

Supplementary Materials for

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

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

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

Published 22 May 2019, *Sci. Adv.* **5**, eaav7574 (2019)
DOI: 10.1126/sciadv.aav7574

This PDF file includes:

- 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)

Supplementary Materials

Supplementary Methods

Size-exclusion chromatography – Multi angle laser light scattering (SEC-MALLS)

50 µg of TmPPase purified in DDM with IDP (1 mM) and IDP+ATC (1 mM each) were run through a Superdex 200 10/300 GL equilibrated with a buffer containing 50 mM MES-NaOH pH 6.5, 50 mM KCl, 5 mM MgCl₂, 2 mM DTT and 0.05% DDM using a HPLC-system (Shimadzu). The eluted samples from the column were analysed using UV-absorbance, refractive index (Optilab rEX, Wyatt) and multi angle laser light scattering (miniDAWN TREOS, Wyatt) detectors and the enzyme molecular mass was calculated using ASTRA software (Wyatt).

Anti-plasmodial activity assay

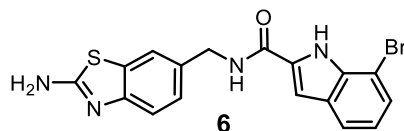
For the *in vitro* anti-plasmodial activity assay, *Plasmodium falciparum* strain 3D7 was maintained in culture as described previously (56). Briefly, parasites were cultured in O+ human erythrocytes at 5% haematocrit in RPMI-1640 medium supplemented with 0.5% Albumax II (Gibco, Carlsbad, CA, USA), 200 µM hypoxanthine (Sigma, St. Louis, MO, USA) and 20 µg ml⁻¹ gentamycin (Gibco). Parasites were synchronized with the sorbitol method as described (57). The test compound was dissolved in DMSO to make a stock solution of 10 mM. Two-fold serial dilutions of the test compound were made in the culture medium to cover the range of 100 µM-100 nM. Fifty microliters of each dilution were mixed in a 96-well plate with 150 µl of 2% haematocrit of 0.5% parasitemia (ring stage). After 72 h incubation at 37 °C, the parasite growth was quantified using fluorescent SYBR Green I[®]-based assay as described (58). The half-maximal inhibitory concentration (IC₅₀) of the test compound was assessed by non-linear regression in Prism 7 (GraphPad Software). The potent anti-malarial drug, artemisinin, was used in parallel as a positive control.

Experimental procedures and characterization data of compounds

All reactions were carried out with commercially available starting materials (Sigma-Aldrich, Schnelldorf, Germany; Biofine International, Vancouver, Canada) and solvents without further purification. Anhydrous DMF was obtained commercially and all solvents for work-up and chromatography were obtained as HPLC quality and used as received. Column chromatography was performed with Merck 230–400 mesh silica gel or with an automated Biotage high performance flash chromatography Isolera One (Uppsala, Sweden) using a 0.1 mm path length flow cell UV-detector/recorder module (fixed wavelength: 254 nm). Analytical thin layer chromatography (TLC) was carried out using 0.2-mm silica gel plates (silica gel 60, F₂₅₄, Merck KGaA, Darmstadt, Germany). Nuclear magnetic resonance spectra (¹H NMR at 400 MHz and ¹³C NMR at 100 MHz) were recorded on a Bruker Ascend 400 (Bruker Corporation, Billerica, Massachusetts, USA). The chemical shifts are reported in parts per million (ppm) on the δ scale using tetramethyl silane (TMS) as an internal standard. The coupling constants *J* are quoted in Hertz (Hz). Data for ¹H NMR spectra are reported as follows: chemical shift (multiplicity, integration, coupling constant(s)). The multiplicity was abbreviated as follows: br, broad signal; s, singlet; d, doublet; t, triplet; q, quartet; quin, quintet; dd, doublet of doublets; dt, doublet of triplets; dq, doublet of quartets; m, multiplet; m_c, centered multiplet. High resolution mass spectra (HRMS) were measured on a Waters Synapt G2 (Waters Corporation, Milford, Massachusetts, USA) and reported for the molecular ions [M+H]⁺. LC-MS purity analyses were performed on a Waters Acquity[®] UPLC system (Waters, Milford MA, USA) attached to an Acquity PDA detector and Waters Synapt G2 HDMS mass spectrometer with an ESI ion source.

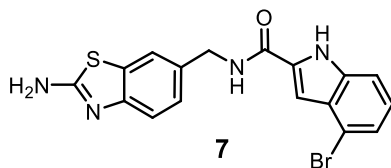
Compounds **1–5** were synthesized according to known literature procedures (38).

***N*-[(2-Aminobenzo[*d*]thiazol-6-yl)methyl]-7-bromo-1*H*-indole-2-carboxamide (6)**



6-(Aminomethyl)benzo[*d*]thiazol-2-amine (38 mg, 0.21 mmol), 7-bromoindole-2-carboxylic acid (0.050 g, 0.21 mmol), EDC·HCl (41 mg, 0.21 mmol) and 4-DMAP (25 mg, 0.21 mmol) were dissolved in dry DMF (6 ml) under argon. The resulting light-yellow suspension was stirred at r.t. for 23 h after which TLC from the reaction mixture showed formation of product. The solution was concentrated with the rotary evaporator and the residue was purified by automated column chromatography on silica gel (eluent: EtOAc/*n*-hexane 1:4→3:4) to give **6** as an off-white solid (42 mg, 50%). ¹H NMR (400 MHz, DMSO-*d*₆): δ 11.37 (1H, br s), 9.06 (1H, br t, *J* = 6.0 Hz), 7.66 (1H, d, *J* = 6.0 Hz), 7.64 (1H, d, *J* = 1.6 Hz), 7.44 (1H, dd, *J* = 0,8 Hz, 7.6 Hz), 7.43 (2H, br s), 7.31 1H, d, *J* = 8.4 Hz), 7.26 (1H, d, *J* = 2.0 Hz), 7.21 (1H, dd, *J* = 1.6 Hz, 8.4 Hz), 7.01 (1H, t, *J* = 7.6 Hz), 4.53 (2H, d, *J* = 6.0 Hz); ¹³C NMR (100 MHz, DMSO-*d*₆): δ 166.4, 160.0, 152.0, 135.1, 133.2, 131.8, 131.1, 128.4, 126.3, 125.3, 121.3, 121.1, 120.0, 117.5, 106.0, 104.5, 42.4; LC-MS: [M+H]⁺, *m/z* 401 (*t*_r = 2.76 min), ≥95%; HRMS-ESI *m/z*: calc. for C₁₇H₁₃BrN₄OS [M+H]⁺: 401.0072, found 401.0075.

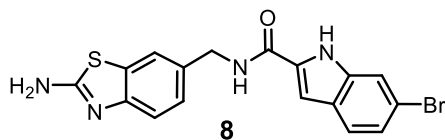
***N*-[(2-Aminobenzo[*d*]thiazol-6-yl)methyl]-4-bromo-1*H*-indole-2-carboxamide (7)**



6-(Aminomethyl)benzo[*d*]thiazol-2-amine (38 mg, 0.21 mmol), 4-bromoindole-2-carboxylic acid (0.050 g, 0.21 mmol), EDC·HCl (41 mg, 0.21 mmol) and 4-DMAP (25 mg, 0.21 mmol) were dissolved in dry DMF (6 ml) under argon. The resulting light yellow suspension was stirred at r.t. for 23 h after which TLC from the reaction mixture showed formation of product. The solution

was concentrated with the rotary evaporator and the residue was purified by automated column chromatography on silica gel (eluent: EtOAc/*n*-hexane 1:4→1:0) to give **7** as an off-white solid (41 mg, 49%). ¹H NMR (400 MHz, DMSO-*d*₆): δ 11.99 (1H, br s), 9.16 (1H, br t, *J* = 6.0 Hz), 7.61 (1H, d, *J* = 1.6 Hz), 7.44 (1H, d, *J* = 8.4 Hz), 7.41 (2H, br s), 7.29 (1H, d, *J* = 8.0 Hz), 7.27 (1H, dd, *J* = 0.8 Hz, 7.6 Hz), 7.23 (1H, m_c) 7.20 (1 H, dd, *J* = 1.6 Hz, 8.0 Hz), 7.11 (1H, dd, *J* = 0.8 Hz, 7.6 Hz), 4.51 (2H, d, *J* = 6.0 Hz); ¹³C NMR (100 MHz, DMSO-*d*₆): δ 166.3, 160.4, 151.9, 136.7, 132.4, 132.0, 131.0, 127.7, 125.2, 124.4, 122.4, 119.8, 117.4, 114.4, 112.0, 102.3, 42.3; LC-MS: [M+H]⁺, *m/z* 401 (*t*_r = 2.84 min), ≥95%; HRMS-ESI *m/z*: calc. for C₁₇H₁₃BrN₄OS [M+H]⁺: 401.0072, found 401.0068.

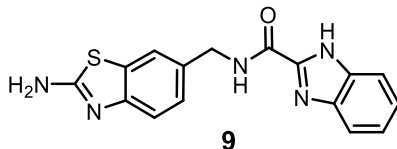
***N*-[(2-Aminobenzo[*d*]thiazol-6-yl)methyl]-6-bromo-1*H*-indole-2-carboxamide (**8**)**



6-(Aminomethyl)benzo[*d*]thiazol-2-amine (38 mg, 0.21 mmol), 6-bromoindole-2-carboxylic acid (0.050 g, 0.21 mmol), EDC·HCl (41 mg, 0.21 mmol) and 4-DMAP (25 mg, 0.21 mmol) were dissolved in dry DMF (6 ml) under argon. The resulting light-yellow suspension was stirred at r.t. for 40 h after which TLC from the reaction mixture showed formation of product. The solution was concentrated with the rotary evaporator and the residue was purified by automated column chromatography on silica gel (eluent: EtOAc/*n*-hexane 1:4→1:0) to give **8** as an off-white solid (21 mg, 25%). ¹H NMR (400 MHz, DMSO-*d*₆): δ 11.74 (1 H, br s), 9.06 (1H, br t, *J* = 6.0 Hz), 7.61–7.58 (3H, m), 7.41 (2H, br s), 7.29 (1H, d, *J* = 8.4 Hz), 7.20 (1H, m_c) 7.18 (1H, d, *J* = 8.4 Hz, overlapping), 7.16 (1H, dd, *J* = 1.6 Hz, 8.4 Hz), 4.51 (2H, d, *J* = 6.0 Hz); ¹³C NMR (100 MHz, DMSO-*d*₆): δ 166.3, 160.6, 151.9, 137.1, 132.5, 132.0, 131.0, 126.1, 125.0, 123.4, 122.7,

119.7, 117.4, 116.0, 114.7, 102.7, 42.2; LC-MS: $[M+H]^+$, m/z 401 ($t_r = 2.95$ min), $\geq 95\%$; HRMS-ESI m/z : calc. for $C_{17}H_{13}BrN_4OS$ $[M+H]^+$: 401.0072, found 401.0075.

***N*-[(2-Aminobenzo[*d*]thiazol-6-yl)methyl]-1*H*-benz[*d*]imidazole-2-carboxamide (**9**)**



6-(Aminomethyl)benzo[*d*]thiazol-2-amine (23 mg, 0.13 mmol), benzimidazole-2-carboxylic acid (21 mg, 0.13 mmol), EDC·HCl (25 mg, 0.13 mmol) and 4-DMAP (16 mg, 0.13 mmol) were dissolved in dry DMF (4 ml) under argon. The resulting light-yellow suspension was stirred at r.t. for 24 h after which TLC from the reaction mixture showed formation of product. The solution was concentrated with the rotary evaporator and the residue was purified by automated column chromatography on silica gel (eluent: EtOAc/*n*-heptane 1:4→1:0) to give **9** as a white solid (12 mg, 29%). 1H NMR (400 MHz, DMF- d_7): δ 13.26 (1H, br s), 9.30 (1H, br t, $J = 6.4$ Hz), 7.74 (1H, s), 7.70 (2H, br s), 7.50 (2H, br s), 7.35–7.29 (4H, m) 4.67 (2H, d, $J = 6.4$ Hz); ^{13}C NMR (100 MHz, DMF- d_7): δ 166.9, 159.1, 152.6, 146.2, 143.3, 135.1, 132.5, 131.8, 125.5, 124.2, 122.7, 120.3, 120.0, 117.8, 112.6, 42.6; LC-MS: $[M+H]^+$, m/z 324 ($t_r = 1.74$ min), $\geq 95\%$; HRMS-ESI m/z : calc. for $C_{16}H_{14}N_5OS$ $[M+H]^+$: 324.0919, found 324.0919.

Supplementary Figures

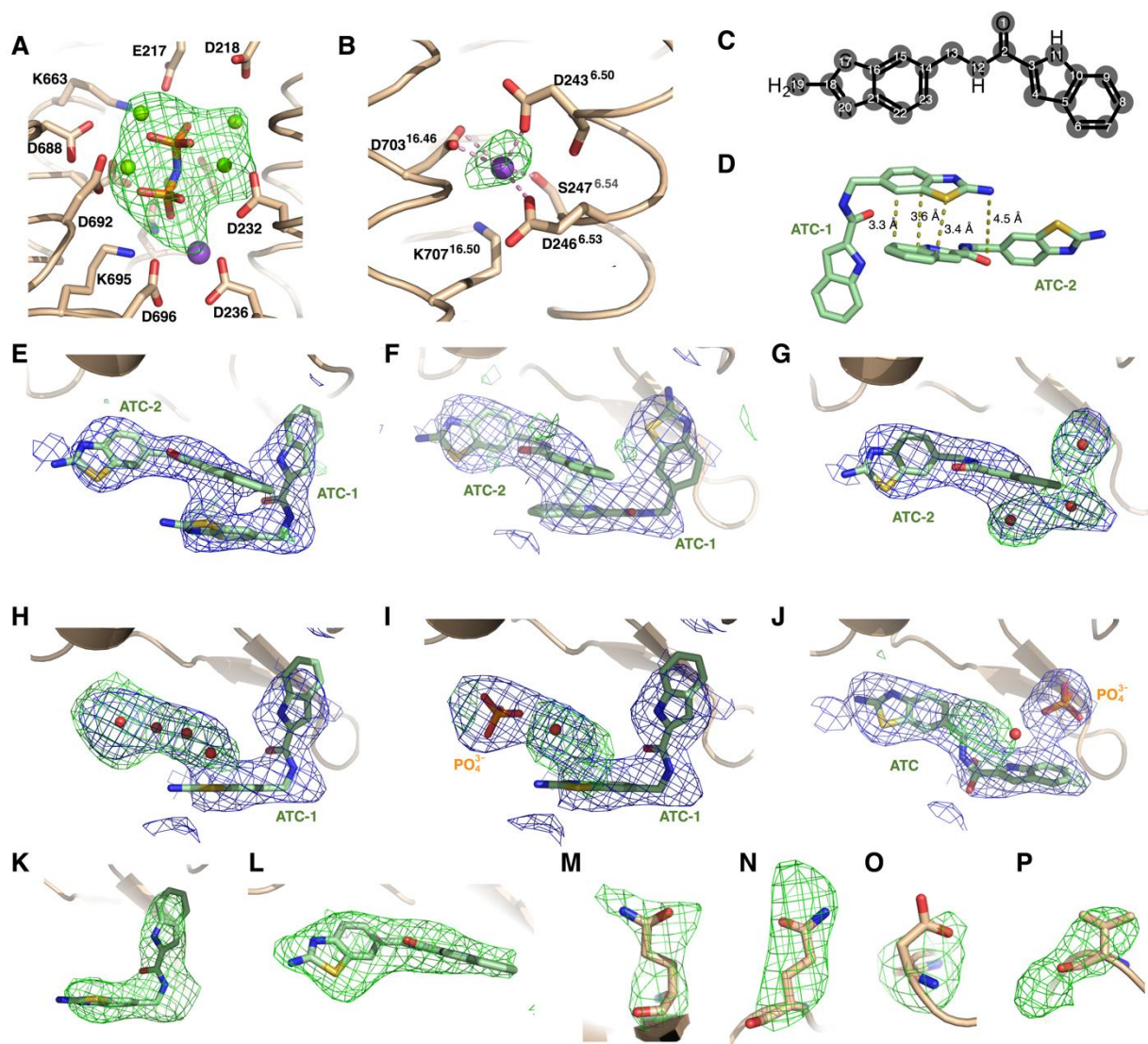


Fig. S1. Electron density maps of ligands in the TmPPase:IDP:ATC structure. (A) Omit ($F_o - F_c$) map contoured at 3σ (green mesh) in the hydrolytic center, into which five magnesium ions (green spheres), a potassium ion (violet-purple sphere) and IDP can be fit. (B) Omit ($F_o - F_c$) map contoured at 3σ (green mesh) electron density map at the gate, which can be fitted with a sodium ion (violet-purple sphere). The sodium ion is coordinated by oxygens from three Asp residues and one serine. (C) Numbering scheme of the atoms in ATC. (D) Distances between dimer interface of ATC-1 and ATC-2. (E) ($2F_o - F_c$) (blue mesh) and ($F_o - F_c$) (green mesh) electron density map of ATC-1 and ATC-2 at 1σ and 3σ , respectively. (F–J) Electron density maps as in D, but with different molecules built into the density and refined: (F): ATC-1 orientation reversed; (G) ATC-

1 replaced with three water molecules (red dots); (**H**) ATC-**2** replaced with four water molecules (red dots); (**I**) ATC-**2** replaced with a water molecule (red dot) and phosphate ion (orange); (**J**) ATC in a different position with a water molecule (red dot) and a phosphate ion (orange). In all the maps (**F–J**), there is considerable positive ($F_o - F_c$) electron density; the only model that fits as well as our preferred model has a reverse ATC-1. (**K–P**) Polder OMIT maps of ATC-**1**, ATC-**2**, Q268, Q277, D351 and V352, demonstrating that the ATC-**1** & ATC-**2** fit the electron density similarly to the protein side chains.

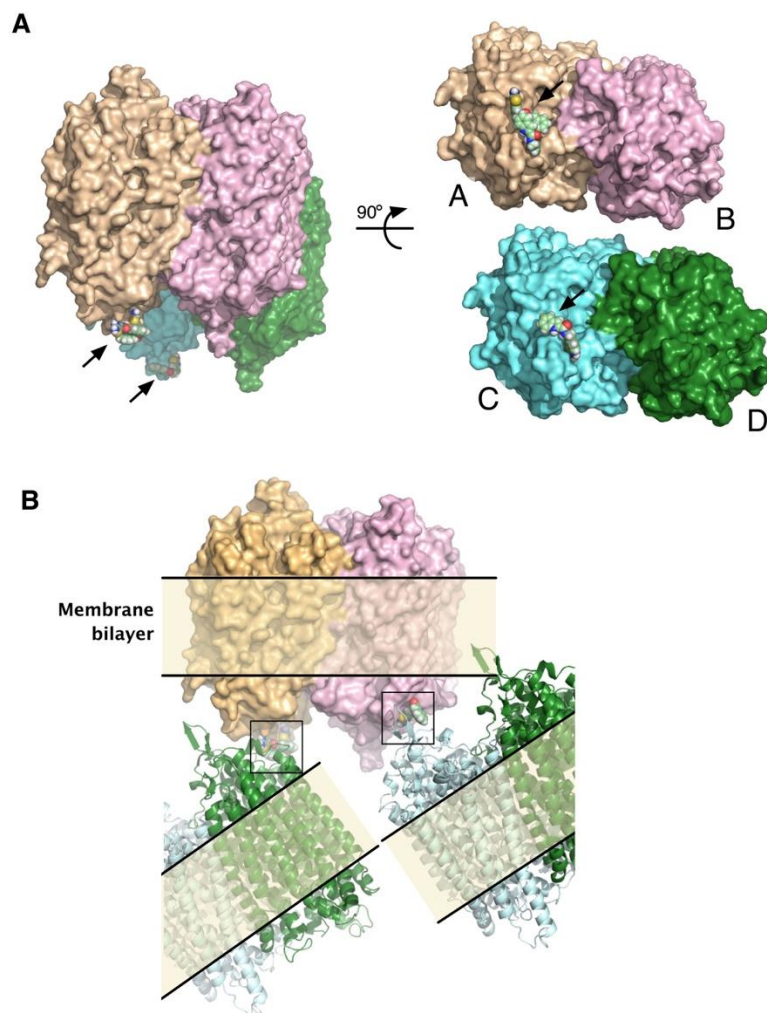


Fig. S2. Asymmetric unit of TmPPase:IDP:ATC. (A) Four TmPPase chains in the asymmetric unit showing the dimer arrangement and ATC binding site (arrow). Color scheme: chain A, wheat; chain B, light pink; chain C, cyan; chain D, forest green. (B) Interaction between chain AB with CD via ATC in the crystal contact. The small boxes indicate the ATC binding sites.

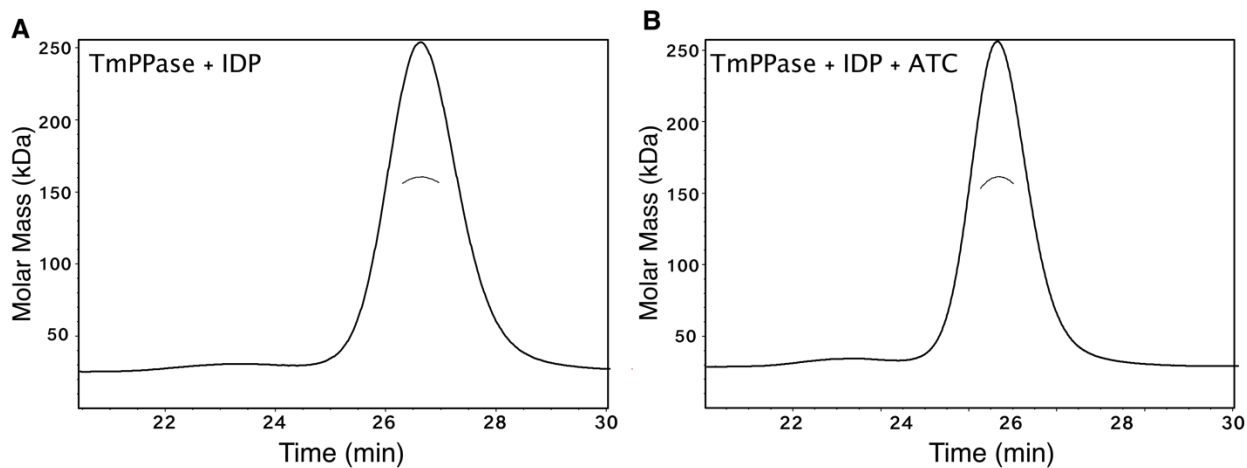


Fig. S3. Oligomerization state of TmPPase upon the addition of inhibitors. Size-exclusion chromatography – multiangle laser light scattering (SEC-MALLS) analysis of (A) TmPPase+IDP and (B) TmPPase+IDP+ATC. Adding ATC to the protein-IDP complex did not change the oligomerization state. Line across the light scattering chromatogram peak shows the molecular weight distribution of the protein.

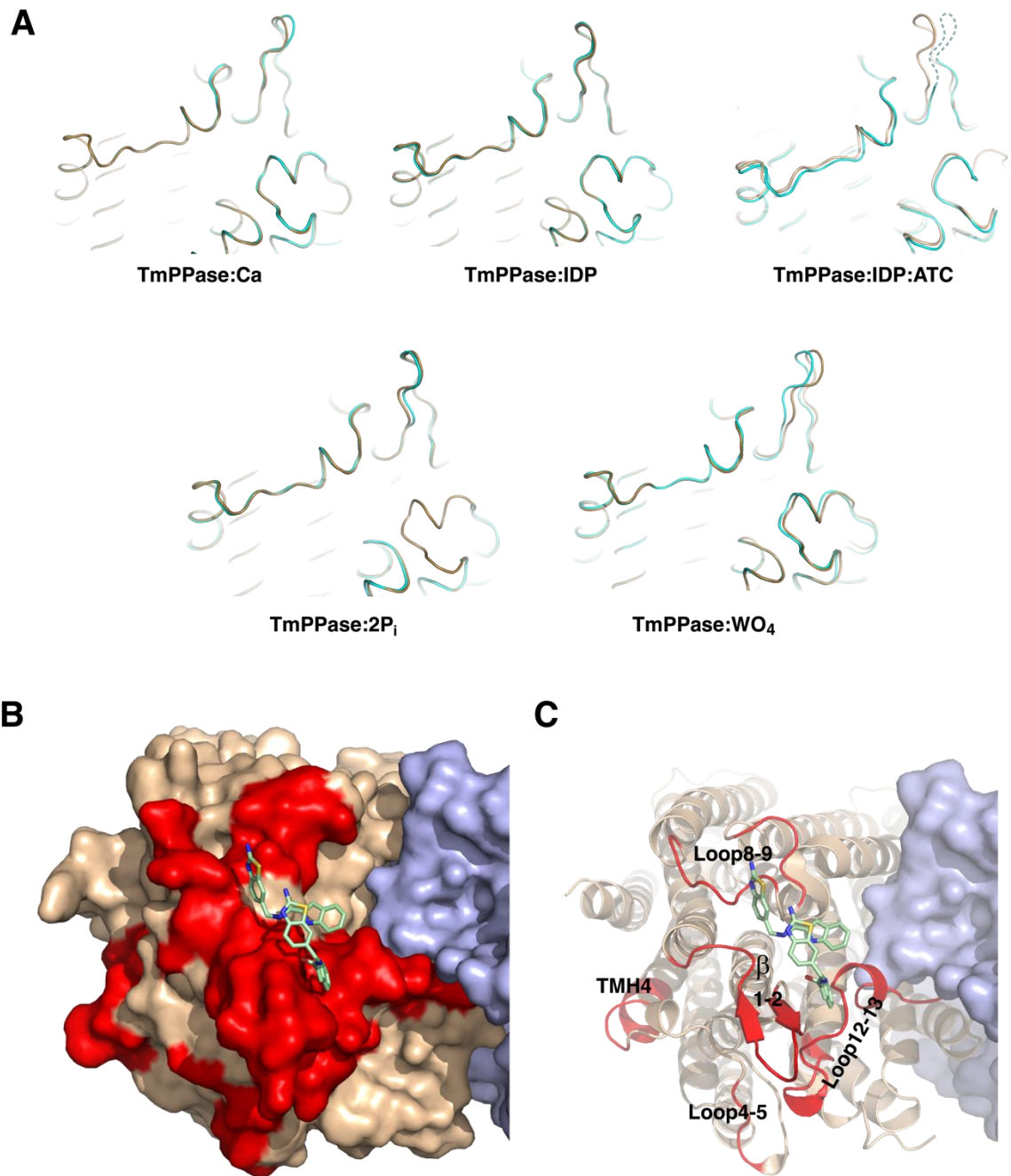


Fig. S4. Comparison of the ATC binding sites in different TmPPase structures. (A) Superposition of ATC binding site between chain A (wheat) and B (cyan) of different TmPPase states. **(B)** Surface and **(C)** cartoon representation of region near exit channel showing the mobile region (red) around the ATC binding site.

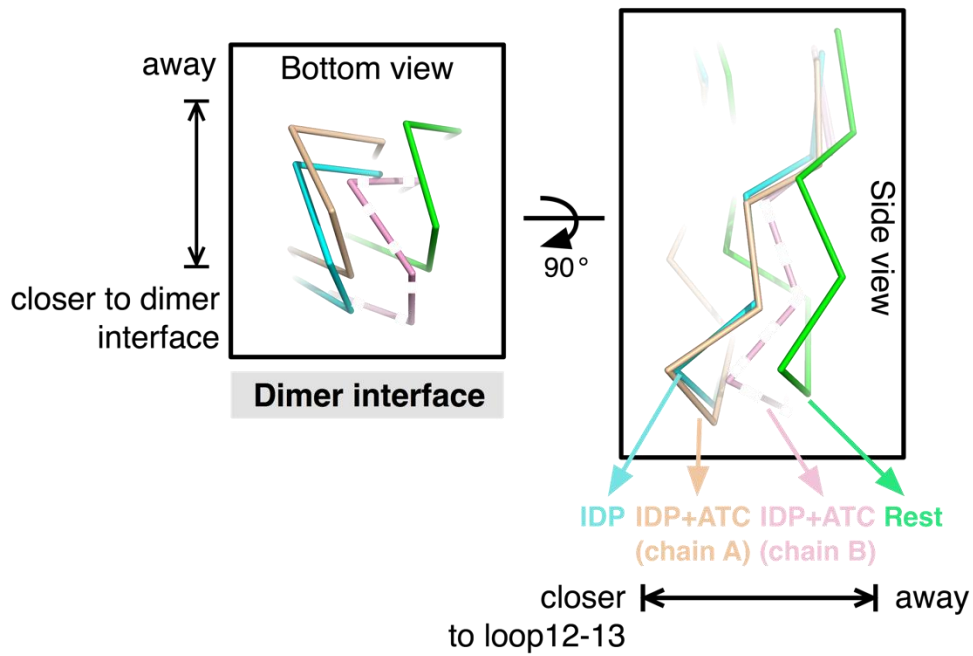


Fig. S5. Differences in the ATC binding site between monomers of the TmPPase:IDP:ATC structure. Superposition of chain A and B of the IDP:ATC-bound state in comparison with chain A of the resting state and IDP-bound state mPPase showing the closer/away movement of β 1–2 strand relative to the loop12–13 and the dimer interface of the protein. β 1–2 strand of chain B is represented with dashed ribbon as this region is flexible and not visible in the structure. Coloring scheme: resting state (TmPPase:Ca:Mg) = green; chain B of IDP:ATC-bound state = light pink; chain A of IDP:ATC-bound state = wheat; IDP-bound state (TmPPase:IDP)= cyan, ATC = pale green.

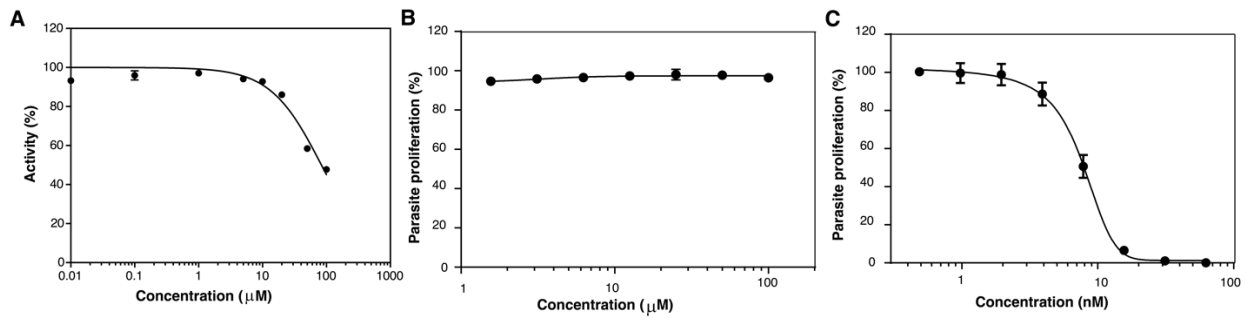


Fig. S6. Effect of ATC on growth of *P. falciparum* and on PfPPase. (A) Inhibition effect of ATC on PfPPase-VP1 (in membrane) from *P. falciparum* expressed in insect cell baculovirus expression system (method and further analysis will be described in more detail elsewhere). ATC inhibits the protein at $\log IC_{50}$ of 1.9 ± 0.03 ($IC_{50} = 83.6 \mu M$). (B) ATC has no anti-plasmodial activity up to $100 \mu M$. (C) Anti-plasmodial activity of artemisinin, as a positive control.

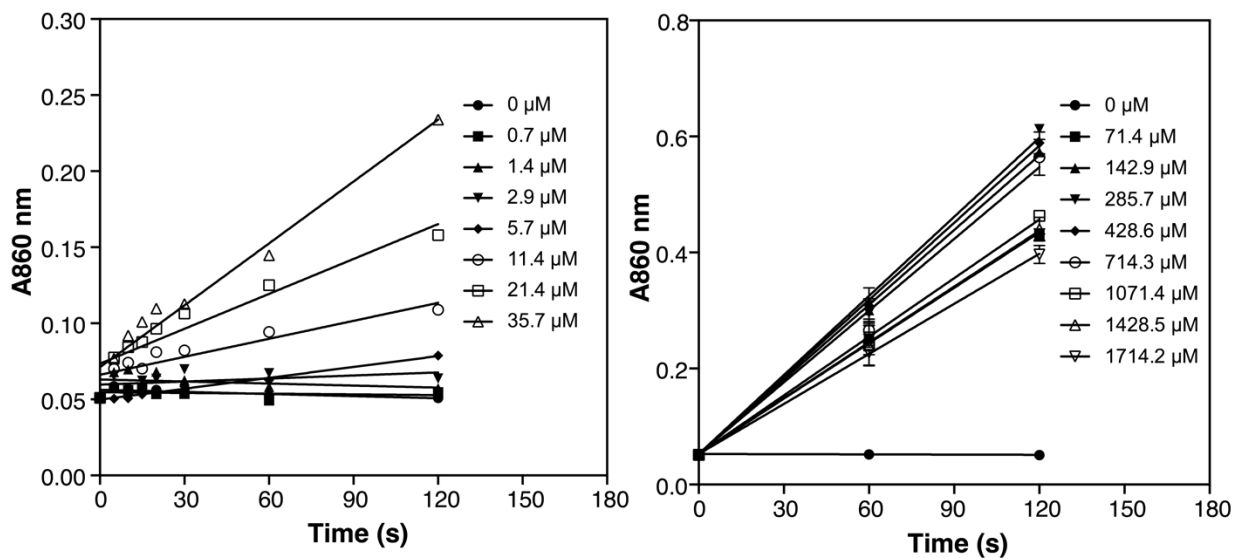


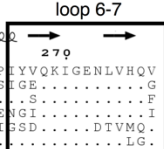
Fig. S7. Linearity of TmPPase activity as a function of time and concentration. Multiple time points were measured for all of the PP_i concentrations used (left panel: 0–35.7 μM ; right panel: 71.4–1714.2 μM). At two minutes reaction time, the absorbance at 860 nm is still in the linear range, so two minutes was used in the kinetic analysis (Fig. 4).

TMH1
1 10 20 30
TmPase (1) MYV.A.....ALFFLIP.LVALGFAA.NF.A.....AVV.....R
PifPase-VP1 (2)NNMDFY.....VFLLLPP.TIGLIFSI..ECIWSNINIKGPE.....KVEKLE..D
VrPase (2) MGAAI L P D L G T . E L I P V C A . V G I A F A L F . Q W L L V S K V K L S A V R D A S P N A A A K N G Y N D Y L I E E E B E G I N
PifPase-VP2 (3)NDNYVHQMTNDHNIQKKEIKKYEKLYKDNHNNHNNNNLINFYMLQNNQTFNQHSNEYEKNNINLIHTK.....NTNNDNRKGDHMKNEIGTN
BvPase (4)METYL.F.....WIVFAS.LLALVLAWY.FY..K.....QMM.....L
CIPase (5)MPL.S.....TLAEVGA.VVALVLAFF.LS..R.....RVM.....R

TMH2
30 40 50 60 70
TmPase (1) K P I G T E R K K P T S S Y I R S C A D S L A H E T K A F K V A I V I A T L M L I F T T W G T C F
PifPase-VP1 (2) G D N P E V E R K M K T I A S Y I A V C A N A E L K K E F P Q Y L A V F I I V F S I L G F F V N S F T A V
VrPase (2) D H N V V V K C A I Q N A I S E C A T S I F L F T E Y K Y V G I F M V A F A I L F L P L G S V E G F S T S P Q A C S Y D K T K T C K P A L A T A I F S T V
PifPase-VP2 (3) E G Y E I S S F D S T G I P I K E C S E G F F T V Q Y N S I F K I S I V F T L L S L Y I . I R G E D T K L P Q G D N Q M T D M N N N N N N N N N N N S S S S N G T I I S S Y A G I I T A I
BvPase (4) E S E G T P T M E K I A S Y V R Q C A M S Y L K Q Y K V V G L V F L G L V L F S I M A Y . G F N L Q N P W V P I
CIPase (5) A D E G T D L M K K I A G A V R K C N A Y L K R Y T G V T A I F F A V M F V L L I L A . I A G F L T F P V V I

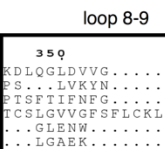
TMH3
80 90 100 110 120 130 140 150 160 170
TmPase (1) A F I I G A V M S A S A I H V C M K M A T R A N V I V A E A R T T K K I G P A L K V A Y Q G S V M G L S V G G F A L L G L V L V L I F G K W M G Q V D N L N I Y T N W L G I N F V P F A M T V S G
PifPase-VP1 (2) S F V I G C L T S I L C G V I C M K I A V A N V R T T N E T W . . K S L D K G F Q V T L N A G T V V G F S L V S F G I I A L G L L I F V Y K T Y V F K N T I P D N Q I Y K I I A G
VrPase (2) S F I I G G V T S I V S G C V I C M K I A T A N A R T T L E A R . . K G V G K A F I T A F R G A V S F L L A A N G L L V L Y I A I N L F K I Y Y G D D W G G L F E A I T G
PifPase-VP2 (3) Y G F G A L C S S I A G V N G H Y V A V R A N V K V A R A A T . Y S Y K K A L I T R S G A V S A I V N V A L A I F G I C S L L L V N L Y P T L A F S Y P P L L I V G
BvPase (4) A F I I G G F F S G L S C H L I C M K I A T A S A R T A N A Q . . H S L N K G L R V A F R S G A V G L V V V G L L D I S F W Y I L L D Y C I P T D A . L N P S A K L C V I T T M T L T
CIPase (5) A F I I G G F F S G L S C H L I C M K I A T A N G S T A N S . . H S L N A G R I A I S A G S V M G F V V V G L L D I S I W Y P L F K F W Y S T V D V M Q N E A A Q V Q A I T S A M L T

TMH5
180 190 200 210 220 230 240 250 260 270
TmPase (1) Y A L G C S I I A M F D R V G G V V T K A A D M A A D I V G H T E L N L P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V
PifPase-VP1 (2) F I G C S S I A L F S R V G C G I T K A A D V A D I V G H T E L N L P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V
VrPase (2) Y G F G A S L A M L Y Q L A G I T K A A D I G A D L V G K V E A G P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V
PifPase-VP2 (3) F G M G A S T Q A L F A R V G C G I T K A A D V G A D L V G K V E A G P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V
BvPase (4) F G M G A S M A L F A R V G C G I T K A A D V G A D L V G K V E A G P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V
CIPase (5) F G M G A S M A L F A R V G C G I T K A A D V G A D L V G K V E A G P E D D D R N P A I A D N V G D N V G D A G L G A D I L E S F V G A I V S S I L A S Y M F P I Y V Q I G E N L V H Q V



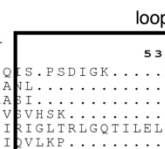
TMH7
280 290 300 310 320 330
TmPase (1) P K E T I Q A L S Y P F F A L V G L G C M L G I L Y I V K . K P S D N . P Q R E N I S L W T S A L T
PifPase-VP1 (2) S P G N A F H C I L F P L F V S F V I C S M I T F Y I T Y S V I N D K K D V E K S K Y L L L S T L V L
VrPase (2) G L N H E L T A P L P I V S S V G L I V C L L T T L F A T D F F I K A V K E I E P A L K K Q L V S T L V L
PifPase-VP2 (3) S E Q T A S Y V I P F V H S M D L L S I T I G I Y L V S K N N N E P F S T L N L E C N T K I I D K K D I N Y F N N S N M H V E E S A M Q P L N N N N K E E S D L L Y A R Y D C D V I
BvPase (4) F K A V I A P L I A A V G I V S I I G I F A V R T K . E N A G M K E L L G S V M S F A V V C A T G T N S E L V L
CIPase (5) F R G V I P S M A A I G I L A S I I G T F F V R T K . E N A S Q K N L L G S R L G T W T S A L T

TMH8
340 350 360 370 380 390 400 410 420
TmPase (1) T V V L T A F E L T Y F K E D L Q G L D V V G F . R F G A . I S P W F S A I T G I F S I L I G F W A E Y T Y S Y R K P T Q F L S K S S I E G T G M V I N G S L G M K S V F P P
PifPase-VP1 (2) Q S L A I L A I G Y V C P S L V K Y N Y L K D I H R . W K I I V P A L V G L W S Q L I G F T T E Y T Y S Y R K P T Q F L S K S S I E G T G M V I N G S L G M K S V F P P
VrPase (2) M T I G A V V V S E V A I P T S F T I F N G V Q K D V K S . W O L F L C V A V G L W A L I G I G F T T E Y T Y S Y R K P T Q F L S K S S I E G T G M V I N G S L G M K S V F P P
PifPase-VP2 (3) E N P I K A L M K A Y F F T C S L G V G S F L C K L F S I D N A K N A I I F S P C G L L V I L F R V Y S I P K V K K I A H A S L S A N I I A G L Y V E S F P P
BvPase (4) I V V A T F L L W A L G L E N N V M S F A V V C A T G T N S E L V L
CIPase (5) I A V A A F P L I Y F G L G A E K I G F Y F A I L S G L A C I L I G V V T E Y T Y S Y R K P T Q F L S K S S I E G T G M V I N G S L G M K S V F P P



TMH11
430 440 450 460 470 480 490 500
TmPase (1) L I L V L G I L F A D Y F A G L V C V A I A I C M L S V A T S V S V G P A D N A G C S E M C E D P E V R K R I T D H L A V E N T T A A K G F A G S A
PifPase-VP1 (2) I I C L S A T L G S Y G L C D I Y G A L A A Y G M L S T I C I L T D A G P S D N A G C A E M A G L P E S V E R R E R T D L D A A G N T T A A K G F A G S A
VrPase (2) I F A I A I S I F V S F T F A M Y G A V A A I C M L S T A T G L A D A G P S D N A G C A E M A G S H R I R E R T D L D A A G N T T A A K G F A G S A
PifPase-VP2 (3) I V A I S I L L S Y L G L K S N I T G D H N I I N G L Y C T S V A T C M L S T A V F L S M S N G P A D N A G C L V E M C K P K Y V R V I T D K L S L G N T T A A K G F A G S A
BvPase (4) V L A V V T G I I L S Y W L A S G F D F A N I S M G L Y G T I A A V G M L S T L G I T L A T D A G P A D N A G C A E M S G L G K E V R K R T D L D S L G N T T A A K G F A G S A
CIPase (5) V I V G A S V L V S Y L S G S D S . Y N N G L Y C V L S A V G M L A T G I T L A T D A G P A D N A G C A E M A H G E E V R K R T D L D S L G N T T A A K G F A G S A

TMH13
510 520 530 540 550 560 570 580
TmPase (1) F A A L S L F A S M F S Q S P S D I G K P P S L V L L N M L D A R V I A G A L L C A A I T V F G G Y L T S A V T K A A M K M V D E I R R O A R E I P G L
PifPase-VP1 (2) L V A F A L F G A P A S S A N L R F N I L N S M V I G L L I G A M L P V L F S A L T M K S V A I A A N S V L N B C L E F P P L I
VrPase (2) I V S L A L F G A P V S R A S I T T V D V L T P V F T G L I V G A M L P W F S A M T M K S V G S A A L K M V E B V R R O F N I P G L
PifPase-VP2 (3) L A C F L F S A P L S E V V H S K M P F S T V D I A T P V F I G G I L G S V V V L F A G W L D A V G K T A E V L K E B V R R O F N E H P G I
BvPase (4) I T G L A L L A S V E E I T G L R L G Q T I L E L P N G T I V D V H N A S T D M L Y I T L M N P R V L G M F L G S M M A L F G C L T M N A V G R A A H M V E B V R R O F R I I K G I
CIPase (5) I T A L A L F A S I T A E I V L K P D E N F M L T I T N P V L N G L F I G G V L P L F A A L T M A V G R C A Q K I V V E B V R R O F R I I K G I

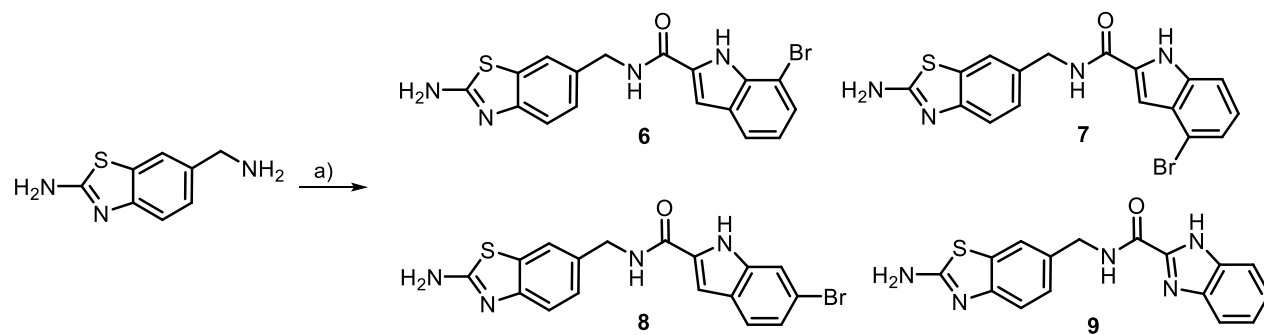


TMH14
590 600 610 620 630 640 650 660 670
TmPase (1) E E K A K P D Y N R K E I T S D N A L K C M G V P A F I A I L T I G C L L S A E F V G C V I G T V L C A M L A L T A N S G A N D N A K K I L S G N L E G Y .
PifPase-VP1 (2) L E C K Q K P D Y E K E I T S D A S L R Q M I V P G L I S V F S P E L M M L K Y A T A G L I G I L L C I Q L A S T S S G A N D N A K K I L S G A L G K E H
VrPase (2) M E C T A K P D Y A T C W K I T S D A S I K E M I P G A L V M L P L V P G I L F C V E T L S C V L A S S L V C V Q I A R S A S N T G A N D N A K K I L S G S E H A R
PifPase-VP2 (3) L I T Y E K P D Y H T C V A I T S K R A L I E T I R P G L L G L S P T I V G L F K Q L G L L Q N N Q L L G A Q A M A S F I M F S T S C I L M A L F L N N A G A N D N A K K I L S G Y G . . .
BvPase (4) L T G E A E P D Y A R C V D I S T K G A Q R E M V P S L L A I A I A I A T G L F I G V P G V I G L I G L S C E F V L A F M A N A G A N D N A K K I L S G N F G . . .
CIPase (5) M T G E A D P D Y A S C V D I C T R S A Q R E M I P A I V A A D I I V G L I L S V N G V A G M L A S A T V C F I L A V M M A N S G A N D N A K K I L S G E Y G . . .

TMH16
680 690 700 710 720
TmPase (1) G K S S P H R A I V I G D T V G D P K D T V G P S L I L I K I M S V S V S V A V S I F K H V H L F .
PifPase-VP1 (2) G K S N A R N V I G D T V G D P K D T S G P S L N I L I K I S A I T S L V F A N V I A T K F T S T R G G P I W L . .
VrPase (2) S L G K S D C H R A A V I G D T V G D P K D T S G P S L N I L I K I M A V E S L V F A P F F A T H G L L F K I F .
PifPase-VP2 (3) G K S G A H V S V I G D T V G D P K D T A G P S I H V L I K I I S I T M V I L
BvPase (4) G K S S V H R A T V T G D T V G D P K D T S G P S L N I L I K I M S M V A I M A G L T V A W S L F .
CIPase (5) G K S S N H R A A V I G D T V G D P K D T S G P S I N I L I K I L S M V A I F A S V I S L F S L I K

Fig. S8. Multiple sequence alignment of mPPases from different organisms. We only show PfPPase-VP1 and PfPPase-VP2 as representative of the 16 parasite mPPases, all of which have 40–70% sequence identity to each other, and none of which have loop 6-7. Numbering after mPPase corresponds to subfamilies: (1): Na⁺-pumping K⁺-dependent (TmPPase from *T. maritima*); (2): H⁺-pumping K⁺-dependent (PfPPase-VP1 from *P. falciparum* and VrPPase from *V. radiata*); (3): H⁺-pumping K⁺-independent (PfPPase-VP2 from *P. falciparum*); (4): Na⁺,H⁺-pumping mPPase (BvPPase from *Bacteroides vulgatus*); and (5): Na⁺-regulated Na⁺,H⁺-pumping mPPase (CIPPase from *Clostridium leptum*). Secondary structure is based on the structure of TmPPase. Black rounded rectangles show loops important for ATC binding. Numbering above the sequence corresponds to the residue number in TmPPase. Red box and white characters: identical residues; blue frame and red characters: similar residues. The figure was generated with ESPript 3.0 (<http://espript.ibcp.fr/ESPript/ESPript/index.php>).

Supplementary Schemes



Scheme S1. Synthesis of analogs of 1 with heavy and basic atoms. a) Reagents and conditions: Br-indole-2-carboxylic acid/ benzimidazole-2-carboxylic acid (1.0 eq.), EDC HCl (1.0 eq.), 4-DMAP (1.0 eq.), DMF, RT, 23-40 h (38).

Supplementary Tables

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

Data Collection	
Space group	P 2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	98.1, 141.6, 252.1
α , β , γ (°)	90, 90, 90
Source	ESRF - ID29
Wavelength (Å)	0.9677
Resolution (Å)	47.49 – 3.42 (3.73 – 3.42)
Anisotropy direction*	
Resolution where CC _{1/2} >0.50	
Overall (Å)	3.42
along h axis	3.57
along k axis	4.11
along l axis	3.41
Measured reflections	186550 (7994)
Unique reflections	36342 (1846)
Completeness (%)	93.7 (57.4)
CC _{1/2}	0.99 (0.56)
Mean <i>I</i> / σ (<i>I</i>)	9.2 (1.5)
Multiplicity	5.1 (4.3)
Wilson B (Å ²)	98.7
R _{merge}	0.11 (0.98)
R _{meas}	0.13 (1.11)
R _{pim}	0.06 (0.52)

Refinement	
Resolution (Å)	47.5 – 3.4
R _{work} (%)/R _{free} (%)	22.8/28.4
No. of atoms	21042
protein	20895
ligands	139
water	8
No. of chains (ASU)	4
B-factors (Å ²)	86.1
Protein	86.0
Ligands/Ion	100.5
R.M.S. Deviations	
Bond lengths (Å)	0.002
Bond angle (°)	0.44
Ramachandran statistics [†]	
Favored (%)	95.71
Allowed (%)	4.29
Outliers (%)	0.00

Statistics for the highest-resolution shell are shown in parentheses.

*The anisotropy statistics were computed with AIMLESS after use of the StarAniso server.

†Ramachandran statistics were calculated with MolProbity.

$CC_{1/2}$, correlation coefficient; h, k, l: indices that define the lattice planes, $I/\sigma(I)$, empirical signal-to-noise ratio; R_{merge} , quality agreement among equivalent reflections to assess X-ray data quality; R_{meas} , multiplicity-corrected R; R_{pim} , expected precision.

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

Monomer	RMSD (Å²)
B	0.31
C	0.26
D	0.33

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

Interaction type	Compound	Chain A		Chain D	
		Residue	BSA* (\AA^2)	Residue	BSA* (\AA^2)
Hydrophobic/ Aromatic	ATC-1	Q268	42.1	I70	18.8
		K269	23.4	T73	19.4
		I270	22.7	W74	20.0
		P530	24.2	W164	26.4
	ATC-2	Q268	32.3	W74	41.0
		Q277	23.8	Y161	32.4
		D351	24.7	F169	18.3
		V352	18.8		
	Total BSA* (\AA^2)		212.0		176.3
Possible hydrogen bond	Compound	Chain A	Distance (\AA)	Chain D	Distance (\AA)
	ATC-1 (O01)	Q268 (N ϵ 2)	2.7		
	ATC-2 (N11)	Q268 (O ϵ 1)	3.0		
	ATC-2 (N12)	Q268 (O ϵ 1)	3.0		
	ATC-2 (N19)			N157 (N δ 2)	3.1
Salt bridge	ATC-2 (N20)	D351 (O δ 1)	3.4		

* BSA: Buried surface area: the solvent-accessible surface area (cut-off = 15 \AA^2) of the corresponding residue that is buried upon interface formation.

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

Loop (residues)	RMSD (Å) ¹				
	TmPPase:IDP:ATC	5LZQ	5LZR	4AV3	4AV6
2-3 (71-75)	0.63	0.18	0.44	0.33	0.26
4-5 (153-172)	1.11	0.24	0.58	0.37	0.32
β1-2 (6-7) (266-276)	1.30 ²	0.49	1.24	0.61	0.50
8-9 (342-359)	0.79 ³	0.41	0.70	0.34	0.35 ⁴
12-13 (520-540)	0.80	0.37	0.41	0.46	0.36

¹ Comparisons were done by superimposing helix 7 (aa. 281-311), which moves least between the structures.

² Missing residues 269-274. RMSD was measured between the remaining residues.

³ Missing residues 346-349. RMSD was measured between the remaining residues.

⁴ Missing residues 349-357. RMSD was measured between the remaining residues.

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

Loop (residues)	RMSD (Å)¹
	Monomer A
2-3 (71-75)	0.81
4-5 (153-172)	0.75
β1-2 (6-7) (266-276)	1.13
8-9 (342-359)	0.70
12-13 (520-540)	1.09

¹ Comparisons were made by superimposing helix 7 (residues 281-311), which moves least between the TmPPase:IDP and TmPPase:IDP:ATC structures

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

PP_i conc. (μM)	Hill constant
35	2.8 ± 0.28
71	2.8 ± 0.22
142	1.9 ± 0.17
285	2.2 ± 0.14
1071	2.2 ± 0.18
1428	1.9 ± 0.21
1721	2.0 ± 0.21

^1H NMR and ^{13}C NMR Spectra

