Biological and structural evaluation of 10*R*- and 10*S*-methylthio-DDACTHF reveals a new role for sulfur in inhibition of glycinamide ribonucleotide transformylase

Stephen Connelly[†], *Jessica K. DeMartino*[‡], *Dale L. Boger*[‡,§] *and Ian A. Wilson*[*,†,§]

Departments of Integrative Structural and Computational Biology[†], Chemistry[‡] and The Skaggs Institute of Chemical Biology[§], The Scripps Research Institute, 10550 North Torrey Pines Rd., La Jolla, CA 92037

* To whom correspondence should be addressed: wilson@scripps.edu

Supporting Information

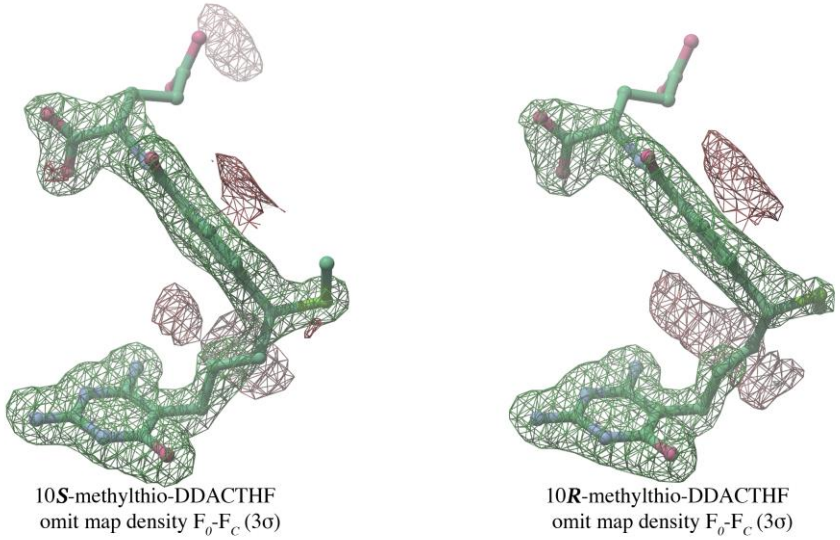

Figure S1. Omit map electron density (F_o-F_c at 3σ) for both the 10*R*- and 10*S*- methylthio-DDACTHF diastereoisomers.

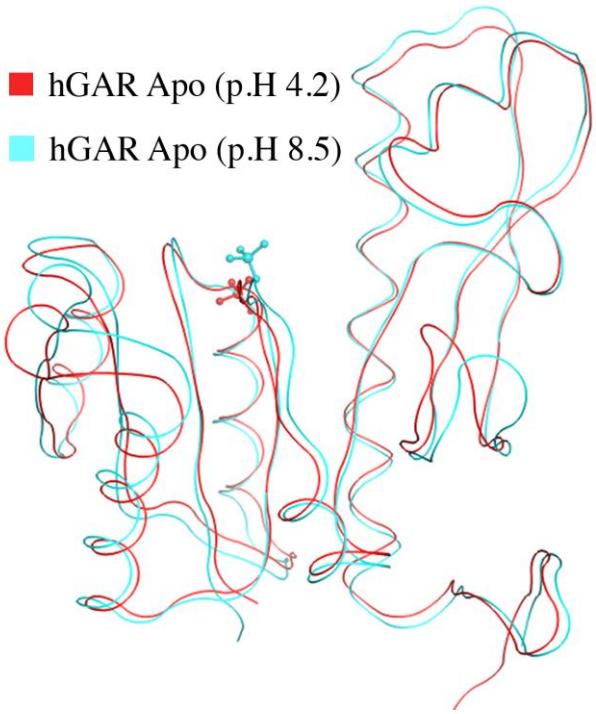

Figure S2. Superposition of apo human GAR Tfase at pH 4.2 (red) with apo human GAR Tfase at pH 8.5 (1MEJ) (light blue). All Cα atoms superpose to 1.2 Å rmsd.

***In-silico* modeling of sulfur-containing hGAR Tfase inhibitors**

In silico docking of the three sulfur-containing folate analogues **9**, **10** and **11** was completed using the Dock simulation in the program MOE (Molecular Operating Environment 2011.10, Chemical Computing Group, Montréal, Canada). The high-resolution coordinates for hGAR Tfase in complex with 10*S*-methylthio-DDACTHF **7** (4EW2), refined with anisotropic *B*-values and hydrogens in riding positions, were used as the template (receptor) coordinates. These hydrogens were added and optimized using the Protonate 3D function within MOE, which assigns protonation from a discrete collection of states. The Generalized Born/Volume Integral (GBVI) electrostatics model was used for longer-range interactions and solvation effects (*1, 2*). Only the two conserved structural water molecules, which play a vital role in binding the pteridine ring, were retained for the simulation (*3*). The target pocket was defined as atoms within 5 Å of the bound ligand 10*S* **7**, which was removed prior to the docking simulation. The MOE Dock simulation uses a number of discrete steps for each docked conformation before a final scoring stage where the results are ranked in a database of 3D coordinates, each of which are described in brief below.

1. Conformational analysis. Dock generates conformations from a single 3D ligand conformer by applying a collection of preferred torsion angles to the rotatable bonds, without altering bond lengths or angles.

2. Placement. A collection of poses is generated from the pool of ligand conformations using the default Alpha Triangle method, in which poses are generated by superposition of ligand atom triplets and triplets of receptor site points. The receptor site points are *alpha sphere* centers, which represent locations of tight packing. During each subsequent iteration, a random conformation is selected, and a random triplet of ligand atoms and a random triplet of alpha sphere centers are used to determine the pose.

3. Initial scoring. Poses generated by the placement methodology are scored. Typically, scoring functions emphasize favorable hydrophobic, ionic and hydrogen bond contacts. For this stage, the London dG scoring function is used (supplemental equation S1) which estimates the free energy of binding of the ligand from a given pose.

4. Refinement. For final pose refinement, the molecular mechanics force field MMF94x was used. By default, backbone atoms are held fixed during refinement, but side chains of the receptor are partially tethered. The weights of the tethers are determined from the individual atom temperature factors (*B*-values). Tethering using *B*-values has the following interpretation: an atom's root-mean-square deviation from its equilibrium position will be equal to the given *B*-value if the atom is floating freely in the potential well of the tether at 300 K.

5. *Final Scoring*. The final poses were then rescored using the GBVI/VSA dG (Generalized Born Volume Integral/Van der Waals Surface Area) scoring function (supplemental Figure S2). This is a force field-based scoring function, which estimates the free energy of binding of the ligand from a given pose. It has been trained using the MMFF94x and AMBER99 force fields on the 99 protein-ligand complexes of the solvated interaction energy (SIE) training set (*4, 5*). The coordinate positions of the top scoring position are discussed herein.

A)

$$\Delta G = c + E_{flex} + \sum_{h-bonds} c_{HB} f_{HB} + \sum_{m-lig} c_M f_M + \sum_{atoms\ i} \Delta D_i$$

B)

$$\Delta D_i = c_i R_i^3 \left\{ \iiint_{u \notin A \cup B} |u|^{-6}\, du - \iiint_{u \notin B} |u|^{-6}\, du \right\}$$

Equation S1. The London dG scoring function, which estimates the free energy of binding of the ligand. A) where c represents the average gain/loss of rotational and translational entropy; E_{flex} is the energy due to the loss of flexibility of the ligand (calculated from ligand topology only); f_{HB} measures geometric imperfections of hydrogen bonds and takes a value in [0,1]; c_{HB} is the energy of an ideal hydrogen bond; f_M measures geometric imperfections of metal ligations and takes a value in [0,1]; c_M is the energy of an ideal metal ligation; and D_i is the desolvation energy of atom i. The difference in desolvation energies is calculated according to the formula. B) where A and B are the protein and/or ligand volumes with atom i belonging to volume B; R_i is the solvation radius of atom i (taken as the OPLS-AA van der Waals sigma parameter plus 0.5 Å); and c_i is the desolvation coefficient of atom i. The coefficients $\{c, c_{HB}, c_M, c_i\}$ were fitted from ~400 x-ray crystal structures of protein-ligand complexes with available experimental pKi data. Atoms are categorized into about a dozen atom types for the assignment of the c_i coefficients. The triple integrals are approximated using Generalized Born integral formulas.

$$\Delta G \approx c + \alpha\left[\frac{2}{3}(\Delta E_{coul} + \Delta E_{sol}) + \Delta E_{vdw} + \beta \Delta SA_{weighted}\right]$$

Equation S2. The GBVI/WSA dG (Generalized Born Volume Integral/van der Waals Surface Area) scoring function, which estimates the free energy of binding of the ligand from a given pose.

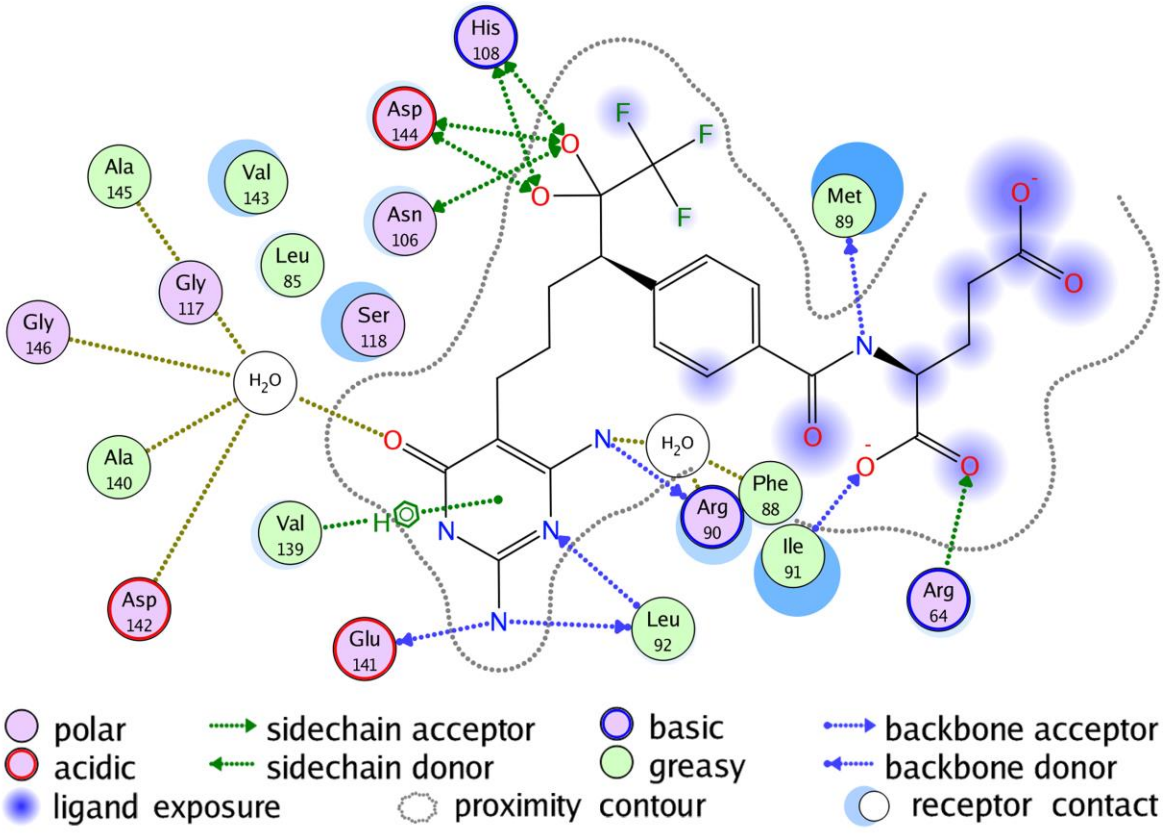

Figure S3. Schematic 2D-projection of the ligand interactions of 10-CF_3CO-DDATHF with GAR Tfase. Legend depicting ligand/receptor contacts as amino-acid type, hydrogen bond donor/acceptor origin, ligand exposure and proximity contour. Bound water molecules that mediate contacts between ligand and receptor are also shown. Figure generated using MOE 2011.10.

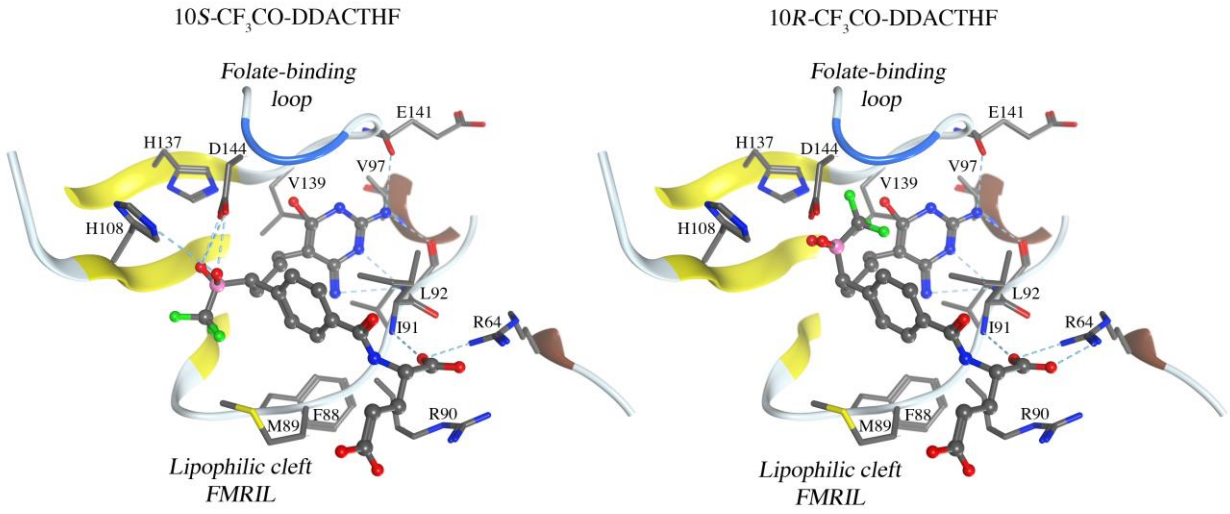

Figure S4. Comparison of C10 diasteriosomers of 10CF_3CO-DDACTHF (shown in ball and stick) bound to human GAR Tfase, with C10 center colored in pink. **A**) Crystal structure of human GAR Tfase with the observed 10*S*-CF_3CO-DDACTHF isomer bound (1NJS) and **B**) the alternate isomer modeled in showing the 10*R* conformation clashes with the folate-binding loop of 141-146 explaining why this is not observed in F_o-F_c density.

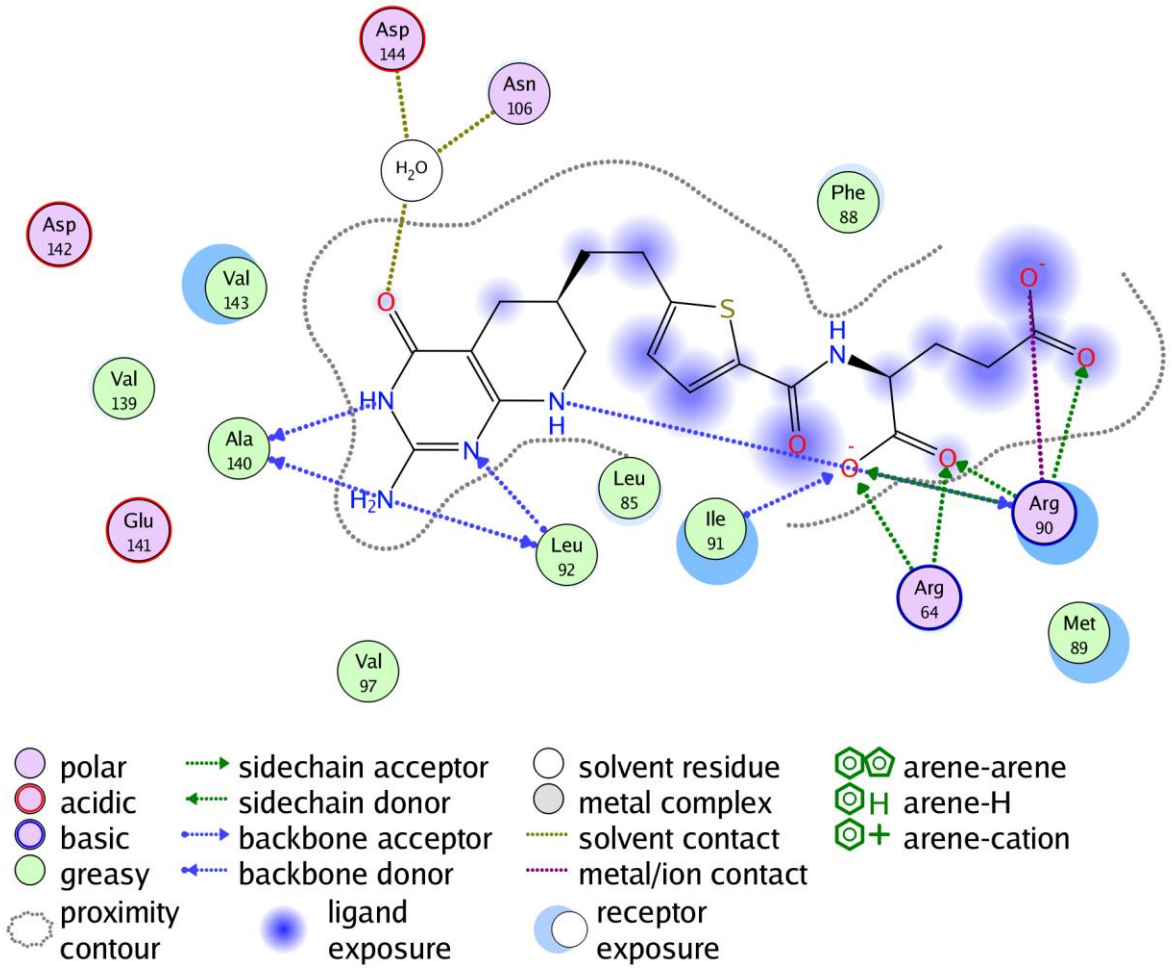

Figure S5. Schematic 2D-projection of the ligand interactions of LY309887. Legend depicting ligand/receptor contacts as amino-acid type, hydrogen bond donor/acceptor origin, ligand exposure and proximity contour. Bound water molecules that mediate contacts between ligand and receptor are also shown. Figures generated using MOE 2011.10.

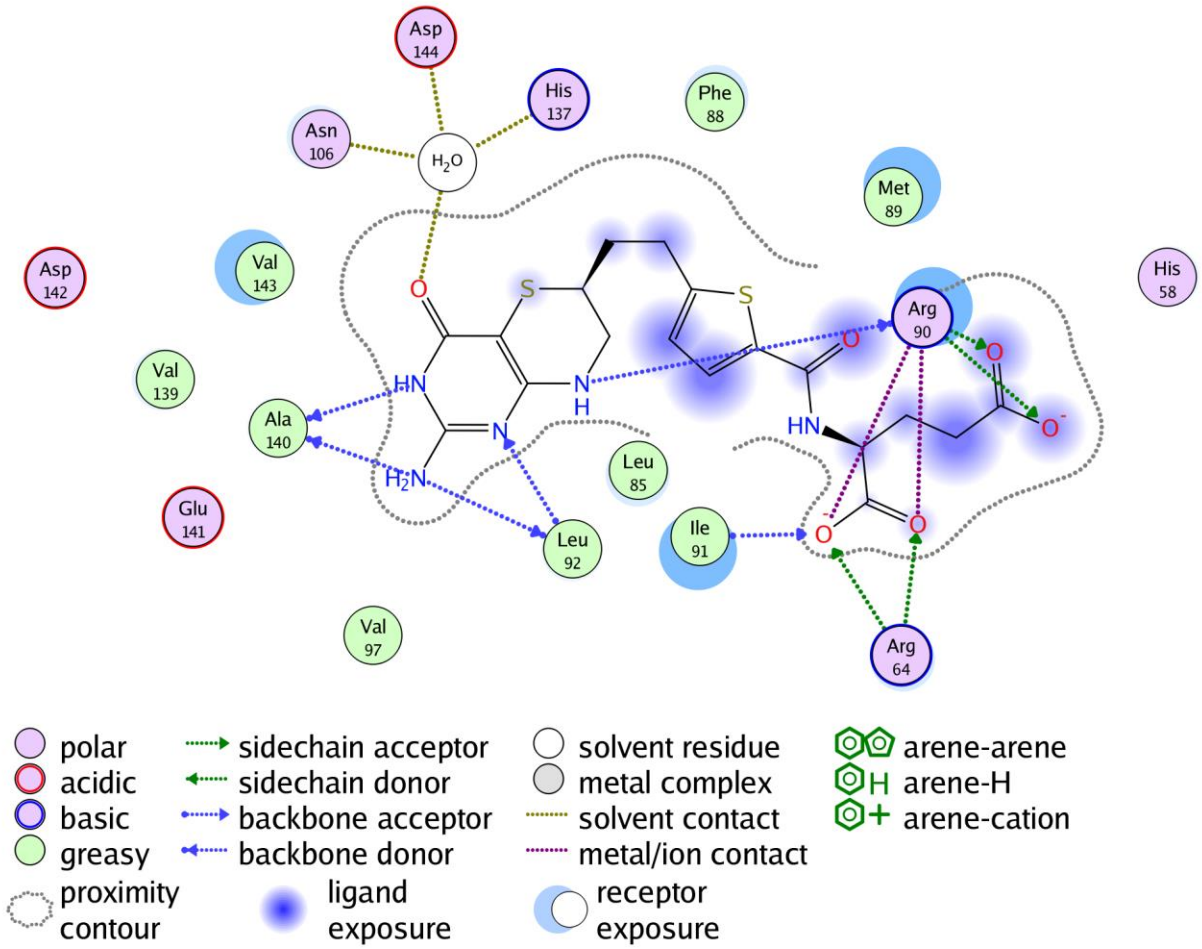

Figure S6. Schematic 2D-projection of the ligand interactions of AG2034 Legend depicting ligand/receptor contacts as amino-acid type, hydrogen bond donor/acceptor origin, ligand exposure and proximity contour. Bound water molecules that mediate contacts between ligand and receptor are also shown. Figures generated using MOE 2011.10.

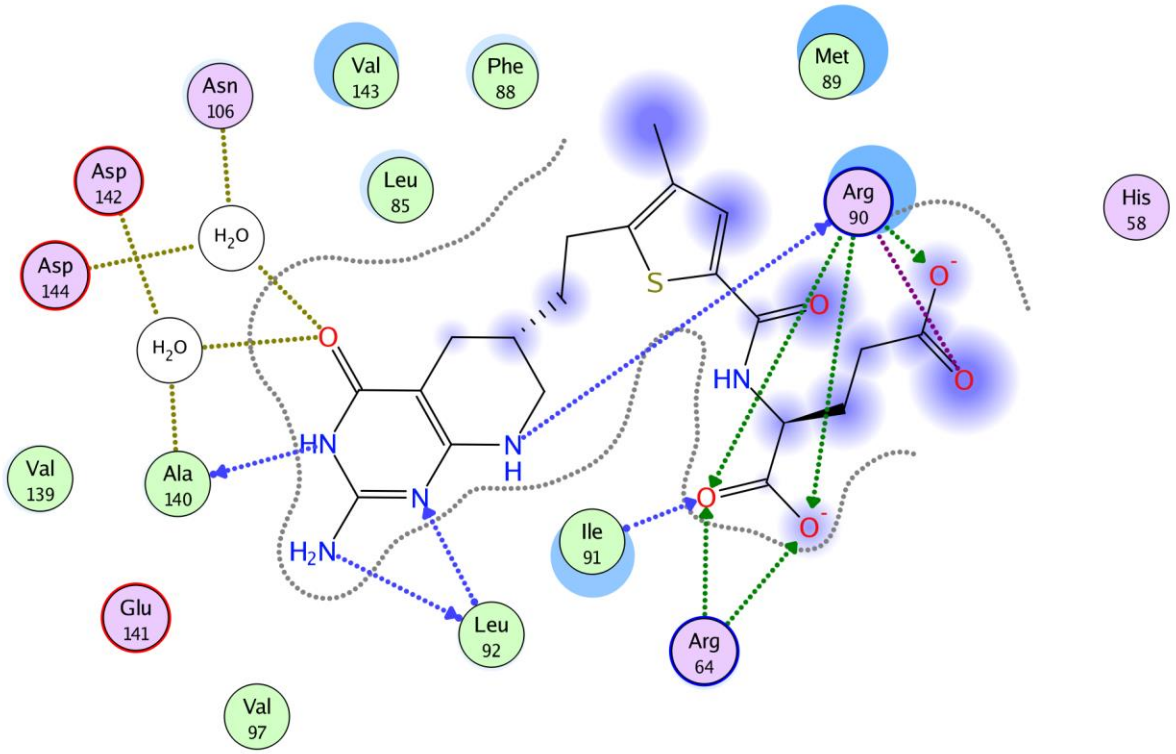

Figure S7. Schematic 2D-projection of the ligand interactions of AG2037 Legend depicting ligand/receptor contacts as amino acid type, hydrogen bond donor/acceptor origin, ligand exposure and proximity contour. Bound water molecules that mediate contacts between ligand and receptor are also shown. Figures generated using MOE 2011.10.

References

1. Labute, P. (2008) The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area, *J Comput Chem 29*, 1693-1698.

2. Labute, P. (2009) Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures, *Proteins 75*, 187-205.

3. Zhang, Y., Desharnais, J., Greasley, S. E., Beardsley, G. P., Boger, D. L., and Wilson, I. A. (2002) Crystal structures of human GAR Tfase at low and high pH and with substrate β-GAR, *Biochemistry 41*, 14206-14215.

4. Naim, M., Bhat, S., Rankin, K. N., Dennis, S., Chowdhury, S. F., Siddiqi, I., Drabik, P., Sulea, T., Bayly, C. I., Jakalian, A., and Purisima, E. O. (2007) Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space, *J Chem Inf Model 47*, 122-133.

5. Feher, M., and Williams, C. I. (2009) Effect of input differences on the results of docking calculations, *J Chem Inf Model 49*, 1704-1714.