among mPPases.** (A) Structure of TmPPase colored according to the sequence conservation among the set of 16 pathogenic mPPases. The pale green surface indicates the location of the ATC binding site. (B) Comparison of the ATC binding sites of the crystal structure of TmPPase (wheat) and a homology model of PpPPase (teal). Carbon atoms of ATC are presented as pale green sticks. Blue, nitrogen; red, oxygen; yellow, sulfur.

inhibited by a wide variety of drugs, leading to drug-induced long QT syndrome and lethal arrhythmia. The molecules, such as quinolone and macrolide antibiotics, bind to hydrophobic pockets on the side of the selectivity filter (30, 31). Recently, a structural study found inhibitors that targeted a hitherto-unknown allosteric transmembrane site in MsbA, an adenosine triphosphate (ATP)-binding cassette (ABC) transporter, by locking MsbA in an inward facing conformation that prevents transition to the outward-facing conformation during lipopolysaccharide transport (32). This binding uncoupled the nucleotide-binding domains, which adopt an asymmetrical conformation.

Similar interactions to the  $\pi$ -stacked ATC dimer that binds to TmPPase have been observed in soluble protein-inhibitor complexes as well. For example, Stornaiuolo *et al.* (33) found an isoquinoline derivative that formed a  $\pi$ -stack of three identical molecules that bound to one binding site of acetylcholine-binding protein. A quinazoline derivative formed a  $\pi$ -stack of five molecules to the lysine-specific demethylase 1 (LSD1)-RE1-silencing transcription factor corepressor (CoREST) complex (34), and a phenylpyrrole derivative forms a head-to-tail  $\pi$ -stacked dimer that binds in the twofold cavity in soluble hexameric MtsPPase. This inhibitor appears to act uncompetitively with the binding site 20 Å from the nearest pyrophosphate (29). The asymmetric binding of the ATC dimer to one monomer of TmPPase is similar in kind.

A recent study has shown that mesophilic mPPases show substrate inhibition with increased  $K_m$  and decreased  $V$  when two molecules of the substrate are bound in the presence of 50 mM  $K^+$  (23). Artukka and co-workers (23) observed that S binds more weakly to  $E_2S$  than to  $E_2$  and that  $k_{cat}$  for  $E_2S_2$  is lower than for  $E_2S$ . They interpreted this to mean that substrate binding to one monomer causes a conformational change in the other that inhibits its interaction with the substrate. For the enzymes they tested, the ratio of  $k_{catS}$  varies between 2.9 and 16 [table 4 of Artukka *et al.* (23)]; our ratio is 20 (Table 1). However, our data differ in that binding of the first substrate molecule appears to be much weaker than that of the second substrate, as  $\alpha$  is 0.12 (Table 1), which is the reverse of what they observed. There are two

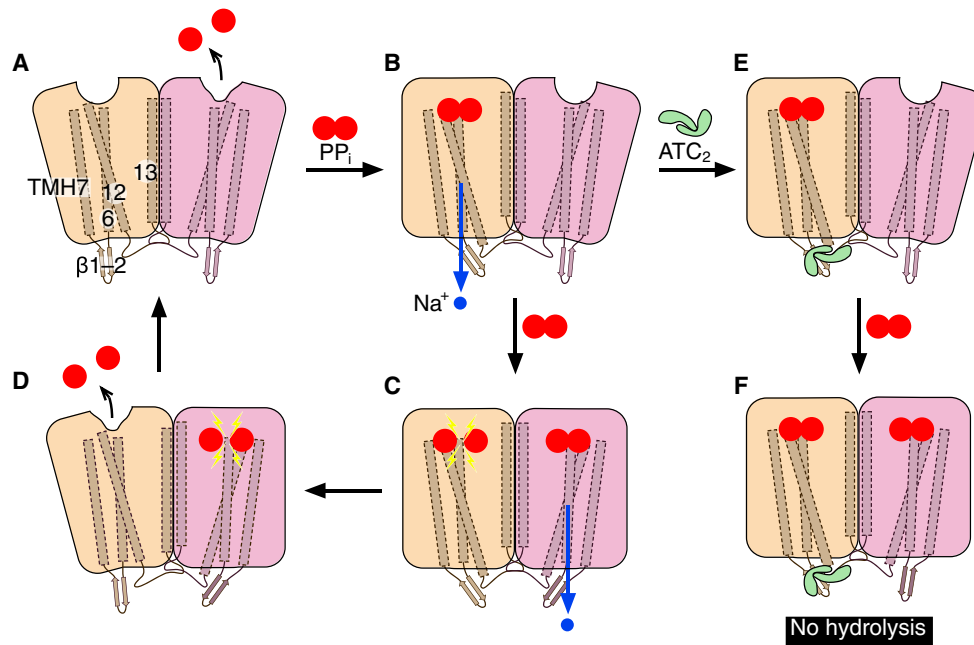
possible explanations: First, they did not study TmPPase, but a range of mesophilic  $K^+$ -dependent mPPases. Second, and we think more likely, our experiments are performed near the temperature optimum of TmPPase (71°C) but in the presence of *n*-dodecyl- $\beta$ -D-maltopyranoside (DDM), not of the bidentate membrane-spanning lipids found in *T. maritima* (35). TmPPase in DDM micelles is almost certainly more conformationally flexible than it is in *T. maritima* lipids, which would weaken binding of the first substrate molecule. If, as both our and the Artukka *et al.* (23) data suggest, mPPases are allosteric, then binding of the first substrate molecule would stabilize the structure of the dimer, increasing the affinity for the second substrate molecule and thus explaining the apparent discrepancy.

However, none of this is germane to the interpretation of the ATC inhibition data. Our ATC-bound structure (Figs. 2 and 3) and inhibition data (Fig. 4, D and E; Table 1; and table S6) are similar: An ATC dimer is required to inhibit the TmPPase<sub>2</sub>(Mg<sub>5</sub>PP<sub>i</sub>)<sub>1-2</sub> dimer in the presence of the substrate. The second site is unavailable in the substrate-bound form, and neither monomer binds ATC before substrate binds. At very low substrate concentrations (Fig. 4E and table S6), the Hill coefficient is 3, suggesting that a third ATC is required to inhibit the enzyme at concentrations below 100  $\mu$ M, when the substrate binds poorly in our assay. We suggest that this is also due to the increased flexibility in DDM at 71°C: ATC binding can stabilize substrate binding.

### Correlation between functional and structural data

As the resolution of the structure is only 3.4 to 4.0 Å, how reliable is our interpretation of the structure? The data were sufficient to identify  $K^+$  in three of the four active sites. This is the first time that they have been observed in a TmPPase structure: The electron density maps are significantly better than 4.0 Å. In addition, the POLDER analysis (fig. S1, K to P) indicates that the ATC fit the density as expected at this resolution.

The link between the structural and functional data then rests on two independent lines of evidence. First, the functional data are consistent only with tight binding to the substrate-bound protein,



**Fig. 7. Catalytic scheme of TmPPase and its inhibition by ATC.** (A) The binding of the substrate (PP<sub>i</sub>; two joined red circles) to monomer A (wheat color) leads to Na<sup>+</sup> pumping (blue circle and blue arrow for the pumping direction) and induces a conformational change in monomer B (pink color). (B) This change increases the affinity of monomer B for the substrate. The Na<sup>+</sup> pumping causes substrate hydrolysis (yellow lightning symbol for this event) in monomer A, while the substrate binding in monomer B leads to its Na<sup>+</sup> pumping (C). (D) After hydrolysis, monomer A releases its hydrolysis products (2P<sub>i</sub>; red circles), while monomer B proceeds to the hydrolysis event, returning the enzyme to state A. The ATC dimer binds to its binding site after substrate binding and Na<sup>+</sup> pumping in monomer A (E), which locks this monomer in the “down” state. Although no hydrolysis can occur in monomer A due to the ATC binding, the conformational change may still induce a conformational change in monomer B, which increases its affinity for the substrate (F). The numbering with white background corresponds to TMH number.

and we could only obtain crystals with ATC bound in the presence of IDP. Second, the functional data indicate that ATC binds as a dimer, and this is what is observed. Close examination of the AB dimer in our structure further reveals that there are more differences between the two monomers than in any other TmPPase structure (table S4), particularly in loop5–6 and loop12–13. It is possible that the changes we see are due to crystal packing, not due to binding of ATC. However, as solution data show that binding of ATC causes changes that prevent catalysis, we suggest that these changes explain the lack of activity of the ATC-bound enzyme. The alternative hypothesis is that binding of an ATC dimer beside the exit channel in a way that closes it and affects the TmPPase:IDP dimer has nothing to do with observation of uncompetitive inhibition by an ATC dimer in solution. Our suggestion that the crystal and solution data are related is a more parsimonious explanation.

### Mechanism of action of ATC

If our model is correct, then how does ATC inhibit TmPPase? The binding of the dimer to chain A via loop6–7, loop8–9, and loop12–13 creates a hydrophobic clamp that locks the exit channel in the closed state and TMH12 in the “down” state. The binding appears to be driven by the multidentate interactions of Q268 with both ATC-1 and ATC-2. We posit that the mechanistic implications of these events are as follows.  $\beta$ 1 extends from TMH6, which contains D243<sup>6,50</sup>, D246<sup>6,53</sup>, and D247<sup>6,54</sup>, all of which coordinate Na<sup>+</sup> in the gate. In the ATC-bound structure, the position and coordination of the Na<sup>+</sup> change. In our current structure, D246<sup>6,53</sup>, D703<sup>16,46</sup>, and Na<sup>+</sup> move up to 1.7 Å from their positions in the TmPPase:IDP struc-

ture (Fig. 1C). This corresponds to a new state that, we believe, keeps the exit channel closed. Our current model (10) posits that, during ion pumping, TMH12 moves downward by at least 2 Å (9). Comparing the relative positions of the ATC-bound and resting states suggests that this movement of TMH12 makes loop12–13 move toward the dimeric interface of the protein and TMH13 move “up” toward the cytoplasmic side (Fig. 3D). This would induce equal conformational changes in monomer B in TmPPase:IDP. However, ATC binding locks monomer A TMH12 in the “down” state. Because of its “down” position, loop12–13 and TMH13 can still trigger the conformational changes of monomer B into a state that binds the substrate more tightly. However, further catalysis is impossible because monomer B cannot undergo a full catalytic cycle: The substrate can bind, but the full motions required for catalysis are impossible. Diagnostic of these motions is the conformation that leads to binding of ATC. In agreement with the structure, the kinetic data show that ATC dimer only binds to one of the monomers and binds most tightly to the TmPPase<sub>2</sub>(Mg<sub>5</sub>PP<sub>i</sub>) state (Table 1).

The structure also explains substrate inhibition and the requirement for potassium for maximal activity. The amino acid responsible for K<sup>+</sup> dependency (A495<sup>12,46</sup> in TmPPase) is located in TMH12, which is directly linked to the dimer interface (loop12–13 and TMH13). We suggest that, in the presence of K<sup>+</sup>, the downward movement of TMH12 upon substrate binding and ion pumping induces conformational changes in the monomer-monomer interface by the “away” movement of loop12–13 and an upward movement of TMH13 (Fig. 3D). This is then propagated to other TMHs. In the absence of K<sup>+</sup>, such coupling



does not occur. These changes increase affinity and activity. They also provide a possible structural explanation for the  $\text{Na}^+/\text{H}^+$  pumping enzymes (23, 36), where  $\text{H}^+$  pumping is not inhibited by 100 mM  $\text{Na}^+$  and 50 mM  $\text{K}^+$ . One way this might happen is through a conformation similar to that observed here: The binding of  $\text{PP}_i$  to monomer A and pumping of, say, a proton would drive changes in loop12–13 and TMH13 such that the conformation of monomer B became suitable for binding  $\text{PP}_i$  and pumping  $\text{Na}^+$ .

### Revised catalytic model

In conclusion, we propose the following model for TmPPase catalysis and its inhibition by ATC (Fig. 7). Substrate binding to monomer A leads to pumping in monomer A and induces a conformational change in monomer B, changing its affinity for  $\text{PP}_i$ . After  $\text{PP}_i$  binds in monomer B,  $\text{PP}_i$  hydrolysis occurs in monomer A and ion pumping in monomer B. The product of catalysis ( $2\text{P}_i$ ) is released from monomer A, allowing hydrolysis and subsequent product release in monomer B. The ATC dimer can only bind when  $\text{PP}_i$  is bound to monomer A. Nonetheless, the A: $\text{PP}_i$  form can still induce the conformational change in monomer B that allows  $\text{PP}_i$  binding. However, since the ATC binding locks monomer A in the down state, no further catalysis can occur.

Our novel structure of TmPPase with IDP and ATC has opened up new routes to drug discovery of neglected diseases by serendipitously demonstrating that there are alternative ways to inhibit mPPase. Although the ATC binding interface is not preserved (Fig. 6 and fig. S8), our data nevertheless indicate that binding at the exit channel can inhibit mPPases. The work shows how extracellular loop binding regions can have significant effects on intracellular substrate binding and vice versa. Last, by stabilizing asymmetric complexes, ATC provides a route into determining other steps along the reaction pathway and thus deriving a full catalytic model—one that may unify the  $\text{Na}^+/\text{H}^+$  and  $\text{Na}^+$  mPPases, as outlined above.

## MATERIALS AND METHODS

### Discovery of ATC and its analogs

The screening process that led to the discovery of compound **1** (ATC) will be described in more detail elsewhere. Briefly, we tested compounds using a recently developed 96-well plate assay with TmPPase as a model system (37). The screening collections included compounds synthesized in our laboratories. The identification of ATC as an inhibitor was followed by the test of analogs **2** to **5**. ATC and its analogs were initially developed for their antimicrobial, antiviral, and anticancer activities, as reported elsewhere (38).

### Compound synthesis

The purity of all the compounds tested was above 95%. The synthesis of ATC (compound **1**) and compounds **2** to **5** has been described elsewhere (38). The synthesis of compounds **6** to **9** is described in the Supplementary Materials. For synthesis, all reactions were carried out using commercially available starting materials (Sigma-Aldrich, Schnellendorf, Germany; BioFine International Inc., Vancouver, Canada; Fluorochem, Hadfield, UK) and solvents without further purification. Column chromatography was performed with an automated Isolera One high-performance flash chromatography system (Biotage, Uppsala, Sweden) using a 0.1-mm path length flow cell ultraviolet (UV) detector/recorder module (fixed wavelength, 254 nm). Analytical thin-layer chromatography was carried out using 0.2-mm silica gel plates (silica gel 60, F<sub>254</sub>; Merck KGaA, Darmstadt, Germany). NMR spectra (<sup>1</sup>H NMR and <sup>13</sup>C NMR) were

recorded on a Bruker Ascend 400 spectrometer (Bruker Corporation, Billerica, MA, USA). <sup>1</sup>H NMR was measured at 400 MHz, and <sup>13</sup>C NMR was measured at 100 MHz. High-resolution mass spectra (HRMS) were measured on a Waters Synapt G2 mass spectrometer (Waters Corporation, Milford, MA, USA) and reported for the molecular ions  $[\text{M} + \text{H}]^+$ .

### Protein expression and purification

Expression and purification of TmPPase are described elsewhere (39, 40). In brief, His-tagged TmPPase in pRS1024 plasmid under the control of the PMA1 promoter was freshly transformed into *Saccharomyces cerevisiae* strain BJ1991. The cells were grown overnight in 250 ml of selective synthetic complete dropout medium and then added to 740 ml of 1.5× YP medium containing 2% glucose. The cells were further grown at 30°C for 8 hours, collected by centrifugation (4000 rpm, 10 min), and lysed at 4°C using a bead beater with 0.2-mm glass beads. The membrane fraction was collected by ultracentrifugation (100,000g, 45 min), and the pellets were resuspended in buffer containing 50 mM MES-NaOH (pH 6.5), 20% (v/v) glycerol, 50 mM KCl, 5.2 mM MgCl<sub>2</sub>, 1.33 mM dithiothreitol (DTT), pepstatin A [2 μg ml<sup>-1</sup> (w/v)] (Sigma), and 0.334 mM phenylmethylsulfonyl fluoride (PMSF; Sigma). The membranes were solubilized using the “hot-solve” method (40) at 75°C for 1.5 hours in solubilization buffer [50 mM MES-NaOH (pH 6.5), 20% (v/v) glycerol, and 5.33% (w/v) DDM (Anatrace)]. After centrifugation to remove denatured proteins, KCl (to a final concentration of 0.3 M) and 2 ml of nickel-nitrilotriacetic acid (Ni-NTA) beads (Qiagen) were added and incubated at 40°C for 1.5 hours to each 40 ml of the solubilized protein and then loaded into an Econo-Pac column (Bio-Rad). Then, the column was washed with 2 × CV (column volume) of washing buffer [50 mM MES-NaOH (pH 6.5), 20% (v/v) glycerol, 50 mM KCl, 20 mM imidazole (pH 6.5), 5 mM MgCl<sub>2</sub>, 1 mM DTT, pepstatin A (2 mg ml<sup>-1</sup>, w/v), 0.2 mM PMSF, and 0.05% DDM (Anatrace)] and eluted with 2 × CV of elution buffer [50 mM MES-NaOH (pH 6.5), 3.5% (v/v) glycerol, 50 mM KCl, 400 mM imidazole (pH 6.5), 5 mM MgCl<sub>2</sub>, 1 mM DTT, pepstatin A (2 mg ml<sup>-1</sup>, w/v), 0.2 mM PMSF and 0.5% octyl glucose neopentyl glycol (OGNPG; Anatrace), 1 mM DTT, pepstatin A (2 mg ml<sup>-1</sup>, w/v), and 0.2 mM PMSF and 0.05% DDM (Anatrace)].

### Activity measurement and kinetic analysis

TmPPase activity, compound screening, and kinetic experiments were done using the molybdenum blue reaction method in a purified protein solubilized in DDM (37). Initial reaction rates of TmPPase were determined using  $\text{PP}_i$  as the substrate and ATC as the inhibitor at varying concentrations. The concentration of MgCl<sub>2</sub> and Na<sub>4</sub>PP<sub>i</sub> required to maintain 5 mM free Mg<sup>2+</sup> at pH 8.0 was approximated as described previously (41). The reaction was done in reaction buffer [60 mM tris-HCl (pH 8.0), 5 mM free Mg<sup>2+</sup>, 100 mM KCl, and 10 mM NaCl] at 71°C for 2 min. We chose a 2-min reaction time as it still produces a reaction product in a linear range (fig. S7).

For analysis, we used a variant of Artukka and co-workers’ method (23), which assumes that the two active sites in an mPPase dimer interact allosterically, with the addition of inhibitor binding to the two active sites (Fig. 5). E<sub>2</sub> represents the dimeric enzyme, S is the substrate (Na<sub>4</sub>PP<sub>i</sub>), I is the inhibitor (ATC), and n is the number of ATC molecules bound to the enzyme. K<sub>S</sub> is a microscopic Michaelis-Menten constant, and k<sub>p</sub> is the per-site maximum rate for the substrate complex. α is the factor relating the binding of the first substrate molecule (K<sub>S</sub>) to

the binding of the second,  $\beta$  is the factor relating the rate of hydrolysis of the  $E_2S$  complex ( $k_p$ ) to that of the  $E_2S_2$  complex, and  $K_{I1obs}$ ,  $K_{I2obs}$ , and  $K_{I3obs}$  are the observed inhibition constants for binding to the  $E_2$ ,  $E_2S$ , and  $E_2S_2$  complexes.

For the mechanism in Fig. 5 with  $n = 2$  (i.e., only inhibitor dimers bind), the rate Eq. 1 (42) is

$$v = \frac{(2k_p + 2\beta k_p [S]/\alpha K_s)}{2(1 + [I]^2/(K_{I2obs})^2) + (K_s/[S])(1 + [I]^2/(K_{I1obs})^2) + ([S]/\alpha K_s)(1 + [I]^2/(K_{I3obs})^2)} \quad (1)$$

where  $[S]$  and  $[I]$  correspond to substrate and inhibitor concentrations, respectively, and the equilibrium and Michaelis-type constants are as in Fig. 5. The kinetic data were fit together to obtain all parameter values and their SEs by nonlinear regression as a function of inhibitor (0, 1.0, 2.0, 4.0, 8.0, and 12.0  $\mu\text{M}$ ) and substrate (0 to 1714  $\mu\text{M}$ ) concentrations using Prism 6.0 (GraphPad Software). Eliminating the  $K_{I1}$  terms had no effect on the overall quality of fit.

### Crystallization, structure determination, and analysis

For crystallization, the purified protein was buffer-exchanged to the crystallization buffer [50 mM MES-NaOH (pH 6.5), 3.5% (v/v) glycerol, 50 mM KCl, 5 mM  $\text{MgCl}_2$ , 2 mM DTT, and 0.5% OGNPG] on a Micro Bio-Spin 6 column (Bio-Rad) and then diluted to a concentration of 10 mg  $\text{ml}^{-1}$ . Before crystallization, 1 mM IDP and 1 mM ATC were added, and the solution was incubated on ice for 10 min and centrifuged for 20 min (16,000g, 4°C). Crystallization trials were done at 22°C by sitting-drop vapor diffusion method using commercial screens, MemGold and PGA screen (Molecular Dimensions), in MRC two-well crystallization plates (Swissci) with a mosquito robot (TTP Labtech), and the drops were monitored using a minstrel DT UV imaging system (Formulatrix). Crystal hits appeared in the MemGold screen under different conditions. Crystallization conditions were then optimized to get better crystals by varying pH, buffer, salt, and polyethylene glycol (PEG) concentrations using the vapor diffusion method in 1  $\mu\text{l}$  + 1  $\mu\text{l}$  (protein–mother liquor) drops in a 24-well plate at room temperature. Harvestable crystals appeared within a week and were frozen after soaking in cryoprotectant (mother liquor + 20% glycerol) or directly from mother liquor. The best diffracting crystal appeared from a solution containing 0.1 M MES (pH 6.5), 0.1 M NaCl, 33% PEG-400, and 4% ethylene glycol.

X-ray diffraction data were collected at the European Synchrotron Radiation Facility, Grenoble (France), on the ID29 beamline at 100 K on a PILATUS 6M detector. One thousand four hundred images were collected at an oscillation angle of 0.1°. The data were merged and scaled using X-ray Detector Software (XDS) (43). Because of the anisotropy of the diffraction data, anisotropic cutoff and correction of the merged intensity data were performed on the STARANISO server (<http://staraniso.globalphasing.org/>) using the DEBYE and STARANISO programs. The structure was solved by molecular replacement with Phaser (44) using the IDP-bound state of the TmPPase structure (5LZQ) (10) as the search model. Phaser found a unique solution in space group  $P2_12_12_1$ , with four monomers in the asymmetric unit forming two dimers. The structure was built and refined using phenix.refine (45) and Coot (46). The final model had  $R_{\text{work}}/R_{\text{free}}$  of 22.8%/28.4%, with 95.7% of residues in the most favored and 4.3% in the allowed region of the Ramachandran plot (table S1).

The structural interface and assembly of the protein were analyzed using the PDBe PISA server ([www.ebi.ac.uk/pdbe/pisa/](http://www.ebi.ac.uk/pdbe/pisa/)) (47). Internal

pocket analysis was done using the SplitPocket server (<http://pocket.med.wayne.edu/patch/>) (48, 49). Structural alignments were done in PyMOL with the default setting using the monomer structure, except if stated differently. All structural figures were produced using PyMOL (<https://pymol.org/2/>).

### Comparison with pathogenic mPPases

The sequences of TmPPase and 16 pathogenic mPPases were aligned using Clustal Omega software (50), and conservation analysis was conducted and visualized with the ConSurf server (51, 52) with default parameters. The initial pairwise sequence alignment of TmPPase and *P. falciparum* mPPase (VP1; PfPPase; sequence identity, 42%) was produced with Discovery Studio 4.5 software (53) using the Align123 algorithm. The sequence alignment figure was generated with ESPript 3.0 (<http://esprict.ibcp.fr/ESPript/ESPript/index.php>) (54). Homology modeling was carried out with Modeller 9v14 (55) with default settings in the Discovery Studio platform. The pairwise alignment was revised during the course of model building; in total, nine models were constructed in three modeling rounds. The Modeller DOPE scores did not differ significantly among these models, and a representative model was selected to minimize strain in the loops connecting TMHs.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/5/eaav7574/DC1>

Fig. S1. Electron density maps of ligands in the TmPPase:IDP:ATC structure.

Fig. S2. Asymmetric unit of TmPPase:IDP:ATC.

Fig. S3. Oligomerization state of TmPPase upon the addition of inhibitors.

Fig. S4. Comparison of the ATC binding sites in different TmPPase structures.

Fig. S5. Differences in the ATC binding site between monomers of the TmPPase:IDP:ATC structure.

Fig. S6. Effect of ATC on growth of *P. falciparum* and on PfPPase.

Fig. S7. Linearity of TmPPase activity as a function of time and concentration.

Fig. S8. Multiple sequence alignment of mPPases from different organisms.

Scheme S1. Synthesis of analogs of **1** with heavy and basic atoms.

Table S1. X-ray data collection and refinement statistics.

Table S2. RMSD of TmPPase monomer in the asymmetric unit relative to monomer A.

Table S3. Major interactions of ATC with chains A and D.

Table S4. RMSD between chains A and B of loops in different TmPPase structures.

Table S5. RMSD of chain A of TmPPase:IDP:ATC loops to chain A of TmPPase:IDP loops.

Table S6. Hill constant of TmPPase inhibition by ATC at different substrate concentrations.

References (56–58)

### REFERENCES AND NOTES

- H. Baltscheffsky, L.-V. von Stedingk, H.-W. Heldt, M. Klingenberg, Inorganic pyrophosphate: Formation in bacterial photophosphorylation. *Science* **153**, 1120–1122 (1966).
- J. Karlsson, Membrane-bound potassium and magnesium ion-stimulated inorganic pyrophosphatase from roots and cotyledons of sugar beet (*Beta vulgaris* L.). *Biochim. Biophys. Acta* **399**, 356–363 (1975).
- P. A. Rea, Y. Kim, V. Sarafian, R. J. Poole, J. M. Davies, D. L. Sanders, Vacuolar  $\text{H}^+$ -translocating pyrophosphatases: A new category of ion translocase. *Trends Biochem. Sci.* **17**, 348–353 (1992).
- A. A. Baykov, A. M. Malinen, H. H. Luoto, R. Lahti, Pyrophosphate-fueled  $\text{Na}^+$  and  $\text{H}^+$  transport in prokaryotes. *Microbiol. Mol. Biol. Rev.* **77**, 267–276 (2013).
- A. Serrano, J. R. Pérez-Castañeira, M. Baltscheffsky, H. Baltscheffsky,  $\text{H}^+$ -PPases: Yesterday, today and tomorrow. *IUBMB Life* **59**, 76–83 (2007).
- T. Kajander, J. Kelloso, A. Goldman, Inorganic pyrophosphatases: One substrate, three mechanisms. *FEBS Lett.* **587**, 1863–1869 (2013).
- M. Baltscheffsky, A. Schultz, H. Baltscheffsky,  $\text{H}^+$ -proton-pumping inorganic pyrophosphatase: A tightly membrane-bound family. *FEBS Lett.* **452**, 121–127 (1999).
- R. Docampo, W. de Souza, P. Miranda, P. Rohloff, S. N. J. Moreno, Acidocalcisomes? Conserved from bacteria to man. *Nat. Rev. Microbiol.* **3**, 251–261 (2005).
- J. Kelloso, T. Kajander, K. Kogan, K. Pokharel, A. Goldman, The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* **337**, 473–476 (2012).

10. K.-M. Li, C. Wilkinson, J. Kellosalo, J.-Y. Tsai, T. Kajander, L. J. C. Jeuken, Y.-J. Sun, A. Goldman, Membrane pyrophosphatases from *Thermotoga maritima* and *Vigna radiata* suggest a conserved coupling mechanism. *Nat. Commun.* **7**, 13596 (2016).
11. S.-M. Lin, J.-Y. Tsai, C.-D. Hsiao, Y.-T. Huang, C.-L. Chiu, M.-H. Liu, J.-Y. Tung, T.-H. Liu, R.-L. Pan, Y.-J. Sun, Crystal structure of a membrane-embedded H<sup>+</sup>-translocating pyrophosphatase. *Nature* **484**, 399–403 (2012).
12. J.-Y. Tsai, J. Kellosalo, Y.-J. Sun, A. Goldman, Proton/sodium pumping pyrophosphatases: The last of the primary ion pumps. *Curr. Opin. Struct. Biol.* **27**, 38–47 (2014).
13. N. R. Shah, K. Vidilaseris, H. Xhaard, A. Goldman, Integral membrane pyrophosphatases: A novel drug target for human pathogens? *AIMS Biophys.* **3**, 171–194 (2016).
14. M. T. McIntosh, Y. M. Drozdowicz, K. Laroija, P. A. Rea, A. B. Vaidya, Two classes of plant-like vacuolar-type H<sup>+</sup>-pyrophosphatases in malaria parasites. *Mol. Biochem. Parasitol.* **114**, 183–195 (2001).
15. Y. M. Drozdowicz, M. Shaw, M. Nishi, B. Striepen, H. A. Liwinski, D. S. Roos, P. A. Rea, Isolation and characterization of TgVP1, a type I vacuolar H<sup>+</sup>-translocating pyrophosphatase from *Toxoplasma gondii*. The dynamics of its subcellular localization and the cellular effects of a diphosphonate inhibitor. *J. Biol. Chem.* **278**, 1075–1085 (2003).
16. C. O. Rodrigues, D. A. Scott, R. Docampo, Presence of a vacuolar H<sup>+</sup>-pyrophosphatase in promastigotes of *Leishmania donovani* and its localization to a different compartment from the vacuolar H<sup>+</sup>-ATPase. *Biochem. J.* **340** (Pt. 3), 759–766 (1999).
17. C. O. Rodrigues, D. A. Scott, R. Docampo, Characterization of a vacuolar pyrophosphatase in *Trypanosoma brucei* and its localization to acidocalcisomes. *Mol. Cell. Biol.* **19**, 7712–7723 (1999).
18. G. Lemercier, B. Espiau, F. A. Ruiz, M. Vieira, S. Luo, T. Baltz, R. Docampo, N. Bakalara, A pyrophosphatase regulating polyphosphate metabolism in acidocalcisomes is essential for *Trypanosoma brucei* virulence in mice. *J. Biol. Chem.* **279**, 3420–3425 (2004).
19. J. Liu, D. Pace, Z. Dou, T. P. King, D. Guidot, Z.-H. Li, V. B. Carruthers, S. N. J. Moreno, A vacuolar-H<sup>+</sup>-pyrophosphatase (TgVP1) is required for microneme secretion, host cell invasion, and extracellular survival of *Toxoplasma gondii*. *Mol. Microbiol.* **93**, 698–712 (2014).
20. Y. Zhang, A. Borrel, L. Ghemtio, L. Regad, G. Boije af Gennäs, A.-C. Camproux, J. Yli-Kauhaluoma, H. Xhaard, Structural isosteres of phosphate groups in the Protein Data Bank. *J. Chem. Inf. Model.* **57**, 499–516 (2017).
21. P. A. Karplus, K. Diederichs, Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* **34**, 60–68 (2015).
22. L. Forlani, P. D. Maria, Tautomerism of aminothiazoles.  $pK_{BH}^+$  values of 2-aminothiazoles and of some model imines. *J. Chem. Soc. Perkin Trans. 2*, 535–537 (1982).
23. E. Artukka, H. H. Luoto, A. A. Baykov, R. Lahti, A. M. Malinen, Role of the potassium/lysine cationic center in catalysis and functional asymmetry in membrane-bound pyrophosphatases. *Biochem. J.* **475**, 1141–1158 (2018).
24. G. Lemercier, S. Dutoya, S. Luo, F. A. Ruiz, C. O. Rodrigues, T. Baltz, R. Docampo, N. Bakalara, A vacuolar-type H<sup>+</sup>-pyrophosphatase governs maintenance of functional acidocalcisomes and growth of the insect and mammalian forms of *Trypanosoma brucei*. *J. Biol. Chem.* **277**, 37369–37376 (2002).
25. R. Gordon-Weeks, S. Parmar, T. G. E. Davies, R. A. Leigh, Structural aspects of the effectiveness of bisphosphonates as competitive inhibitors of the plant vacuolar proton-pumping pyrophosphatase. *Biochem. J.* **337** (Pt. 3), 373–377 (1999).
26. C. O. Rodrigues, D. A. Scott, B. N. Bailey, W. De Souza, M. Benchimol, B. Moreno, J. A. Urbina, E. Oldfield, S. N. Moreno, Vacuolar proton pyrophosphatase activity and pyrophosphate (PP<sub>i</sub>) in *Toxoplasma gondii* as possible chemotherapeutic targets. *Biochem. J.* **349** (Pt. 3), 737–745 (2000).
27. M. B. Martin, J. S. Grimley, J. C. Lewis, H. T. Heath, B. N. Bailey, H. Kendrick, V. Yardley, A. Caldera, R. Lira, J. A. Urbina, S. N. J. Moreno, R. Docampo, S. L. Croft, E. Oldfield, Bisphosphonates inhibit the growth of *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania donovani*, *Toxoplasma gondii*, and *Plasmodium falciparum*: A potential route to chemotherapy. *J. Med. Chem.* **44**, 909–916 (2001).
28. M. B. Martin, J. M. Sanders, H. Kendrick, K. de Luca-Fradley, J. C. Lewis, J. S. Grimley, E. M. Van Brussel, J. R. Olsen, G. A. Meints, A. Burzynska, P. Kafarski, S. L. Croft, E. Oldfield, Activity of bisphosphonates against *Trypanosoma brucei rhodesiense*. *J. Med. Chem.* **45**, 2904–2914 (2002).
29. A. H. Pang, A. Garzan, M. J. Larsen, T. J. McQuade, S. Garneau-Tsodikova, O. V. Tsodikov, Discovery of allosteric and selective inhibitors of inorganic pyrophosphatase from *Mycobacterium tuberculosis*. *ACS Chem. Biol.* **11**, 3084–3092 (2016).
30. U. Bischoff, C. Schmidt, R. Netzer, O. Pongs, Effects of fluoroquinolones on HERG currents. *Eur. J. Pharmacol.* **406**, 341–343 (2000).
31. W. Wang, R. MacKinnon, Cryo-EM structure of the open human ether-à-go-go related K<sup>+</sup> channel hERG. *Cell* **169**, 422–430.e10 (2017).
32. H. Ho, A. Miu, M. K. Alexander, N. K. Garcia, A. Oh, I. Zilberleyb, M. Reichelt, C. D. Austin, C. Tam, S. Shriver, H. Hu, S. S. Labadie, J. Liang, L. Wang, J. Wang, Y. Lu, H. E. Purkey, J. Quinn, Y. Franke, K. Clark, M. H. Beresini, M.-W. Tan, B. D. Sellers, T. Maurer, M. F. T. Koehler, A. T. Weckler, J. R. Kiefer, V. Verma, Y. Xu, M. Nishiyama, J. Payandeh, C. M. Koth, Structural basis for dual-mode inhibition of the ABC transporter MsbA. *Nature* **557**, 196–201 (2018).
33. M. Stornauiolo, G. E. De Kloe, P. Rucktooa, A. Fish, R. van Elk, E. S. Edink, D. Bertrand, A. B. Smit, I. J. P. de Esch, T. K. Sixma, Assembly of a  $\pi$ - $\pi$  stack of ligands in the binding site of an acetylcholine-binding protein. *Nat. Commun.* **4**, 1875 (2013).
34. V. Speranzini, D. Rotili, G. Ciossani, S. Pilotto, B. Marrocco, M. Forgione, A. Lucidi, F. Forneris, P. Mehdipour, S. Velankar, A. Mai, A. Mattevi, Polymyxins and quinazolines are LSD1/KDM1A inhibitors with unusual structural features. *Sci. Adv.* **2**, e1601017 (2016).
35. J. S. S. Damsté, W. I. C. Rijpstra, E. C. Hopmans, S. Schouten, M. Balk, A. J. M. Stams, Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order *Thermotogales*. *Arch. Microbiol.* **188**, 629–641 (2007).
36. H. H. Luoto, A. A. Baykov, R. Lahti, A. M. Malinen, Membrane-integral pyrophosphatase subfamily capable of translocating both Na<sup>+</sup> and H<sup>+</sup>. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1255–1260 (2013).
37. K. Vidilaseris, J. Kellosalo, A. Goldman, A high-throughput method for orthophosphate determination of thermostable membrane-bound pyrophosphatase activity. *Anal. Methods* **10**, 646–651 (2018).
38. S. Montalvão, T. O. Leino, P. S. Kiuru, K.-E. Lillsunde, J. Yli-Kauhaluoma, P. Tammela, Synthesis and biological evaluation of 2-aminobenzothiazole and benzimidazole analogs based on the clathrocin structure. *Arch. Pharm.* **349**, 137–149 (2016).
39. J. Kellosalo, T. Kajander, M. G. Palmgren, R. L. Lopéz-Marqués, A. Goldman, Heterologous expression and purification of membrane-bound pyrophosphatases. *Protein Expression Purif.* **79**, 25–34 (2011).
40. R. L. López-Marqués, J. R. Pérez-Castañeira, M. J. Buch-Pedersen, S. Marco, J.-L. Rigaud, M. G. Palmgren, A. Serrano, Large-scale purification of the proton pumping pyrophosphatase from *Thermotoga maritima*: A “hot-solve” method for isolation of recombinant thermophilic membrane proteins. *Biochim. Biophys. Acta* **1716**, 69–76 (2005).
41. A. A. Baykov, N. P. Bakuleva, P. A. Rea, Steady-state kinetics of substrate hydrolysis by vacuolar H<sup>+</sup>-pyrophosphatase. *Eur. J. Biochem.* **217**, 755–762 (1993).
42. I. H. Segel, *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* (John Wiley & Sons, 1975), 957 pp.
43. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
44. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
45. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
46. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
47. E. Krissinel, K. Henrick, Inference of Macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
48. Y. Y. Tseng, C. Dupree, Z. J. Chen, W.-H. Li, SplitPocket: Identification of protein functional surfaces and characterization of their spatial patterns. *Nucleic Acids Res.* **37**, W384–W389 (2009).
49. T. Y. Yuan, W.-H. Li, Identification of protein functional surfaces by the concept of a split pocket. *Proteins* **76**, 959–976 (2009).
50. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
51. H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
52. M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 (2005).
53. Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.5. Dassault Systèmes, Biovia (2015).
54. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
55. A. Šali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
56. W. Trager, J. B. Jensen, Human malaria parasites in continuous culture. *Science* **193**, 673–675 (1976).
57. H. C. Hoppe, J. A. Verschoor, A. I. Louw, *Plasmodium falciparum*: A comparison of synchronisation methods for in vitro cultures. *Exp. Parasitol.* **72**, 464–467 (1991).

58. M. Smilkstein, N. Sriwilajaroen, J. X. Kelly, P. Wilairat, M. Riscoe, Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. *Antimicrob. Agents Chemother.* **48**, 1803–1806 (2004).

**Acknowledgments:** We thank V. Manole for technical help, N. Sipari for help with the mass spectrometry analyses, and J. Kelloso, A. Malinen, and P. Henderson for fruitful discussions during the project. **Funding:** This work was supported by grants from the Jane and Aatos Erkko Foundation and the BBSRC (BB/M021610) (to A.G.), the Academy of Finland (nos. 308105 to K.V., 310297 to H.X., 265481 to J.Y.-K., and 279755 to S.M.), and University of Helsinki Research Funds (to G.B.a.G.). **Author contributions:** Conceptualization: K.V., A.G., H.X., J.Y.-K., and G.B.a.G.; methodology: K.V., A.G., A.T., A.Ki., N.G.J., and T.O.L.; investigation: K.V., A.Ki., A.T., and A.Kh.; resources: A.G., H.X., S.M., and J.Y.-K.; writing—original draft: K.V., A.T., H.X., and A.Ki.; writing—review and editing: K.V., A.Ki., A.T., A.Kh., N.G.J., T.O.L., P.S.K., G.B.a.G., S.M., J.Y.-K., H.X., and A.G.; visualization: K.V., A.T., and A.Kh.; supervision and project administration: A.G., H.X., G.B.a.G., J.Y.-K., and S.M.; funding acquisition: A.G., H.X., K.V., J.Y.-K., G.B.a.G., and

S.M. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. The atomic coordinates and structure factors of the TmPPase:IDP:ATC complex have been deposited in the PDB, [www.rcsb.org](http://www.rcsb.org) (PDB ID: 6QXA).

Submitted 17 October 2018

Accepted 16 April 2019

Published 22 May 2019

10.1126/sciadv.aav7574

**Citation:** K. Vidilaseris, A. Kiriazis, A. Turku, A. Khattab, N. G. Johansson, T. O. Leino, P. S. Kiuru, G. Boije af Gennäs, S. Meri, J. Yli-Kauhaluoma, H. Xhaard, A. Goldman, Asymmetry in catalysis by *Thermotoga maritima* membrane-bound pyrophosphatase demonstrated by a nonphosphorus allosteric inhibitor. *Sci. Adv.* **5**, eaav7574 (2019).

## Asymmetry in catalysis by *Thermotoga maritima* membrane-bound pyrophosphatase demonstrated by a nonphosphorus allosteric inhibitor

Keni Vidilaseris, Alexandros Kiriakis, Ainoleena Turku, Ayman Khattab, Niklas G. Johansson, Teppo O. Leino, Paula S. Kiuru, Gustav Boije af Gennäs, Seppo Meri, Jari Yli-Kauhaluoma, Henri Xhaard and Adrian Goldman

*Sci Adv* 5 (5), eaav7574.  
DOI: 10.1126/sciadv.aav7574

ARTICLE TOOLS	<a href="http://advances.sciencemag.org/content/5/5/eaav7574">http://advances.sciencemag.org/content/5/5/eaav7574</a>
SUPPLEMENTARY MATERIALS	<a href="http://advances.sciencemag.org/content/suppl/2019/05/20/5.5.eaav7574.DC1">http://advances.sciencemag.org/content/suppl/2019/05/20/5.5.eaav7574.DC1</a>
REFERENCES	This article cites 55 articles, 16 of which you can access for free <a href="http://advances.sciencemag.org/content/5/5/eaav7574#BIBL">http://advances.sciencemag.org/content/5/5/eaav7574#BIBL</a>
PERMISSIONS	<a href="http://www.sciencemag.org/help/reprints-and-permissions">http://www.sciencemag.org/help/reprints-and-permissions</a>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.