Supplementary Information

Decrypting a Cryptic Allosteric Pocket in \textit{H. pylori} Glutamate Racemase

Pratik Rajesh Chheda\textsuperscript{1}, Grant T. Cooling\textsuperscript{1}, Sondra F. Dean\textsuperscript{1}, Jonah Propp\textsuperscript{1}, Kathryn F. Hobbs\textsuperscript{2} and M. Ashley Spies\textsuperscript{1,2*}

\textsuperscript{1}Division of Medicinal and Natural Products Chemistry, Department of Pharmaceutical Sciences and Experimental Therapeutics, The University of Iowa, Iowa City, Iowa 52242, United States of America

\textsuperscript{2}Department of Biochemistry, Carver College of Medicine, The University of Iowa, Iowa City, Iowa 52242, United States of America
Supplementary Figures:

**Supplementary Figure 1.**
Acidification of Cα carbon of glutamate substrate due to protonation of Cα-carboxylate oxygen by catalytic Cys-181 is necessary to reduce the pKₐ and enable stereo-inversion.

**Supplementary Figure 2.**
Overlap of compound A pose generated by FlexX onto the co-crystallized ligand in PDB 2JFZ. RMSD of the two poses is 0.883 Å which is lower than the crystal structure resolution of 1.86 Å.
Supplementary Figure 3.

A-C) Representative binding curves of top five hits to glutamate racemase as measured by SPR. Data was fit with a 1:1 binding model shown in orange with raw data shown in black and the $K_d$ was determined by fitting to the same model. Error values represent standard deviation (SD). A) Binding curve for NP-004431: measured $K_d$ of $228 \pm 2 \mu M$. B) Binding curve for NP-004604: measured $K_d$ of $170 \pm 2 \mu M$. C) Binding curve for NP-008029: measured $K_d$ of $910 \pm 10 \mu M$. D-F) Representative IC$_{50}$ curves for natural product hits evaluating inhibitory activity of the hits against H. pylori GR employing a previously established coupled-enzyme assay\(^1\). D) IC$_{50}$ curve for NP-004431: measured IC$_{50}$ of 705.3
µM. **E)** IC$_{50}$ curve for NP-004604: measured IC$_{50}$ of 425.3 µM. **F)** IC$_{50}$ curve for NP-000205: measured IC$_{50}$ of 512.8 µM.

**Supplementary Figure 4.**

**A)** Docking pose of NP-020560 occupying the allosteric binding pocket and highlighting key interactions with *H. pylori* GR; NP-020560 is represented as pink sticks and allosteric residues are depicted as green sticks; protein surface showing the occupancy of allosteric site. **B-E)** Ligand Interaction maps for top hits generated in Molecular Operating Environment (MOE) highlighting key ligand-protein interactions with their nature detailed in legend for panel B. **B)** Ligand interaction map for NP004604. **C)** Ligand interaction map for NP000205. **D)** Ligand interaction map for NP004431. **E)** Ligand interaction map for NP008029.
Supplementary Figure 5.

A) The changes in RMSF (Å) between the inhibitor-bound system (GR-D-Glu-Compound A) and the inhibitor-free system (GR-D-Glu) are shown for MD simulations as a function of the residue number for Monomer A. 

B) The changes in RMSF (Å) between the inhibitor-bound system (GR-D-Glu-NP-020560) and the inhibitor-free system (GR-D-Glu) are shown for MD simulations as a function of the residue number for Monomer B. 

C) The changes in RMSF (Å) between the inhibitor-bound system (GR-D-Glu-Compound A) and the inhibitor-free system (GR-D-Glu) are shown for MD simulations as a function of the residue number for Monomer B. 

D) The changes in normalized B-factors between the inhibitor-bound structure (2JFZ: GR-D-Glu-Compound-A) and the inhibitor-free system (2JFX: GR-D-Glu) are plotted as a function of residue number for Monomer B.
Supplementary Figure 6.

All panels are plotted as function of residue number. **A)** Normalized B-factors for inhibitor-free system (2JFX: GR-D-Glu). **B)** Normalized B-factors for inhibitor-bound structure (2JFZ: GR-D-Glu-Compound-A). **C)** RMSF (Å) for monomer A of inhibitor-bound system (GR-D-Glu-Compound A). **D)** RMSF (Å) for monomer B of inhibitor-bound system (GR-D-Glu-Compound A). **E)** RMSF (Å) for monomer A of inhibitor-free system (GR-D-Glu). **F)** RMSF (Å) for monomer B of inhibitor-free system (GR-D-Glu). The oval highlights residues the uninhibited state of GR shows increased flexibility, specifically in residues 240-255 which form the C-terminal α-helix and a major part of the allosteric inhibitor binding site.
Supplementary Figure 7.

Topology map of the difference between the DCCM for the uninhibited system (GR-D-Glu) and the inhibited system (GR-D-Glu-Compound A). The difference DCCM topology map, from subtracting the DCCM for GR-D-Glu-compound A from the DCCM for the native GR-D-Glu. Positive changes in coupled motion represent motions that are lost upon binding of the allosteric inhibitor, compound A. The salient pattern is very similar to what is seen for NP-020560, showing a loss in coupled motion between the monomers.

Supplementary Figure 8.
A) DCCM map for uninhibited system (GR-D-Glu). B) DCCM map for inhibited system (GR-D-Glu-NP 020560). C) DCCM map for inhibited system (GR-D-Glu-Compound A).

Supplementary Figure 9.

Root Mean Square Deviation (RMSD) of amino acid backbone during Molecular Dynamic simulation run for A. 2JFZ and B. 4B1F. 2JFZ achieves equilibrium around 60 ns while 4B1F achieves equilibrium around 10 ns.

Supplementary Table

| Smiles Strings                                                                 | Active or Decoy |
|-------------------------------------------------------------------------------|-----------------|
| CC(C)CN1C(=O)N(C)C(=O)c2c1nn(Cc3cccc4cccccc34)c2c5ccncc5                  | Active          |
| CN1C(=O)c2c(nn(Cc3ccncc4ccc(Cl)cc34)c2c5cc(Cn5C)n(C6CC6)C1=O               | Active          |
| CN1C(=O)c2c(nn(Cc3ccnc4ccc(Cl)cc34)c2c5cc(o5)S(C)(=O)=O)N(C6CC6)C1=O      | Active          |
| CN1C(=O)c2c(nn(Cc3ccnc4ccc(Cl)cc34)c2c5cc(cn5C)S(C)(=O)=O)N(C6CC6)C1=O    | Active          |
| Cc1oc(cc1c2c3C(=O)N(C)C(=O)N(CC4CC4)c3nn2Cc5ccncc6ccc(Cl)cc56)S(C)(=O)=O | Active          |
| Cc1oc(cc1c2c3C(=O)N(C)C(=O)N(CC4CC4)c3nn2Cc5cncc6ccc(Cl)cc56)S(C)(=O)=O  | Active          |
| CN1C(=O)c2c(nn(Cc3ccncc4ccc(Cl)cc34)c2c5cc(cn5C)S(N)(=O)=O)N(CC6CC6)C1=O | Active          |
| CN1C(=O)c2c(nn(Cc3ccncc4ccc(Cl)cc34)c2c5cc(cn5C)S(N)(=O)=O)N(CC6CC6)C1=O | Active          |
| CNS(=O)                                                                     | Active          |
| CNS(=O)                                                                     | Active          |
| CNS(=O)                                                                     | Active          |
| CNS(=O)                                                                     | Active          |
| CNS(=O)                                                                     | Active          |
| CN1C(=O)c2c(nn(Cc3ccncc4ccc(Cl)cc34)c2c5cc(cn5C)S(C)(=O)=O)N(CC6CC6)C1=O | Active          |
Supplementary Table 1. SMILES strings for the library of 65 active pyrazolopyridimidinedione analogs and respective inactive/decoy compounds generated using the Database of Useful Decoys – Enhanced (DUD–E) website that were used for ROC analysis.

References:

1. Rej, R. A convenient continuous-rate spectrophotometric method for determination of amino acid substrate specificity of aminotransferases: application to isoenzymes of aspartate aminotransferase. *Analytical biochemistry* **119**, 205-210 (1982).