Supporting Information

A Binding Pose Flip Explained via Enthalpic and Entropic Contributions

Michael Schauperl[1], Paul Czodrowski[2], Julian E. Fuchs[1], Roland G. Huber[3], Birgit J. Waldner[1], Maren Podewitz[2], Christian Kramer[1], Klaus R. Liedl[1]*

[1] Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80-82, A-6020 Innsbruck, Tyrol, Austria

[2] Discovery Technologies, Merck Serono Research, Merck Serono R&D, Merck KGaA, Frankfurter Strasse 250, 64293 Darmstadt, Germany

[3] Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), #07-01 Matrix, 30 Biopolis Street, 138671 Singapore

*Klaus.Liedl@uibk.ac.at

Ligand partial charges:

All charges were obtained with the AM1-BCC[1, 2] charge model using the antechamber tool of the AMBER software package.

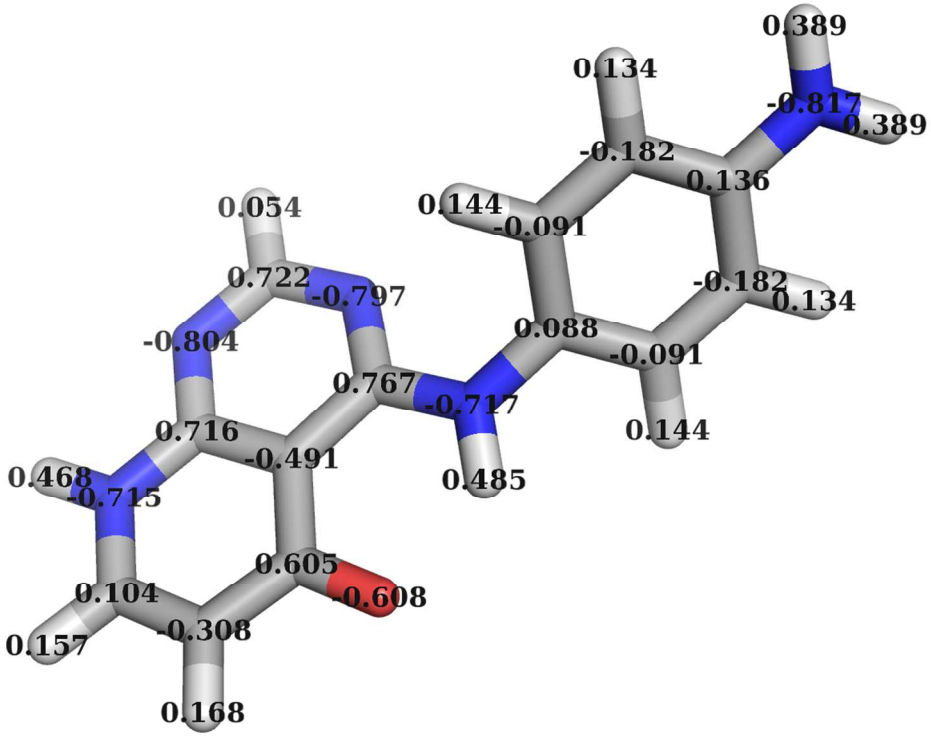

Figure S1: Partial charges of Compound C in the neutral form.

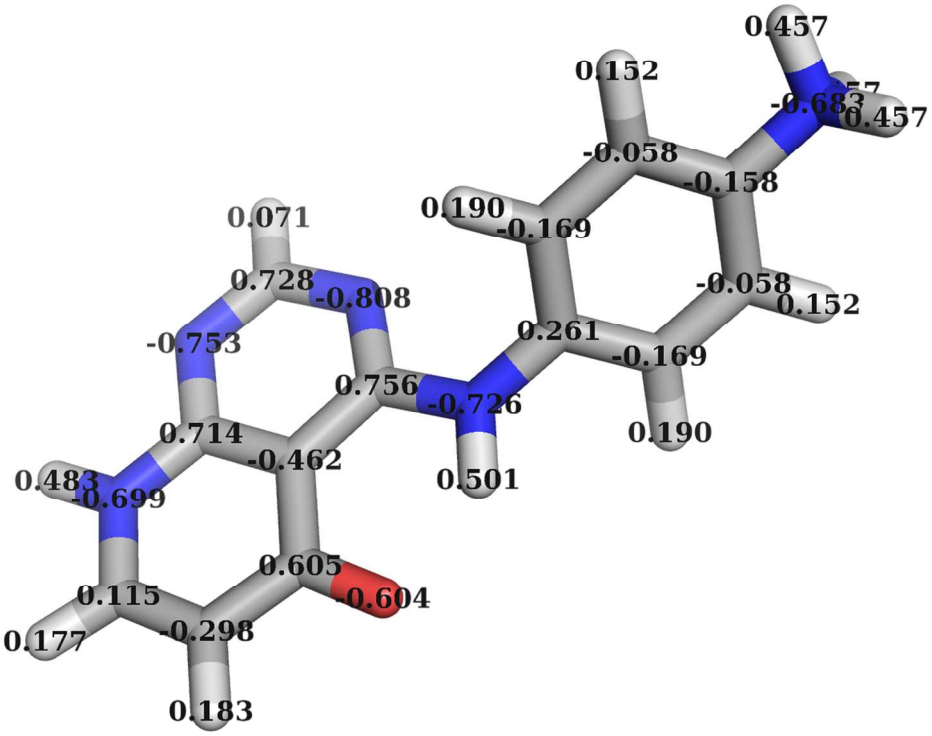

Figure S2: Partial charges of Compound C in the charged (+1) form.

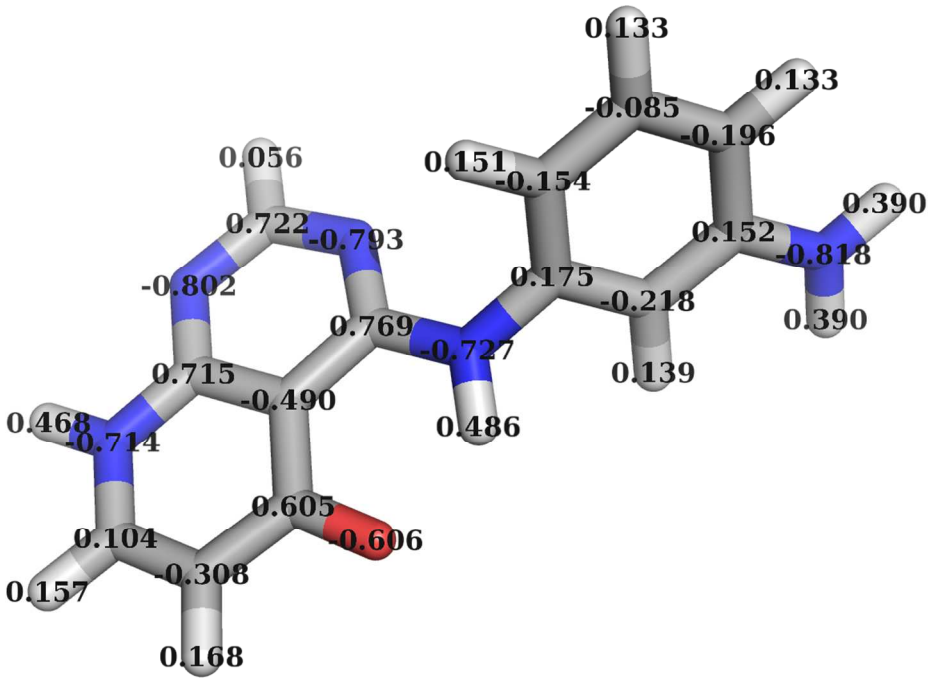

Figure S3: Partial charges of Compound D in the neutral form.

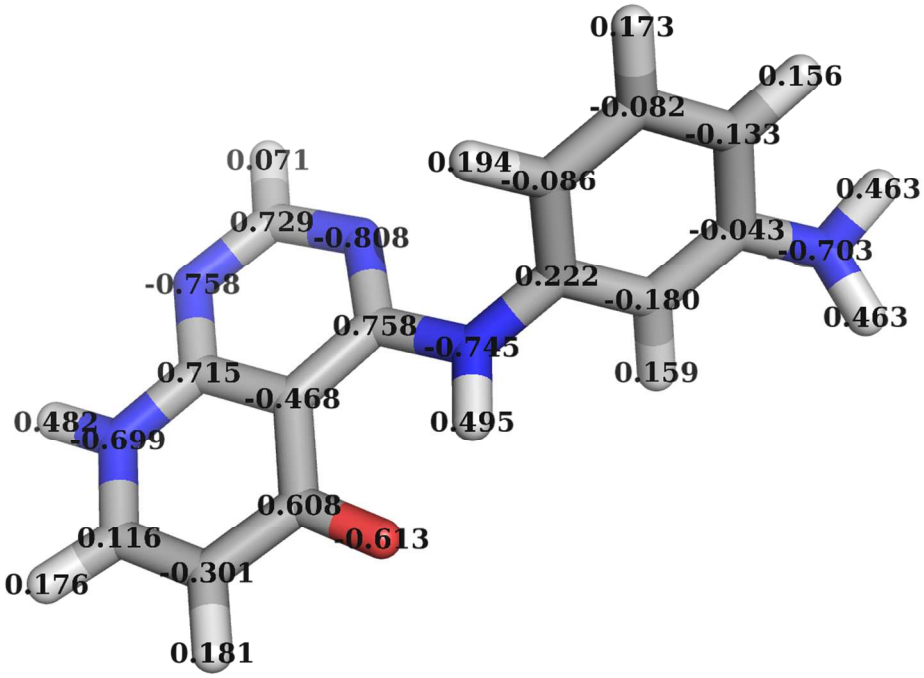

Figure S4: Partial charges of Compound D in the charged (+1) form.

MM/GB(PB)SA calculations

MM-GBSA and MM-PBSA energies were calculated for the eight configurations outlined in Table S1. Compounds C and D were each placed in pose C and D, respectively. These calculations were performed with neutral and positively charged configurations of compound C and D. Free energy calculations were performed using *MMPBSA.py* [3] as included in Amber14.[4] Calculation method and Born radii were chosen according to the GB set II described by Onufriev et al.[5], which is implemented in Amber 14 as *mbondi2* and *igb=5* respectively. MM-PB(GB)SA calculations were performed on 100 equally spaced frames extracted from the trajectory described earlier using a single-trajectory approach.[6]

Table S1. Results of MM-GBSA and MM-PBSA energy calculations in kcal/mol.

MM-PBSA		**Ligand**			
		Neutral		**Positive**	
		C	**D**	**C**	**D**
Pose	**C**	-2.43	-4.16	-5.43	-7.37
	D	-8.64	-15.8	-5.98	-11.42
MM-GBSA		**Ligand**			
		Neutral		**Positive**	
		C	**D**	**C**	**D**
Pose	**C**	-29.68	-30.70	-32.04	-31.64
	D	-36.73	-40.57	-34.20	-42.81

MM/GB(PB)SA calculations predict the neutral form of compound C in pose D to be significantly more stable than pose C, which is in contradiction to the experimentally observed findings, where compound C crystallizes in pose C. Only for the positively charged compound C results are in agreement with the experimental findings; both poses are similar in terms of free energy. Therefore, this method strongly supports our hypothesis that compound C is protonated in complex, as otherwise compound C would also likely to be crystallized in pose D.

Additional enthalpy calculations

In the main manuscript we used the LIE[7, 8] tool to examine the differences in interaction energy between the ligand and its environment for the bound and unbound state. An alternative approach would be to only calculate the interaction energy between the ligand and the protein, as the interaction between the ligand and the protein is not covered by GIST but the interaction between the water and the solutes is. This seems to be a valid approach at first glance. However, the results contradict our suggested mechanism, as we find that for compound C that pose D (-128.9 kcal/mol vs -102.5 kcal/mol) is the more stable one (Table S2; column LIG-PRO).

At second glance we see that we are still missing one important energy contribution due to our strong "in-silico" modifications. By introducing a charged amino group in p/m-position of the ligand pointing into the backpocket of TGFBR1, we also form salt-bridges with either ASP-351 or/and GLU-245 but at the same time we sacrifice the salt bridges between LYS-232 and ASP-351 or/and GLU-245. Therefore, we also have to take the interaction energy between these residues into account to get the total energy gain for the ligand pose. To measure the interaction energy between the sidechains we used again the LIE tool provided with the AmberTools package. Table S2 comprises the results.

Table S2: Results of the additional performed enthalpy calculations. Values in kcal/mol.

Compound	Pose	Charge	LIG-PRO	LYS-GLU,ASP	Total
C	C	neutral	-55.8	-155.9	-211.7
		positive	-102.5	-155.8	-258.3
	D	neutral	-70.2	-153.9	-224.1
		positive	-128.9	-98.7	-227.5
D	C	neutral	-53.5	-157.1	-210.6
		positive	-92.3	-156.6	-248.9
	D	neutral	-80.0	-156.0	-236.0
		positive	-149.5	-69.6	-219.1

If we only look at the interaction between ligand and protein (column LIG - PRO), pose D is always energetically more favored than pose C for compound C. If we also incorporate the interaction between GLU-245, ASP-351 and LYS-232 (column LYS - GLU, ASP) we obtain a different picture. For the positively charged compound C the results including both contributions (column Total) show the energetic gain for compound C in pose C (-258.3 kcal/mol) is more favorable than pose D (-227.5

kcal/mol). This is again in line with our presented results in the main manuscript. For compound D the energy gain of -12.9 kcal/mol from the neutral pose D (-236.0 kcal/mol) to the positive pose C (-248.0 kcal/mol) may not be sufficient to protonate the ligand as compound D is even more difficult to protonate than compound C.

References

1. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I., Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.,* **2000**, 21, 132-146.
2. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.,* **2002**, 23, 1623-41.
3. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.,* **2012**, 8, 3314-3321.
4. Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E. I.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. *AMBER 2015*, University of California, San Francisco, 2015.
5. Onufriev, A.; Bashford, D.; Case, D. A., Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Bioinf.,* **2004**, 55, 383-394.
6. Massova, I.; Kollman, P. A., Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspect. Drug Discovery Des.,* **2000**, 18, 113-135.
7. Aqvist, J.; Marelius, J., The Linear Interaction Energy Method for Predicting Ligand Binding Free Energies. *Comb. Chem. High Throughput Screening,* **2001**, 4, 613-26.
8. Gutierrez-de-Teran, H.; Aqvist, J., Linear Interaction Energy: Method and Applications in Drug Design. *Methods Mol. Biol.,* **2012**, 819, 305-23.