An $n \log n$ approximation based on the natural organization of biomolecules for speeding up the computation of long range interactions

Ramu Anandakrishnan and Alexey Onufriev
Virginia Tech, Blacksburg, VA

Abstract

Estimating potential, energy or forces due to long range interactions in biomolecular systems is generally computationally demanding. Presented here is a method, the multilevel charge approximation (MLCA), for speeding up long range electrostatic computations in biomolecular systems. The approximation is based on multiple levels of natural partitioning of biomolecular structures into a hierarchical set of its constituent structural components, e.g. amino acids and nucleotides (groups), protein and DNA chains (subunits), and molecular complexes. The charge distribution for each component in this hierarchical set is approximated by a small number of point charges, which, for the highest level component, are much fewer than the number of atoms in the component. For short distance interactions, the electrostatic potential is calculated exactly using the full set of atomic charges. For long distance interactions, the smaller set of approximate charges are used instead. For a structure consisting of n atoms, the computational cost of the MLCA can scale as $O(n \log n)$ compared to $O(n^2)$ for an all-atom computation, depending on the specific hierarchical organization of the structure. The MLCA is tested on a representative set of 600 biomolecular structures ranging in size from approximately 300 atoms to 3 million atoms. The accuracy is estimated relative to the exact all-atom calculation and is compared to that of the commonly used spherical cutoff and particle mesh Ewald (PME) methods. For the set of biomolecular structures tested, the MLCA exhibits relative force error of 0.024, which, on average, approaches the accuracy of the more complex, industry standard, PME method using typical parameters. The MLCA is also significantly more accurate than the spherical cutoff method which has a relative force error of 0.089. We also show that the MLCA can be multiple orders of magnitude faster than the exact all-atom computation for large structures. A critical benefit of this approximation, compared to PME, is that it may be used with implicit solvent models to further reduce computational costs.