Open-source GPU-enabled *ab initio* QM/MM with QUICK/AMBER including long-range electrostatics

Madushanka Manathunga, Timothy J Giese, Vinicus Wilian D Cruzeiro, Yipu Miao, Dawei Mu, Kristopher Keipert, Kamesh Arumugam, Darrin M York, Andreas W. Goetz and Kenneth M. Merz, Jr.

The AMBER software for biomolecular simulations has a long history of QM/MM implementations that couple the strength of quantum chemistry with the speed of molecular mechanics and thus enable simulations of local molecular properties and reaction mechanisms in complex environments such as enzymes and other condensed phase systems. Supported Hamiltonians range from built-in efficient but approximate semiempirical and density functional tight-binding methods to accurate but computationally more expensive *ab initio* methods via robust interfaces to external quantum chemistry software packages. A major drawback of the latter is that electrostatic interactions between the QM and MM regions are truncated beyond a cutoff.

Here we present recent developments that enable both accurate and efficient *ab initio* Hartree-Fock and density functional theory-based QM/MM simulations that properly couple the QM and MM regions in condensed phase under periodic boundary conditions without truncation of electrostatics. Because Ewald approaches with simple charge mapping schemes are not numerically stable with extended basis sets, we incorporate long-range electrostatic interactions via the ambient-potential composite Ewald method. The QM charge density thus interacts directly with the MM environment, but in a manner that avoids the use of dense Fourier transform grids. We demonstrate the accuracy of the method with a range of applications including reaction free energy profiles in the condensed phase. We describe this new method via an integration of AMBER with the open-source, multi-GPU enabled Gaussian basis-set code QUICK via an application programming interface that can also be used with other software. In QUICK all numerically intensive parts of a QM/MM calculation have been ported to GPUs, leading to a performance on a single data center GPU that exceeds the performance of the corresponding MPI parallel CPU code on a state-of-the-art multi-core CPU compute node. This opens the door to large-scale QM/MM simulations that require both high accuracy and extended sampling, such as reaction free energy profiles in the condensed phase.