Benchmarking Methods For Free Energy Estimation

Jocelyn Sunseri and David Ryan Koes

Carnegie Mellon – University of Pittsburgh Joint Program in Computational Biology, University of Pittsburgh

Abstract

Estimating binding free energy is crucial for judging the favorability of a binding reaction. This is an important step in a drug discovery pipeline, where an initial query may return many hits that must be subsequently filtered and ranked, and the initial search may have used a scoring function as a proxy for the binding free energy. A number of free energy estimation methods use molecular dynamics simulations of the protein-ligand complex. The simulation is then post-processed to provide an estimate. The quality of sampling the simulation provides is highly dependent on the force field parametrization, particularly the parameters estimated for the ligand. We report our first results in a comparison of the accuracy of these methods and how that accuracy varies when using different methods for partial charge parametrization.

Methods

Amber14 was used to set up and run molecular dynamics simulations with the **ff14sb** force field and **tip3p** water model. Every complex was simulated for **100ns**. A subset of the DUD-E database with binding affinity information in the PDB was used for testing. A total of **49 complexes** were successfully built, simulated, and evaluated by all methods.

Five methods built into AmberTools were evaluated:

MM/GBSA (1traj) – Binding free energy is estimated as the average free energy calculated from trajectory frames of the complex, minus the average free energy of solitary protein and ligand conformations stripped from the trajectory of the complex. Generalized Born and surface area solvation is used.

MM/GBSA (3traj) – Similar to above, except separate simulations are run for the solitary ligand and protein to obtain configurations independent of the trajectory of the complex.

MM/PBSA (1traj) –Poisson-Boltzmann and surface area solvation is used instead of GBSA.

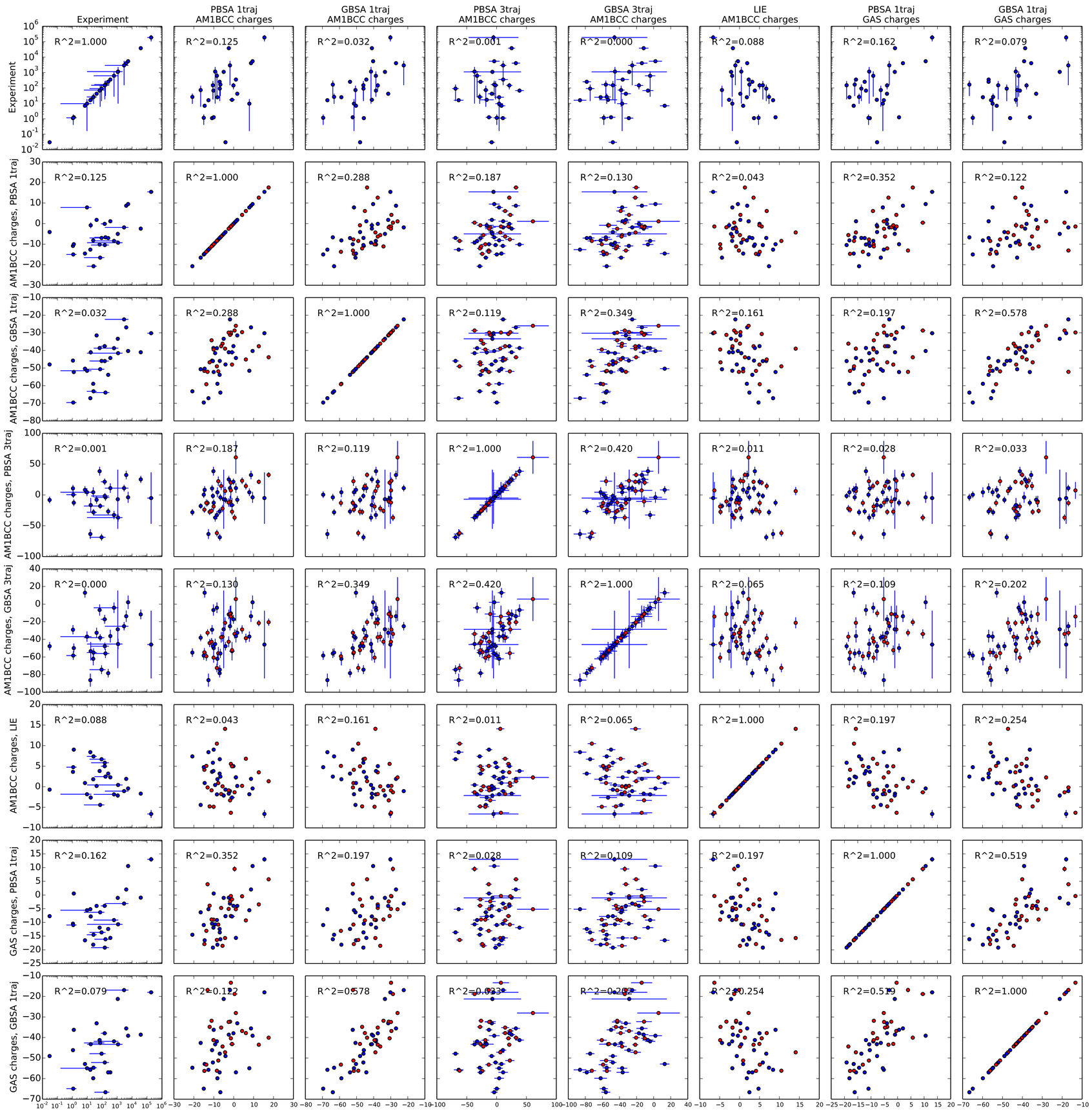
MM/PBSA (3traj) – Similar to above, except separate simulations are run for the ligand and protein.

Linear Interaction Energy – LIE estimates average Van der Waals and electrostatic energies over both complex and ligand simulations and computes the binding free energy as a weighted sum of the differences between these terms.

Future Work

•Use Paramfit to Gaussian calculated potentials to obtain parameters, including RESP charges and bond terms.

•Include other methods for free energy estimation, including thermodynamic integration and the weighted histogram analysis method.



This work was supported by the National Institute of Health [R01GM108340]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.