Free energy simulations with quantum/deep-learning potentials for drug discovery and enzyme design

Darrin York

Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ

Free energy simulations have been used extensively in applications to drug discovery and enzyme design. In the case of drug discovery, so-called alchemical free energy simulations are used to make predictions about ligand-protein binding affinities that are responsible for potency and selectivity of drugs. In the case of enzyme design, free energy simulations are used to probe the landscape of chemical reactions in order to predict mechanistic pathways and identify factors that regulate activity. The predictive capability of these methods relies critically on the accuracy of the force fields and the precision of the calculations themselves. For complex biological systems, the latter requires extensive sampling of macromolecular conformations and solvent configurations that can be prohibitively expensive using *ab intio* quantum mechanical (QM) methods even within a combined quantum mechanical/molecular mechanical (QM/MM) framework. Nonetheless, fast semiempirical quantum methods traditionally lack the high accuracy demanded by these applications. Herein we discuss strategies whereby new semiempirical (density-functional tight binding) quantum mechanical methods with machine learning potential corrections (QM/Δ-MLPs) afford a robust, accurate and efficient alternative. These models can be used to extensively sample the important degrees of freedom required to calculate accurate free energy surfaces and values, and make reliable predictions about catalytic mechanisms and binding free energies. The QM/Δ-MLPs are integrated with new enhanced sampling, path finding and free energy analysis methods, including a new generalized weighted thermodynamic perturbation methods that ultimately corrects, if necessary, the QM/Δ-MLPs to the target QM level. The methods are illustrated in binding free energy simulations of designed drugs to protein targets, and catalytic mechanisms of nucleic acid enzymes and including prediction of kinetic isotope effects from (ring polymer) path integral molecular dynamics (PIMD). Insight gained from these studies reveal important design principles for new nucleic acid enzymes used in biotechnology applications and as therapeutics.