

Supporting Information:

Simulating Protein–Ligand Binding with Neural Network Potentials

Shae-Lynn J. Lahey and Christopher N. Rowley*

Department of Chemistry, Memorial University of Newfoundland, St. John's, NL, Canada

E-mail: crowley@mun.ca

Table S1: Table of crystallographic data. Ligands that did not have a common drug name are listed by their Drugbank ID or their compound ID number(CID).

PDB ID	Protein	EC	Ligand	Resolution (Å)	Temp (K)
1XOZ ^{S1}	phosphodi-esterase 5A	3.1.4.17	tadalafil	1.37	93
3EYG ^{S2}	JAK1	2.7.10.2	tofacitinib	1.9	100
2W6N ^{S3}	biotin carboxylase	6.3.4.14	DB08315	1.87	100
4HJO ^{S4}	EGFR	2.7.10.1	erlotinib	2.21	110
4NCT ^{S5}	Human DYRK1A	2.7.12.1	midostaurin	2.6	100
2HYY ^{S6}	Abl kinase	2.7.10.2	imatinib	2.4	100
3EIG ^{S7}	dihydrofolate reductase	1.5.1.3	methotrexate	1.7	93
3ETA ^{S8}	IGF-1R	2.7.10.1	CID 45272927	2.6	93

Table S2: Table of RMSD of the calculated structures of the ligands relative to the PDB structure.

PDB ID	RMSD (Å)	
	CGenFF	NNP/MM
1XOZ	0.19	0.13
3EYG	0.24	0.50
2W6N	0.54	0.57
4HJO	0.79	0.59
4NCT	1.1	0.60
2HYY	0.39	0.42
3EIG	0.37	0.28
3ETA	0.26	0.35

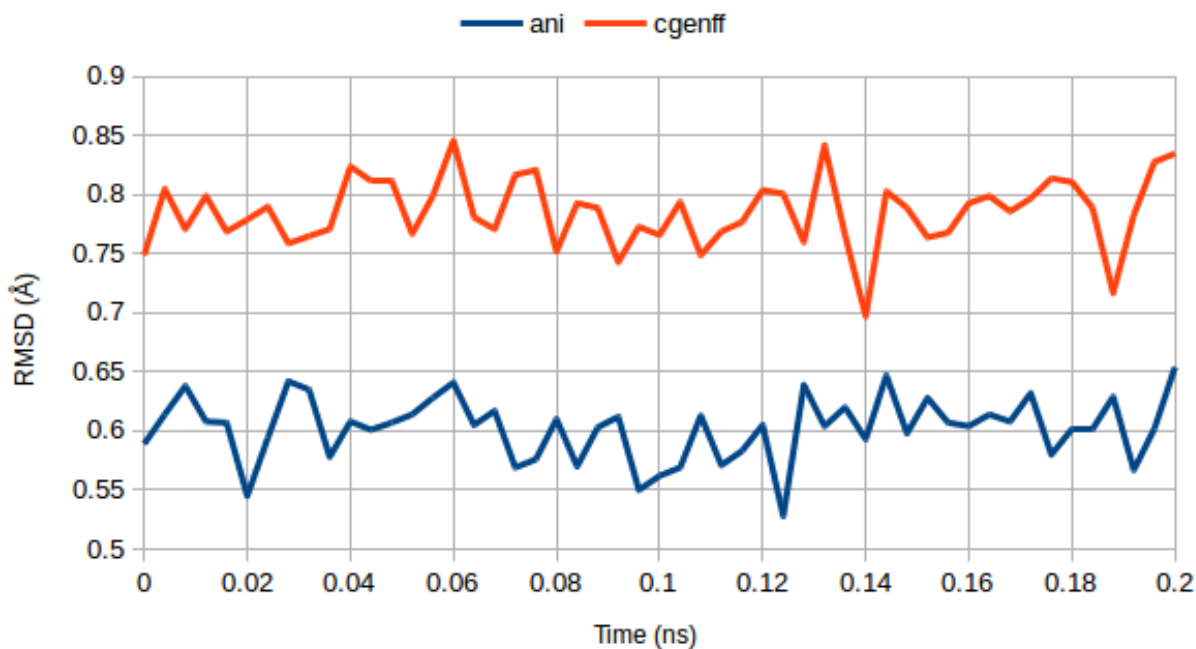


Figure S1: Trajectory of the RMSD of the calculated structure of 4HJO vs the PDB Structure vs time.

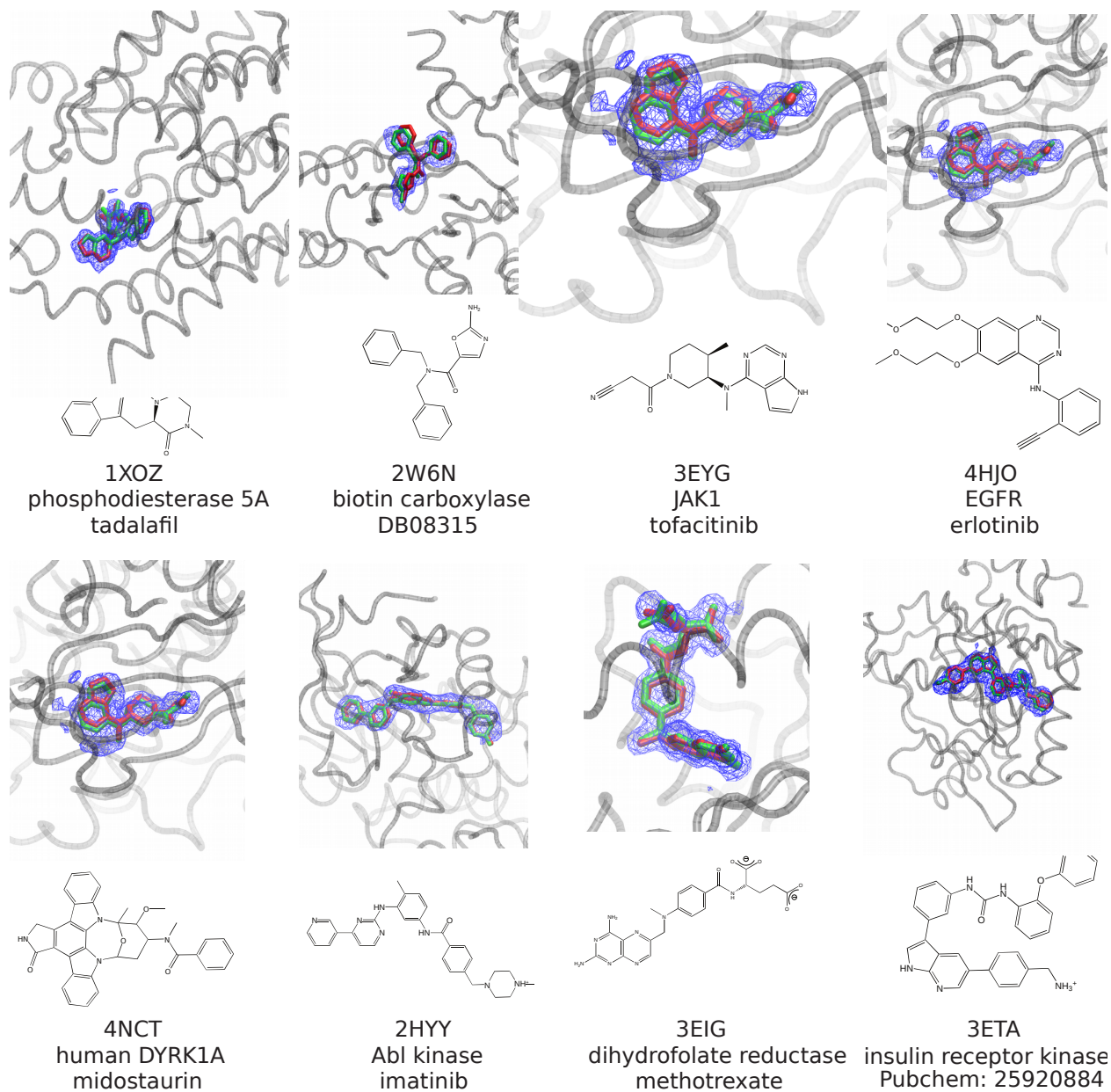


Figure S2: Calculated poses of ligands. CGenFF is in green and ANI is in red. The crystallographic electron density of the ligands are shown in blue. The PDB ID, protein name, and ligand name are included beneath the image.

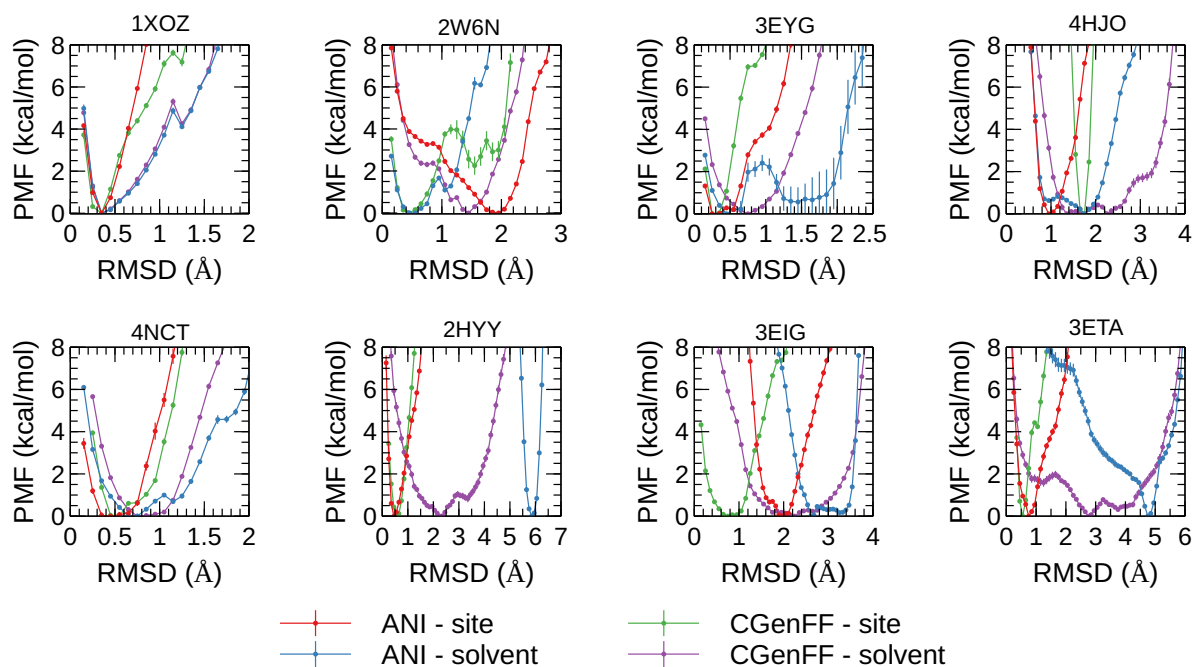


Figure S3: The potential of mean force for the deviation of the structure of a ligand from its bound conformation when it is bound to its protein target.

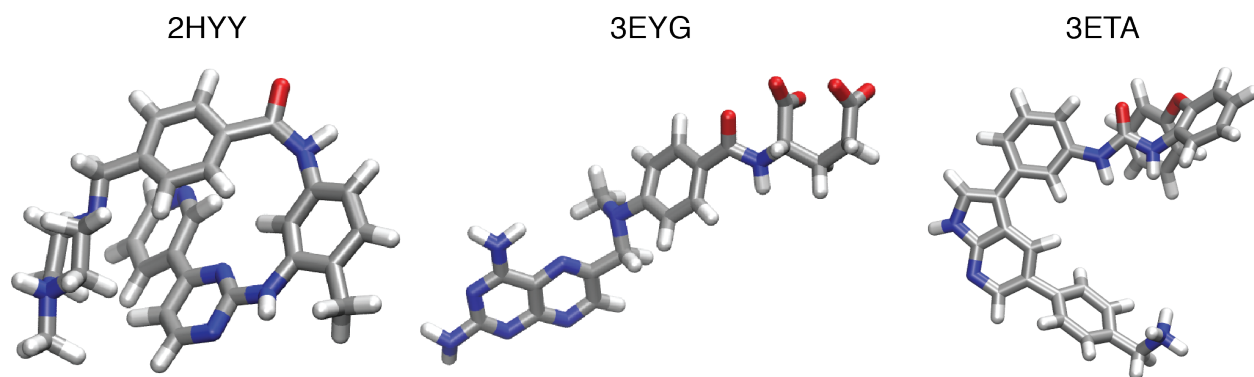


Figure S4: The representative structures of the ionic ligands from the lowest energy of the PMF for the NNP/MM simulations of the ligands in explicit water (the solvent is not shown for clarity). The charged functional groups of the ligands form spurious intramolecular contacts

NNP Technical Details

The ANI-1ccX model was used in all instances where NNP calculations. The TorchANI implementation was used.^{S9} No additional intramolecular force terms for the ligand were included. In this NNP, there are 16 radial elements and the radial cutoff is 5.2 Å. The complete details are described in the Supplementary Materials of Smith et al.[?]

The parameters of the NNP used here are defined in the TorchANI parameter file,^{S10}

Figure S5: Parameters used by the ANI-1ccX NNP (rHCNO-5.2R_16-3.5A_a4-8.params)

```
TM = 1
Rcr = 5.2000e+00
Rca = 3.5000e+00
EtaR = [1.6000000e+01]
ShfR = [9.0000000e-01,1.1687500e+00,1.4375000e+00,1.7062500e+00,
1.9750000e+00,2.2437500e+00,2.5125000e+00,2.7812500e+00,
3.0500000e+00,3.3187500e+00,3.5875000e+00,3.8562500e+00,
4.1250000e+00,4.3937500e+00,4.6625000e+00,4.9312500e+00]
Zeta = [3.2000000e+01]
ShfZ = [1.9634954e-01,5.8904862e-01,9.8174770e-01,1.3744468e+00,
1.7671459e+00,2.1598449e+00,2.5525440e+00,2.9452431e+00]
EtaA = [8.0000000e+00]
ShfA = [9.0000000e-01,1.5500000e+00,2.2000000e+00,2.8500000e+00]
Atyp = [H,C,N,O]
```

NAMD Input File Section for NNP/MM Simulation

```
qmforces on
qmParamPDB qmmm.pdb
qmSoftware custom
qmexecpath client.py
qmBaseDir /dev/shm/
QMColumn occ
qmChargeMode none
qmElecEmbed off
```

Sample ORCA RI-MP2 Input File

```
! RI-MP2 RIJCOSX def2-TZVP def2-TZVP/C def2/J TIGHTSCF Opt PAL8
% maxcore 15000

* xyzfile 0 1 pes.xyz
%geom
Constraints
{D 2 3 4 5 0.0 C }
{D 3 4 5 6 0.0 C }
end
end
```

Sample ORCA DLPNO-CCSD(T) Input File

```
! DLPNO-CCSD(T) def2-TZVP def2-TZVP/C TIGHTSCF
% pal nprocs 8 end
% maxcore 15000

* xyzfile 0 1 pes.xyz
end
end
```

References

- (S1) Card, G. L.; England, B. P.; Suzuki, Y.; Fong, D.; Powell, B.; Lee, B.; Luu, C.; Tabrizizad, M.; Gillette, S.; Ibrahim, P. N.; Artis, D. R.; Bollag, G.; Milburn, M. V.; Kim, S.-H.; Schlessinger, J.; Zhang, K. Y. Structural Basis for the Activity of Drugs that Inhibit Phosphodiesterases. *Structure* **2004**, *12*, 2233–2247.
- (S2) Williams, N. K.; Bamert, R. S.; Patel, O.; Wang, C.; Walden, P. M.; Wilks, A. F.; Fantino, E.; Rossjohn, J.; Lucet, I. S. Dissecting Specificity in the Janus Kinases: The Structures of JAK-Specific Inhibitors Complexed to the JAK1 and JAK2 Protein Tyrosine Kinase Domains. *Journal of Molecular Biology* **2009**, *387*, 219–232.
- (S3) Mochalkin, I.; Miller, J. R.; Narasimhan, L.; Thanabal, V.; Erdman, P.; Cox, P. B.; Prasad, J. V. N. V.; Lightle, S.; Huband, M. D.; Stover, C. K. Discovery of An-

- tibacterial Biotin Carboxylase Inhibitors by Virtual Screening and Fragment-Based Approaches. *ACS Chemical Biology* **2009**, *4*, 473–483.
- (S4) Park, J. H.; Liu, Y.; Lemmon, M. A.; Radhakrishnan, R. Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. *Biochemical Journal* **2012**, *448*, 417–423.
- (S5) Alexeeva, M.; Åberg, E.; Engh, R. A.; Rothweiler, U. The structure of a dual-specificity tyrosine phosphorylation-regulated kinase 1A–PKC412 complex reveals disulfide-bridge formation with the anomalous catalytic loop HRD(HCD) cysteine. *Acta Crystallographica Section D* **2015**, *71*, 1207–1215.
- (S6) Cowan-Jacob, S. W.; Fendrich, G.; Floersheimer, A.; Furet, P.; Liebetanz, J.; Rummel, G.; Rheinberger, P.; Centeleghe, M.; Fabbro, D.; Manley, P. W. Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. *Acta Crystallographica Section D Biological Crystallography* **2006**, *63*, 80–93.
- (S7) Volpato, J. P.; Yachnin, B. J.; Blanchet, J.; Guerrero, V.; Poulin, L.; Fossati, E.; Berghuis, A. M.; Pelletier, J. N. Multiple Conformers in Active Site of Human Dihydrofolate Reductase F31R/Q35E Double Mutant Suggest Structural Basis for Methotrexate Resistance. *Journal of Biological Chemistry* **2009**, *284*, 20079–20089.
- (S8) Patnaik, S.; Stevens, K. L.; Gerding, R.; Deanda, F.; Shotwell, J. B.; Tang, J.; Hamajima, T.; Nakamura, H.; Leesnitzer, M. A.; Hassell, A. M.; Shewchuck, L. M.; Kumar, R.; Lei, H.; Chamberlain, S. D. Discovery of 3, 5-disubstituted-1H-pyrrolo[2, 3-b]pyridines as potent inhibitors of the insulin-like growth factor-1 receptor (IGF-1R) tyrosine kinase. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 3136–3140.
- (S9) Roitberg Group, TorchANI. 2019; <https://github.com/aiqm/torchani>.
- (S10) Roitberg Group, ANI-1ccX Parameters. 2019; https://github.com/isayev/ASE_ANI/blob/master/ani_models/ani-1ccx_8x/rHCN0-5.2R_16-3.5A_a4-8.params.