Supporting Information

Fragment-Based Discovery of Selective Inhibitors of the *Mycobacterium tuberculosis* Protein Tyrosine Phosphatase PtpA

Katherine A. Rawls,[a] P. Therese Lang,[b] Jun Takeuchi,[a] Shinichi Imamura,[a] Tyler D. Baguley,[a] Christoph Grundner,[b,*] Tom Alber,[b,*] and Jonathan A. Ellman[a,*]

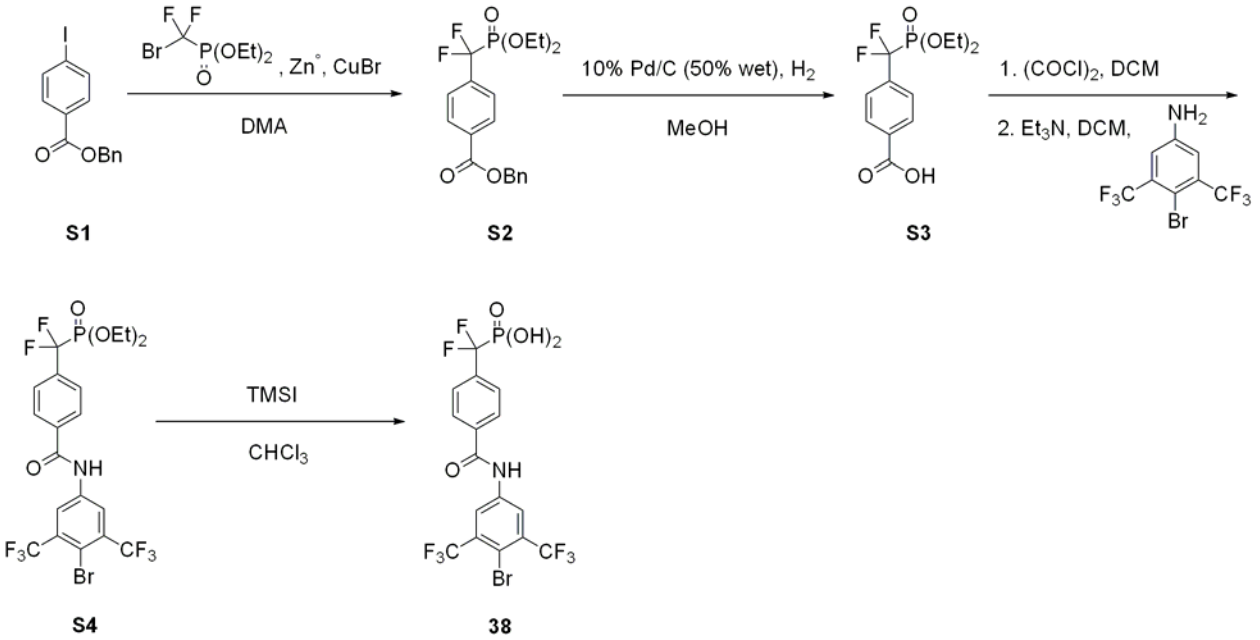
[a]*Department of Chemistry, University of California, Berkeley, California, 94720-1460*
[b]*Department of Molecular and Cell Biology, University of California, Berkeley, California, 94720-3200*

Table of Contents

1. General synthetic methods.

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. Tetrahydrofuran (THF), dichloromethane (CH_2Cl_2), and diethyl ether (Et_2O) were dried over alumina under a nitrogen atmosphere. Methanol was dried over calcium hydride under a nitrogen atmosphere. Dimethylacetamide (DMA) was purchased as anhydrous grade in a sure-seal bottle. All reactions, unless otherwise stated, were performed under inert atmosphere using syringe and cannula techniques, with flame-dried glassware. All ^{1}H, ^{13}C, ^{19}F, and ^{31}P NMR spectra were measured with a Bruker DRX-500, AVB-400, AVQ-400 or AV-300 spectrometer. NMR chemical shifts are reported in ppm relative to 1,2-difluorobenzene (-138.9) for ^{19}F NMR and trimethylphosphate (3.0) for ^{31}P NMR. Mass spectrometry (HRMS) was carried out by the University of California, Berkeley Mass Spectrometry Facility.

2. Synthesis and analytical data for benzanilide inhibitors.



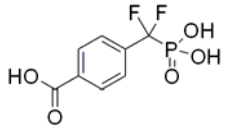
Synthesis of S2. A solution of diethyl-(bromodifluoromethyl)phosphonate (11.8 g, 44.0 mmol) in DMA (20 mL) was slowly added to a stirred suspension of activated Zn dust (2.88 g. 44 mmol) in DMA (20 mL) at 60 °C. After addition was complete, the mixture was sonicated at room temperature for 3 h, followed by addition of CuBr (6.31 g, 44 mmol) in one portion. A solution of benzyl 4-iodobenzoate (**S1**, 5.41 g, 16 mmol) in DMA (5 mL) was added dropwise, and the resulting mixture was stirred for 38 h at room temperature. The mixture was diluted with water (50 mL) and ether (50 mL), and was then passed through Celite. The organic layer was separated, washed with brine (1 x 100 mL), dried over anhydrous $MgSO_4$ (s), and filtered. The solvent was removed under reduced pressure to afford crude product, which was then purified via column chromatography to yield **S2** as a colorless oil (4.68 g, 73% yield). **Analytical data.** ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.16 (d, J = 8.0 Hz, 2H), 7.70 (d, J = 8.0 Hz, 2H), 7.46-7.36 (m, 5H), 5.39 (s, 2H), 4.26-4.11 (m, 4H), 1.30 (t, J = 7.2 Hz, 6H); ^{19}F NMR (376 MHz, $CDCl_3$): δ -108.59 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, $CDCl_3$): δ 5.76 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{19}H_{22}F_2O_5P$, 399.1172; found, 399.1177.

Synthesis of S3. A solution of **S2** (2.50 g, 6.28 mmol) in MeOH (5mL) was added to 10% Pd/C (835 mg, ca. 50% wet) in MeOH (30mL). The reaction mixture was stirred for 16 h under H_2 atmosphere. The catalyst was then removed by filtration through Celite, and the solvent was removed under reduced pressure to give crude **S3**. The crude product was purified by recrystallization from EtOAc/hexanes to yield **S3** as a white powder (1.67 g, 86% yield). **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.15 (d, J = 8.0 Hz, 2H), 7.70 (d, J = 8.0 Hz, 2H), 4.26-4.13 (m, 4H), 1.31 (t, J = 7.2 Hz, 6H); ^{19}F NMR (376 MHz, CD_3OD): δ -111.51 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 5.73 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + Na]^+$ calcd for $C_{12}H_{15}F_2O_5PNa$, 331.0523; found, 331.0531.

Synthesis of S4. To a solution of **S3** (98 mg, 0.32 mmol) in dry CH_2Cl_2 (3 mL) was added oxalyl chloride (55 μL, 0.64 mmol) and a catalytic amount of DMF. The reaction mixture was stirred for 1 h at room temperature, followed

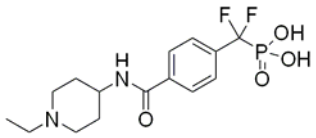
by removal of the solvent under reduced pressure, and drying under high vacuum. The resulting acid chloride was dissolved in dry CH_2Cl_2 (2 mL) and slowly added to a solution of 4-bromo-3,5-bis(trifluoromethyl)aniline (114 mg, 0.37 mmol) and triethylamine (67 µL, 0.48 mmol) in dry CH_2Cl_2 (3 mL). The reaction mixture was stirred at room temperature for 18 h. The solvent was evaporated under reduced pressure to afford crude **S4**. The crude product was purified via column chromatography to give **S4** as a white solid (96 mg, 50% yield). **Analytical data**. ^{1}H NMR (400 MHz, CD_3OD): δ 9.07 (br s, 1H), 8.40 (s, 2H), 7.92 (d, J = 8.2 Hz, 2H), 7.65 (d, J = 7.8 Hz, 2H), 4.29-4.16 (m, 4H), 1.33 (t, J = 7.1 Hz, 6H); ^{19}F NMR (376 MHz, CD_3OD): δ -61.47 (s), -108.27 (d, J_{PF} = 114 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.86 (t, J_{PF} = 114 Hz). MS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{20}H_{18}BrF_8NO_4P$, 596.9951; found, 598.0 and 599.0. MS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{20}H_{18}BrF_8NO_4P$, 597.9951; found, 598.0 and 600.0.

Synthesis of 38. To a stirred solution of **S4** (35 mg, 0.06 mmol) in $CHCl_3$ (3 mL) was added TMSI (33 µL, 0.22 mmol). The mixture was stirred for 3 h at room temperature. Volatiles were removed under reduced pressure and the residue was dissolved in MeOH (3 mL) and stirred at room temperature for 18 h. The solvent was removed under reduced pressure to give crude **S4**. The crude product was purified by recrystallization from EtOAc/hexanes to yield **36** as a white powder (13 mg, 41% yield). **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.54 (s, 2H), 8.06 (d, J = 8.0 Hz, 2H), 7.77 (d, J = 8.0 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -64.40, -112.46 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.55 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{16}H_{10}{}^{79}BrF_8NO_4P$, 541.9403; found, 541.9406.



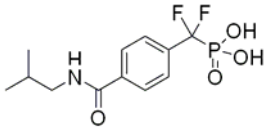
1

Synthesis of 1. Compound **1** was synthesized by subjecting **S3** to the procedure described for the synthesis of **38** from **S4**. **Analytical data.** ^{1}H NMR (CD_3OD): δ 7.93 (d, 2H, J = 7.4 Hz), 7.72 (d, 2H, J = 7.4 Hz); ^{19}F NMR (CD_3OD): δ -110.78 (d, J_{FP} = 111 Hz); ^{31}PNMR (CD_3OD): δ 3.38 (br m). HRMS-ESI (*m/z*): $[M - H]^-$ calcd for $C_8H_6O_5F_2P$, 250.9926; found, 250.9916.



11

Synthesis of 11. Compound **11** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.85 (d, 2H, J = 8.3 Hz), 7.70 (d, 2H, J = 8.3 Hz), 4.13 (m, 1H), 4.03 (quintet, 2H, J = 7.1 Hz), 3.43 (br m, 2H), 3.12 (br m, 2H), 2.17 (br m, 2H), 1.79 (br m, 2H), 1.24 (t, 3H, J = 7.1 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -108.82 (d, J_{FP} = 99 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 3.20 (br t, J_{PF} = 99 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{22}O_4N_2F_2P$, 363.1280; found, 363.1273.

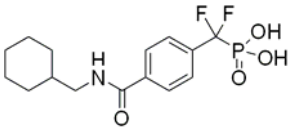


12

Synthesis of 12. Compound **12** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.89 (d, 2H, J = 8.3 Hz), 7.69 (d, 2H, J = 8.3 Hz), 3.20 (d, 2H, J = 6.9 Hz), 1.93 (nonet, 1H, J = 6.9 Hz), 0.96 (d, 6H, J = 6.9 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -110.60 (d, J_{FP} = 110 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.37 (br t, J_{PF} = 112 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{12}H_{17}O_4NF_2P$, 308.0858; found, 308.0852.

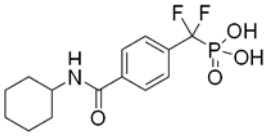
13

Synthesis of 13. Compound **13** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.89 (d, 2H, J = 8.1 Hz), 7.68 (d, 2H, J = 8.1 Hz), 3.09 (s, 3H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.67 (d, J_{FP} = 110 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.34 (br t, J_{PF} = 110 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_9H_{11}O_4NF_2P$, 266.0388; found, 266.0385.



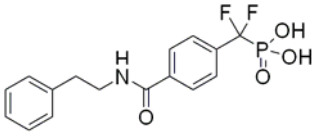
14

Synthesis of 14. Compound **14** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.89 (d, 2H, J = 7.1 Hz), 7.69 (d, 2H, J = 7.1 Hz), 3.22 (d, 2H, J = 6.9 Hz), 1.77 (br m, 4H), 1.71-1.62 (br m, 2H), 1.34-1.18 (br m, 3H), 1.02 (br m, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.65 (d, J_{FP} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.63 (br m). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{21}O_4NF_2P$, 348.1171; found, 348.1166.



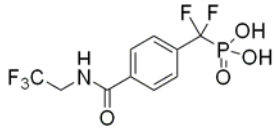
15

Synthesis of 15. Compound **15** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.87 (d, 2H, J = 8.0 Hz), 7.67 (d, 2H, J = 8.0 Hz), 3.88-3.84 (m, 1H), 1.95 (br m, 2H), 1.81 (br m, 2H), 1.68 (br m, 1H), 1.45-1.29 (br m, 4H), 1.27-1.19 (br m, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.63 (d, J_{FP} = 110 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.72 (br m). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{19}O_4NF_2P$, 334.1014; found, 334.1020.



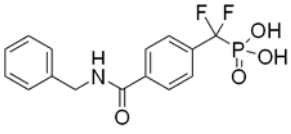
16

Synthesis of 16. Compound **16** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.94 (d, 2H, J = 8.2 Hz), 7.84 (d, 2H, J = 8.2 Hz), 7.34-7.17 (m, 5H), 3.60 (t, 2H, J = 7.4 Hz), 2.91 (t, 2H, 7.4 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -110.73 (d, J_{FP} = 112 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.49 (br m). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{16}H_{17}O_4NF_2P$, 356.0858; found, 356.0865.



17

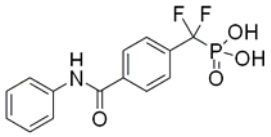
Synthesis of 17. Compound **17** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.93 (d, 2H, J = 7.6 Hz), 7.72 (d, 2H, J = 8.0 Hz), 4.10 (q, 2H, J_{HF} = 9.2 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -72.92 (t, J_{FH} = 9.4 Hz), -110.78 (d, J_{FP} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.48 (br m). HRMS-ESI (*m/z*): $[M - H]^-$ calcd for $C_{10}H_8O_4NF_5P$, 332.0117; found, 332.0105.



18

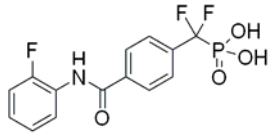
Synthesis of 18. Compound 1**8** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 7.93 (d, 2H, J = 8.1 Hz), 7.70 (d, 2H, J = 8.1 Hz), 7.36-7.22 (m,

5H), 4.58 (s, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.57 (d, J_{FP} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.09 (br t, J_{PF} = 109 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{15}O_4NF_2P$, 342.0701; found, 342.0712.



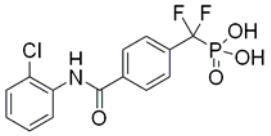
20

Synthesis of 20. Compound **20** was synthesized in a manner analogous to the procedures for **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.02 (d, *J* = 8.0 Hz, 2H), 7.74 (d, *J* = 8.0 Hz, 2H), 7.69 (d, *J* = 8.0 Hz, 2H), 7.36 (t, *J* = 7.6 Hz, 2H), 7.15 (t, *J* = 7.6 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.39 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.65 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + Na]^+$ calcd for $C_{14}H_{12}F_2NO_4PNa$, 350.0370; found, 350.0366.



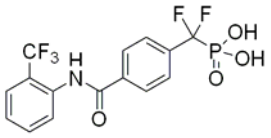
22

Synthesis of 22. Compound **22** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.03 (d, *J* = 7.1 Hz, 2H), 7.80-7.71 (m, 3H), 7.29-7.14 (m, 3H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.70 (d, J_{PF} = 111 Hz), -124.20; ^{31}P NMR (162 MHz, CD_3OD): δ 4.27 (t, J_{PF} = 102 Hz). MS-ESI (*m/z*): $[2M + H]^+$ calcd for $C_{28}H_{23}F_6N2O_8P_2$, 691.0760; found, 691.0.



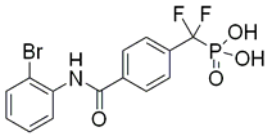
23

Synthesis of 23. Compound **23** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.02 (d, *J* = 8.0 Hz, 2H), 7.89 (t, *J* = 2.0 Hz, 1H), 7.75 (d, *J* = 8.0 Hz, 2H), 7.60 (dd, *J* = 8.0, 2.0 Hz, 1H), 7.34 (d, *J* = 8.0 Hz, 1H), 7.15 (dd, *J* = 8.0, 2.0 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.44 (d, J_{PF} = 109 Hz). ^{31}P NMR (162 MHz, CD_3OD): δ 2.64 (t, J_{PF} = 109 Hz). HRMS-EI (m/z): $[M + H]^+$ calcd for $C_{14}H_{12}{}^{35}ClF_2NO_4P$, 362.0160; found, 362.0169.



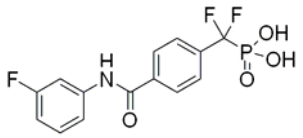
24

Synthesis of 24. Compound **24** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.03 (d, *J* = 8.0 Hz, 2H), 7.78-7.71 (m, 4H), 7.62-7.60 (m, 1H), 7.54-7.52 (m, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -61.52, -110.88 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.60 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{12}F_5NO_4P$, 396.0424; found, 396.0424.



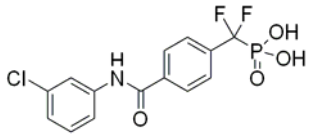
25

Synthesis of 25. Compound **25** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.07 (d, *J* = 8.0 Hz, 2H), 7.77 (d, *J* = 8.0 Hz, 2H), 7.72-7.69 (m, 2H), 7.43 (t, *J* = 7.6 Hz, 1H), 7.21 (d, *J* = 7.6 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.45 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.64 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{12}{}^{79}BrF_2NO_4P$, 405.9655; found, 405.9666.



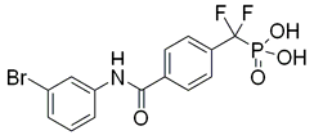
26

Synthesis of 26. Compound **26** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.02 (d, J = 8.0 Hz, 2H), 7.75 (d, J = 8.0 Hz, 2H), 7.69-7.66 (m, 1H), 7.45 (d, J = 8.0 Hz, 1H), 7.38-7.32 (m, 1H), 6.88 (t, J = 8.0 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.41 (d, J_{PF} = 109 Hz), -115.00; ^{31}P NMR (162 MHz, CD_3OD): δ 2.80 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{10}F_3NO_4P$, 344.0300; found, 344.0292.



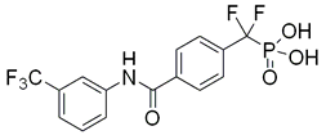
27

Synthesis of 27. Compound **27** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.02 (d, J = 8.0 Hz, 2H), 7.89 (t, J = 2.0 Hz, 1H), 7.75 (d, J = 8.0 Hz, 2H), 7.60 (dd, J = 8.0, 2.0 Hz, 1H), 7.34 (d, J = 8.0 Hz, 1H), 7.15 (dd, J = 8.0, 2.0 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.44 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.64 (t, J_{PF} = 109 Hz). HRMS-EI (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{12}{}^{35}ClF_2NO_4P$, 362.0160; found, 362.0169.



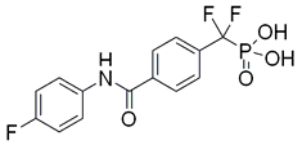
28

Synthesis of 28. Compound **28** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.03-8.00 (m, 4H), 7.75 (d, J = 8.0 Hz, 1H), 7.67-7.65 (m, 1H), 7.29-7.26 (m, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.27 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.64 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{12}{}^{79}BrF_2NO_4P$, 405.9655; found, 405.9650.



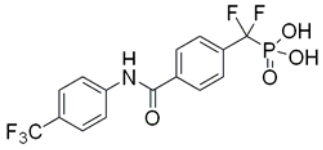
29

Synthesis of 29. Compound **29** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.17 (s, 1H), 8.04 (d, J = 8.0 Hz, 2H), 7.95 (d, J = 8.0 Hz, 1H), 7.76 (d, J = 8.0 Hz, 2H), 7.56 (t, J = 8.0 Hz, 1H), 7.44 (d, J = 8.0 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -65.08, -112.39 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.64 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + Na]^+$ calcd for $C_{15}H_{11}F_5NO_4PNa$, 418.0244; found, 418.0243.



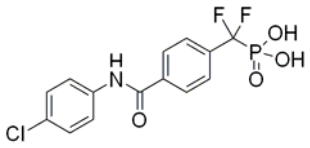
30

Synthesis of 30. Compound 30 was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.01 (d, J = 8.0 Hz, 2H), 7.74 (d, J = 8.0 Hz, 2H), 7.70 (dd, J = 8.8, 4.8 Hz, 2H), 7.10 (t, J = 8.8 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.38 (d, J_{PF} = 109 Hz), -120.72; ^{31}P NMR (162 MHz, CD_3OD): δ 2.71 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + Na]^+$ calcd for $C_{14}H_{10}F_3NO_4PNa$, 390.0095; found, 390.0102.



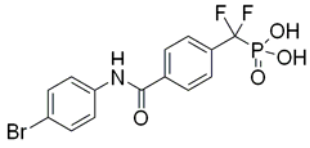
31

Synthesis of 31. Compound **31** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.03 (d, J = 8.0 Hz, 2H), 7.94 (d, J = 8.0 Hz, 2H), 7.76 (d, J = 8.0 Hz, 2H), 7.66 (d, J = 8.0 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -64.41, -112.37 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.59 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{12}F_5NO_4P$, 396.0424; found, 396.0423.



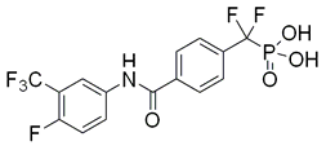
32

Synthesis of 32. Compound **32** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.01 (d, J = 8.0 Hz, 2H), 7.74 (d, J = 8.0 Hz, 2H), 7.72 (d, J = 8.0 Hz, 2H), 7.36 (d, J = 8.0 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.42 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.60 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + 2Na - H]^+$ calcd for $C_{14}H_{10}{}^{35}ClF_2NO_4PNa_2$, 405.9799; found, 405.9815.



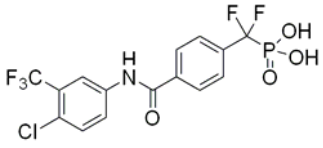
33

Synthesis of 33. Compound **33** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.01 (d, J = 8.0 Hz, 2H), 7.74 (d, J = 8.0 Hz, 2H), 7.67 (d, J = 8.8 Hz, 2H), 7.50 (d, J = 8.8 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -112.39 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.74 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{12}{}^{79}BrF_2NO_4P$, 405.9655; found, 405.9650.



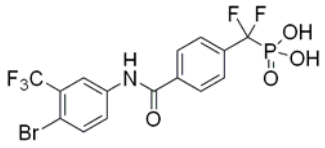
34

Synthesis of 34. Compound **34** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.14 (m, 1H), 8.05-7.92 (m, 3H), 7.82 (s, 1H), 7.75 (m, 1H), 7.31 (m, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -62.06 (d, J_{FP} = 13 Hz), -121.21. MS-ESI (*m/z*): $[M - H]^-$ calcd for $C_{15}H_9F_6NO_4P$, 413.0252; found 413.0.



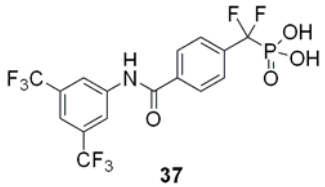
35

Synthesis of 35. Compound **35** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.27 (s, 1H), 8.04 (d, J = 8.0 Hz, 2H), 7.99 (d, J = 8.8 Hz, 1H), 7.75 (d, J = 8.0 Hz, 2H), 7.59 (d, J = 8.8 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -64.96, -112.45 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.62 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{11}{}^{35}ClF_5NO_4P$, 430.0034; found, 430.0029.



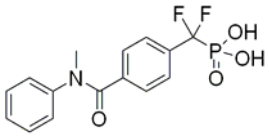
36

Synthesis of 36. Compound **36** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.27 (s, 1H), 8.04 (d, J = 8.0 Hz, 2H), 7.92 (d, J = 7.2 Hz, 1H), 7.79-7.74 (m, 3H); ^{19}F NMR (376 MHz, CD_3OD): δ -64.96, -112.46 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.62 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{11}{}^{79}BrF_5NO_4P$, 473.9529; found, 473.9542.



37

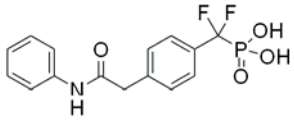
Synthesis of 37. Compound **37** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): δ 8.42 (s, 2H), 8.07 (d, J = 8.0 Hz, 2H), 7.77 (d, J = 8.0 Hz, 2H), 7.71 (s, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -65.38, -112.46 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.54 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{16}H_{11}F_8NO_4P$, 464.0297; found, 464.0301.



39

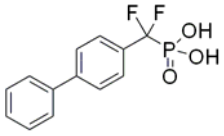
Synthesis of 39. Compound **39** was synthesized by following the general procedures for the synthesis of **38**. **Analytical data.** ^{1}H NMR (400 MHz, CD_3OD): 7.43 (d, J = 8.0 Hz, 2H), 7.39 (d, J = 8.0 Hz, 1H), 7.28 (t, J = 7.6 Hz, 2H), 7.22- 7.15 (m, 3H), 3.47 (s, 3H); ^{19}F NMR (376 MHz, CD_3OD): -112.01 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): 2.62 (t, J_{PF} = 109 Hz). HRMS-FAB (*m/z*): $[M + 2Na - H]^+$ calcd for $C_{15}H_{13}F_2NO_4PNa_2$, 386.0346; found, 386.0354.

3. Analytical data for other inhibitors.



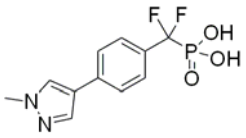
2

Compound 2. ^{1}H NMR (400 MHz, CD_3OD): δ 3.74 (2H, s), 7.09 (1H, t, J = 7.4 Hz), 7.30 (2H, t, J = 8.0 Hz), 7.47 (2H, d, J = 8.0 Hz), 7.52-7.56 (2H, m), 7.59 (2H, d, J = 8.0 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -110.03 (d, J = 112.8 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.75 (t, J = 112.6 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{15}F_2NO_4P$, 342.0701; found, 342.0715.



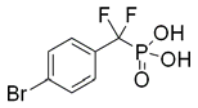
3

Compound 3. ^{1}H NMR (400 MHz, CD_3OD): δ 7.75-7.63 (m, 6H), 7.45 (t, J = 7.6 Hz, 2H), 7.34 (t, J = 7.4 Hz, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -111.8 (d, J_{PF} = 113 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 3.18 (t, J_{PF} = 115 Hz). MS-ESI (*m/z*): $[M - H]^-$ calcd for $C_{13}H_{10}F_2O_3P$, 283.0414; found, 283.0.



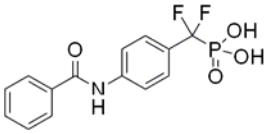
4

Compound 4. ^{1}H NMR (400 MHz, DMSO-d_6) δ 3.87 (s, 3H), 7.48 (d, J = 8.0 Hz, 2H), 7.66 (d, J = 8.4 Hz, 2H), 7.92 (s, 1H), 8.21 (s, 1H); ^{19}F NMR (376 MHz, DMSO-d_6): δ -107.00 (d, J_{PF} = 108 Hz); ^{31}P NMR (162 MHz, DMSO-d_6): δ 3.04 (t, J_{PF} = 109 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{11}H_{12}F_2N_2O_3P$, 289.0548; found, 289.0549.



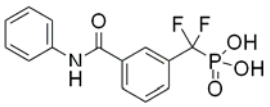
5

Compound 5. ^{1}H NMR (400 MHz, CD_3OD): δ 7.52 (2H, d, J = 8.0 Hz), 7.65 (2H, d, J = 8.4 Hz); ^{19}F NMR (376 MHz, CD_3OD): δ -110.49 (d, J = 110.9 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.15 (t, J = 109.9 Hz). MS-ESI (*m/z*): $[M + H]^+$ calcd for $C_7H_7BrF_2O_3P$, 286.9206; found 287.0.



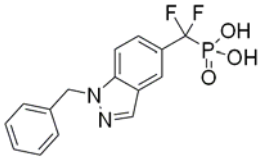
7

Compound 7. ^{1}H NMR (400 MHz, CD_3OD): δ 7.93 (d, J = 7.6 Hz, 2H), 7.83 (d, J = 8.0 Hz, 2H), 7.61-7.49 (m, 5H); ^{19}F NMR (376 MHz, CD_3OD): δ -111.17 (d, J_{PF} = 113 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 3.31 (t, J_{PF} = 113 Hz). HRMS-FAB (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{13}F_2NO_4P$, 328.0550; found, 328.0545.



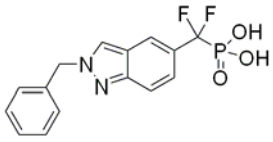
8

Compound 8. ^{1}H NMR (400 MHz, CD_3OD): δ 7.16 (t, J = 7.4 Hz, 1H), 7.37 (t, J = 7.8 Hz, 2H), 7.64 (t, J = 7.8 Hz, 1H), 7.70 (d, J = 7.6 Hz, 2H), 7.82 (d, J = 7.6 Hz, 1H), 8.06 (d, J = 7.6 Hz, 1H), 8.15 (s, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.56 (d, J_{PF} = 110.2 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 4.30 (t, J_{PF} = 110.4 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{13}F_2NO_4P$, 328.0545; found, 328.0551.



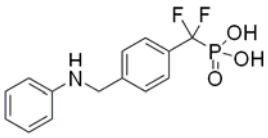
9

Compound 9. ^{1}H NMR (400 MHz, CD_3OD): δ 5.68 (s, 2H), 7.16-7.32 (m, 5H), 7.62 (s, 2H), 8.06 (s, 1H), 8.17 (s, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -108.15 (d, J_{PF} = 114.7 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 5.22 (m). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{14}F_2N_2O_3P$, 339.0705; found, 339.0714.



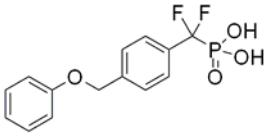
10

Compound 10. ^{1}H NMR (400 MHz, CD_3OD): δ 5.67 (s, 2H), 7.28-7.39 (m, 5H), 7.52 (d, J = 9.2 Hz, 1H), 7.67 (d, J = 8.8 Hz, 1H), 8.02 (s, 1H), 8.44 (s, 1H); ^{19}F NMR (376 MHz, CD_3OD): δ -108.72 (d, J_{PF} = 114.7 Hz); ^{31}P NMR (162 MHz, CD_3OD) δ 5.05 (t, J_{PF} = 116.0 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{14}F_2N_2O_3P$, 339.0705; found, 339.0711.



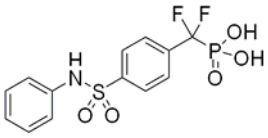
40

Compound 40. ^{1}H NMR (400 MHz, CD_3OD): δ 7.48 (*AB*q, *J* = 8.4 Hz, 4H), 7.10 (t, *J* = 7.8 Hz, 2H), 6.68 (d, *J* = 8.0 Hz, 2H), 6.64 (t, *J* = 7.4 Hz, 1H), 4.34 (s, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -106.70 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 2.78 (t, J_{PF} = 108 Hz). HRMS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{14}H_{15}F_2NO_3P$, 314.0758; found, 314.0766.



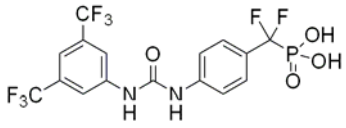
41

Compound 41. ^{1}H NMR (400 MHz, CD_3OD): δ 7.60 (d, *J* = 8.1 Hz, 2H), 7.47 (d, *J* = 8.0 Hz, 2H), 7.24 (t, *J* = 7.9 Hz, 2H), 6.93-6.90 (m, 3H), 5.08 (s, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -107.45 (d, J_{PF} = 109 Hz); ^{31}P NMR (162 MHz, CD_3OD): δ 7.01 (br m). HRMS-ESI (*m/z*): $[M + Na]^+$ calcd for $C_{14}H_{13}F_2O_4PNa$, 337.0412; found, 337.0415.



42

Compound 42. ^{1}H NMR (400 MHz, CD_3OD): δ 7.04-7.13 (m, 3H), 7.19-7.25 (m, 2H), 7.70 (d, *J* = 8.4 Hz, 2H), 7.85 (d, J = 8.4 Hz, 2H); ^{19}F NMR (376 MHz, CD_3OD): δ -110.80 (d, J_{PF} = 106.0 Hz). MS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{13}H_{13}F_2NO_5PS$, 364.0142; found, 364.0.



43

Compound 43. ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.7 (br s, 1H), 10.1 (br s, 1H), 8.22 (m, 2H), 8.10 (m, 2H), 7.61 (m, 1H), 7.32 (m 2H). MS-ESI (*m/z*): $[M + H]^+$ calcd for $C_{16}H_{12}F_8N_2O_4P$, 479.0239; found, 479.0.

4. Expression and purification of PtpA.

The gene for PtpA was amplified from *Mtb* genomic DNA and cloned into the pET28b vector (Novagen). Protein was expressed in BL21(DH3) cells (Invitrogen). Transformed bacteria were grown to an OD600 of 0.8 in terrific broth and protein expression was induced by the addition of 100 uM isopropyl β-D-1-thiogalactopyranoside. After 18 hours of expression at 20 °C, cells were harvested and resuspended in buffer A (20 mM Tris pH 7.5, 50 mM NaCl), and protease inhibitor AEBSF. Cell suspensions were sonicated, the lysates centrifuged for 1 hour at 15,000 g, and the cleared lysate loaded onto a metal affinity column. After elution with an imidazole gradient, the His6 tag was cleaved by treatment with 1:1,000 (w/w) trypsin for 10 minutes at rt. The protein was further purified by gel filtration on a S75 Superdex column in buffer A and concentrated to 2.3 mg/mL for phosphatase assays.

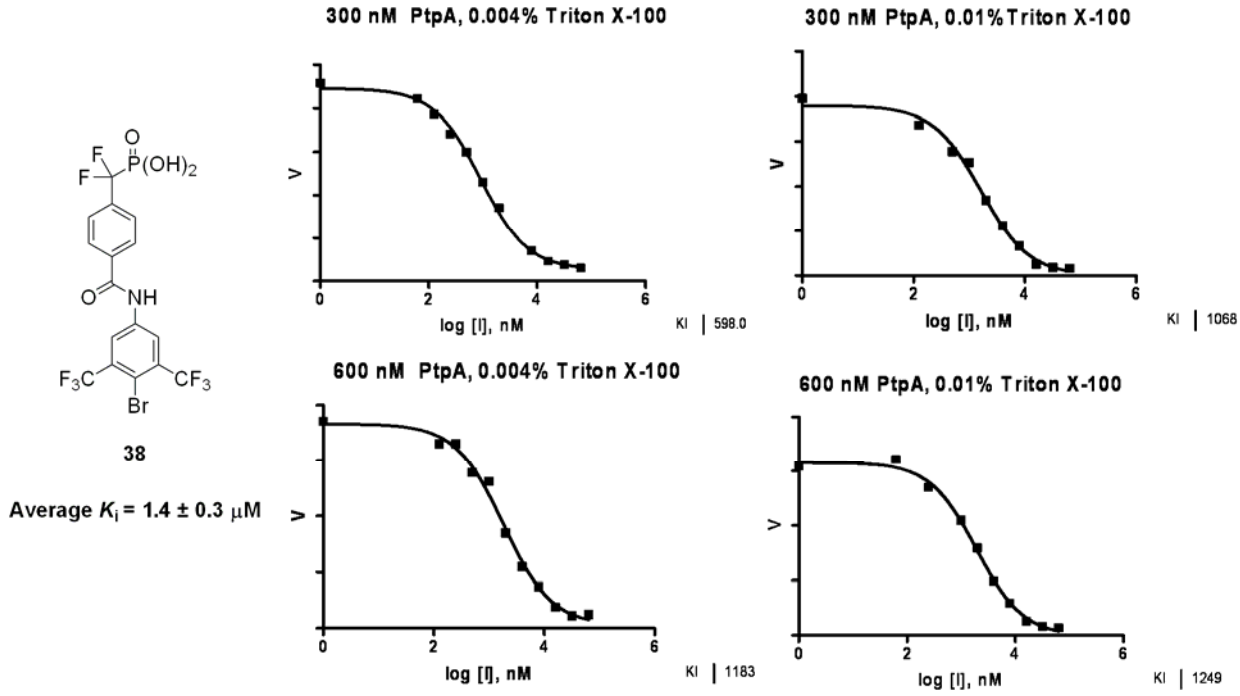
5. Determination of inhibitor K_i.

96-well plates were used to run K_i assays, with reaction volumes of 100 μL per well. 45 μL of water was added to each well, followed by 20 μL of sodium citrate buffer (stock solution: 100 mM sodium citrate, pH 6.2, 0.02% Triton X-100), 5 μL of 20 mM EDTA stock solution, 5 μL of 20 mM DTT stock solution, and 10 μL of 1 μM PtpA stock solution. Then 5 μL of the appropriate inhibitor stock solutions, serially diluted 2-fold for a total of 10 different concentrations in DMSO, was added to the wells, and the plate was covered and incubated at 37 ^{o}C in a UV-Vis plate reader. The reaction was started after 5 minutes of incubation by addition of 10 μL of 2 mM pNPP substrate stock, and reaction progress was monitored at 405 nm with continued incubation at 37 ^{o}C. The initial rate data collected was

used for the determination of K_i values. The kinetic values were obtained from nonlinear regression of substrate-velocity curves in the presence of various concentrations of inhibitor using the equation $v=V_{max}*[S]/K_M(1+[I]/K_i)+[S]$.

6. K_i of inhibitor 38.

Inhibitor **38** K_i values were determined using at least four independent measurements. Assays were run at two enzyme concentrations (300 and 600 nM) and two concentrations of Triton X-100 (0.004% and 0.01%). Representative K_i curves for benzanilide inhibitor **38** are shown. The "log[0]" point corresponds to no inhibitor added, included as a control.



7. Inhibitor-PtpA modeling.

Receptor relaxation. The X-ray crystal structure of PtpA (PDB ID 1U2P) was used for all modeling studies.[1] To allow the protein to relax, a short molecular dynamics simulation was run in AMBER 9.0 using the ff03 force field.[2,3] The structure was prepared by removing crystallographic waters and adding an 8.0 Å octagon of TIP3P water and sodium ions using the LEaP accessory.[2] The system was minimized, slowly melted to 300K, and allowed to equilibrate for 25 ps. The simulation was continued for an additional 150 ps and selected as the final snapshot for docking.

Active site identification. To identify where inhibitors might bind, waters were removed from the structure produced by the molecular dynamics simulation, and protein surface invaginations were identified using spheres generated by the DOCK accessory SPHGEN.[4] The putative binding site was characterized by selecting all spheres within a 12 Å radius of the chlorine atom bound to the active-site cysteine nucleophile in the X-ray structure.

Receptor preparation for docking. An octagon of TIP3P waters was built around the receptor using the Chimera AmberTools module, followed by removal of any waters >5 Å from any receptor atom, resulting in approximately two shells of water molecules.[5] All waters $\leq$3 Å from the active site spheres described above were then removed. The Chimera (version 1.3) Dock Prep module was used to complete the receptor preparation.[6,7] To account for the receptor contribution to the score during DOCKing, grids were precomputed to store the van der Waals and electrostatic values for the receptor using the DOCK accessory GRID.[7]

Compound preparation for docking. To validate observed structure-activity relationships, structures **20**, **38**, and **43** were docked onto PtpA. Each compound was drawn and converted to SMILE strings using the JME molecular editor.[8] The SMILE strings were used to create rotamer ensembles of three-dimensional structures in OMEGA.[9] All generated conformations were kept for docking, resulting in any average of 11 conformations per compound. Each conformation was protonated and assigned AM1-BCC charges using the Chimera (version 1.3) AddH and AddCharge modules.[10]

Docking procedure. The compound conformations were docked using Grid Score in DOCK 6.4 using default parameters.[9] Each conformation was then rescored and ranked using the PB/SA score. The top scoring conformation for each compound was used for comparisons. For compounds **20**, **39**, and **43**, molecular dynamics simulations were also performed on the top-scoring conformation to explore the validity of the docked poses. The same protocol described in the "Receptor Relaxation" section was used. All three simulations were equilibrated after 100 ps, and the simulations were run for a further 50 ps. The snapshot with the ligand heavy atom RMSD closest to the docked pose from the final 50 ps of each simulation was selected for analysis.

8. Modeling figures.

View of adjacent enzyme pocket. An unfilled pocket was observed upon docking inhibitors into the active site of PtpA (Figure S1). Extension of functionality into this pocket may provide an avenue for further compound improvement.

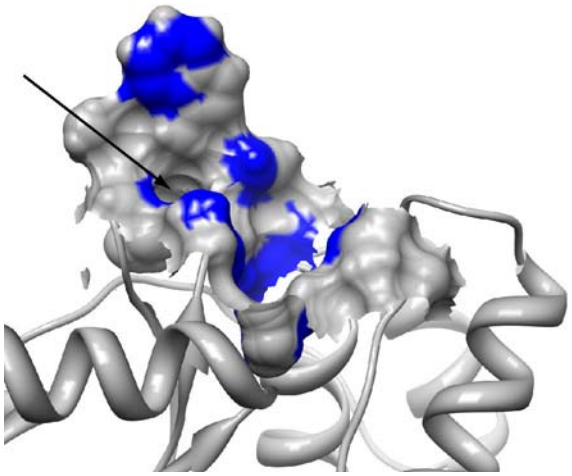


Figure S1. PtpA binding pocket with inhibitor-enzyme contact points shown. The arrow indicates the position of an unfilled enzyme pocket adjacent to the docked inhibitor.

PtpA structure overlays. The crystal structure of PtpA[1] was overlayed with the structures of HCPtpA (PDB ID 5PNT) and PtpB (PDB ID 1YWF), using Chimera,[7,11] for comparison of the PTP active-site P-loop and the variable loop of each enzyme (Figure S2). PtpA and HCPtpA have high homology in both loops, but differ in the residue located at position 48 (Trp for PtpA, Tyr for HCPtpA), which was predicted to be important for pi-stacking interactions with our inhibitors. PtpA and PtpB have differences in the P-loop and differ significantly in the sequence and position of the variable loop. These distinctions likely account for the high selectivity of inhibitor **38** for PtpA over PtpB.

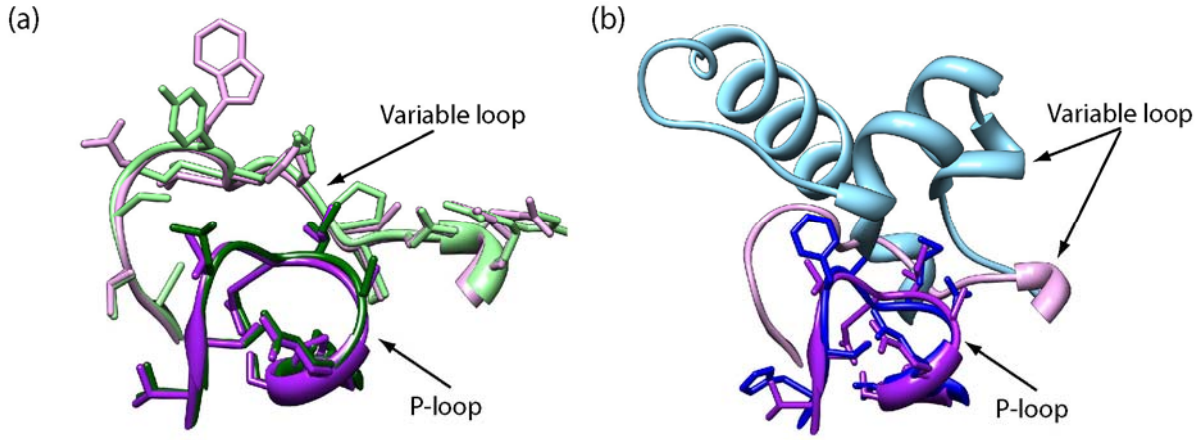


Figure S2. Structure overlay of *Mtb* PtpA (purple) with (a) HCPtpA (green) and (b) *Mtb* PtpB (blue).

9. References

1. Madhurantakam, C.; Rajakumara, E.; Mazumdar, P. A.; Saha, B.; Mitra, D.; Wiker, H. G.; Sankaranarayanan, R.; Das,

A. K. *J. Bacteriol.* **2005**, *187*, 2175-2181.
2. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999-2012.
3. The ff03 force field was designed by and is available from: Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P., *AMBER 9*, University of California, San Francisco, **2006**.
4. Lee, M. C.; Duan, Y. *Proteins* **2004**, *55*, 620-634.
5. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
6. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269-88.
7. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605-1612.
8. Moustakas, D. T.; Lang, P. T.; Pegg, S.; Pettersen, E.; Kuntz, I. D.; Brooijmans, N.; Rizzo, R. C. *J. Comput. Aided Mol. Des.* **2006**, *20*, 601-619.
9. Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA-Publ. RNA Soc.* **2009**, *15*, 1219-1230.
10. Shoichet, B.; Bodian, D. L.; Kuntz, I. D. *J. Comp. Chem.* **1992**, *13*, 380-397.
11. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081).