Automated control in computational chemical discovery

Fang Liu,[1] Chenru Duan,[1,2] Heather J. Kulik[1]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

[2]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139*

There are several prominent challenges in conducting quantum chemistry studies of transition metal complexes: generation of molecule structure, selection of electronic structure method, workflow control, and post-processing of simulation results. molSimplify is a user-friendly open-source toolkit that enables seamless generation of candidate inorganic molecule structures, preparation and execution of Density Functional Theory (DFT) calculations, and post-processing DFT results. Here we focus on extending the capability of molSimplify, enabling automated selection of electronic structure methods including both DFT and wavefunction based methods and automated control of simulations. For automatic method selection, we develop a python module, MultirefPredict, which is a high-level cross-platform workflow that calculates widely used multi-reference diagnostics for a given molecular system, without users handling the input and output of quantum chemistry packages. The backend of MutlrefPredict supports several quantum chemistry packages based on both DFT and wavefunction theory methods, with calculation performed on both CPU and GPU. MultirefPredict guides a user to choose between single-reference and multi-reference based methods, which is essential for obtaining accurate results for transition metal containing systems. For *in-situ* simulation control, we developed an automated workflow that checks the current status of geometry optimization using a "dynamic classifier" that predicts the probability of failure based on geometric and electronic structure information collected on the fly. This workflow terminates the simulation when it is predicted to fail, saving numerous computational resources. These open-source tools, which are accessible to the whole transition metal chemistry community, are anticipated to make transition metal chemistry simulation more automated and transparent.