Minimal-Parameter Implicit Solvent Model for Large-Scale DFT Calculations

J. Dziedzic[1,2,3], A. Bhandari[1,3], C.-K. Skylaris[1,3]

[1] *School of Chemistry, University of Southampton, Southampton, UK*

[2] *Faculty of Applied Physics and Mathematics, Gdansk University of Technology, Poland*

[3] *The Faraday Institution, UK*

Accurate studies of important biochemical processes, such as protein folding or protein-ligand binding, necessitate the inclusion of the solvent environment. Including the solvent explicitly with full atomistic detail is usually computationally unfeasible, not only because it substantially increases the number of atoms in the system, but also because it requires extensive averaging over the solvent degrees of freedom. The implicit solvent (IS) approach addresses this difficulty by representing the solvent as an unstructured dielectric continuum. This simplified description is often sufficient for capturing the relevant phenomena, yielding solvation energies in good agreement with experiment, making IS the prime candidate for use within density functional theory (DFT) calculations on biological systems.

A multitude of IS models of differing sophistication have been proposed to date. One of the approaches for determining the electrostatic component of solvation involves solving the Poisson-Boltzmann equation (PBE) for a system where the solute is placed in a suitably defined cavity with the dielectric kept outside of the cavity. The exact shape of the cavity usually depends on a number of parameters, such as atomic radii. Alternatively, the cavity can be defined in terms of the electronic density, which eliminates the need to carefully tune the many cavity parameters present in most IS models.

We present an implementation of an PBE-based solvation model in the ONETEP linear-scaling density functional theory program. ONETEP achieves CPU and memory requirements that increase linearly with the number of atoms by employing a density matrix formulation of DFT using localized functions (Non-orthogonal Generalized Wannier Functions or NGWFs). Our IS model supports open and periodic boundary conditions, charged and neutral solutes, it approximately includes dispersion-repulsion interactions, and uses novel approaches to describe the Boltzmann electrolyte. It is a fully self-consistent approach, where the electronic density is optimized in the presence of the implicit solvent. Solutions to the PBE are obtained using a high-order defect-corrected multigrid method, which we demonstrate is efficient and accurate. We show how the proposed approach consistently gives predictions in excellent agreement with experiment on a test set of 100+ molecules, outperforming competing IS models, and demonstrate how it can be used to obtain binding energies in solvent for entire proteins with thousands of atoms, while maintaining good parallel scalability.