Benchmarking Polarizable Continuum Models For Macromolecular Analysis

Benjamin J. Albrecht and John M. Herbert

(Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio)

THE OHIO STATE UNIVERSITY

Abstract

In quantum chemistry, molecular characteristics, such as energy, vibrational frequency, and geometry, are predicted and modeled using computational chemistry software. In computing these characteristics, calculations can become cumbersome for increasing orders of accuracy or system size. Thus affordability of calculation times has become the largest constraint in quantum chemistry. With no all-encompassing optimal computational method for calculating molecular characteristics, methods have to be chosen depending on the molecular property of interest, the environment of the species, and the desired accuracy of the result. Here, the ability of various state of the art implicit solvent models, known as polarizable continuum models (PCMs), are benchmarked in their ability to accurately and efficiently compute solvation energies. With the collected benchmark data, it can be determined what empirical or physically motivated corrections can be implemented to effectively reduce solvation energy errors in PCMs, potentially expanding the accuracy and timescale efficiency in computing solvation energies in macromolecular systems.

Motivation

In both experimental and theoretical studies of chemistry, focus tends to lie on gas phase properties due to a wide variety of challenges faced in liquid phase properties. One of these factors is the immense scaling of computational cost in *ab initio* quantum chemistry methods with respect to both accuracy and number of atoms. This poses a large problem for computing molecules in solution because the number of atoms in the entire system can become expensive easily. PCMs, being a category of implicit solvent models, alleviate this burden and give rise to greater capabilities in computing molecular properties of large systems in solution, such as biological macromolecules. One large challenge in PCMs has been the inability to afford smooth potential energy surfaces. However, this challenge has, in recent years, been partially overcome with the introduction of the Switching/Gaussian (SWIG) approach. It is now of great interest to benchmark the ability of various modern PCMs with the newly implemented SWIG approach and reevaluate their practicality and effectiveness as implicit solvent models.

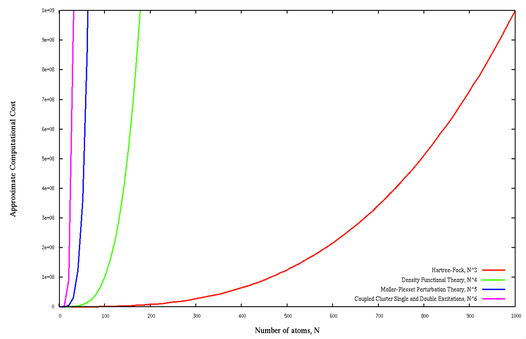

Fig 1 – Scaling of computational cost with respect to system size in number of atoms for various *ab initio* quantum chemistry methods. Cost of more accurate methods become prohibitive at a much higher rate.

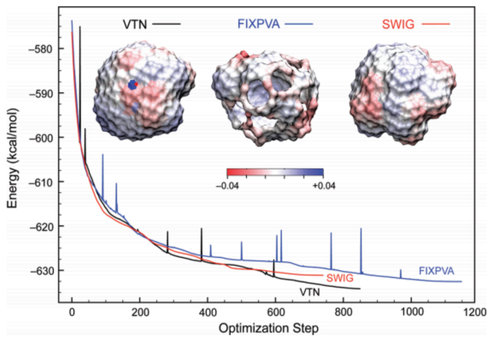

Fig. 2 – Geometry optimization of (adenine)$(H_2O)_{52}$ in bulk water, using three different implementations of CPCM: Variable Tesserae Number (VTN), Fixed-point Variable Area (FIXPVA), and Switching/Gaussian (SWIG) Methods. The vertical scale represents binding energy. Also shown are the cavity surfaces for the optimized structures. (A.W. Lange and J.M. Herbert, 2010)

Polarizable Continuum Models

Dielectric Continuum Solvation Model

PCMs are effective at reducing computational cost in isotropic bulk-solvent solutions at equilibrium because they implicitly treat the solvent molecules by representing the solvent as a dielectric continuum, as opposed to treating them explicitly. By constructing a cavity around the solute and imposing the condition that the charge density of the solute is effectively screened by the dielectric continuum at the cavity surface, the electrostatic interactions can be solved explicitly using Poisson's Equation:

$$-\vec{\nabla}[\epsilon(\vec{r})\vec{V}(\vec{r})] = 4\pi\rho_M(\vec{r})$$

Where ρ, the charge density of the solute, is obtained either quantum mechanically (QM) or molecular mechanically (MM). By imposing these conditions, the problem is effectively mapped from a 3-dimensional space onto a 2-dimensional surface, reducing the complexity and computational cost of the problem.

Cavity Construction

In reality, the cavities defined in PCMs do not exist. Therefore, various unique schemes of defining the cavity have emerged:

Solvent Excluded Surface (SES)
-pictured to the right
Solvent Accessible Surface (SAS)
-pictured to the right
Isodensity Contour
-locus of points with electron density, alpha
Marching Cubes
-similar to isodensity, but marching outward
Van Der Waals Radii (vdW)
-scaled Bondi radii

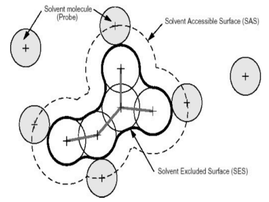

Fig 3. – Solvent accessible surface (SAS), also known as a Connolly surface, traced by the center of the probe. The solvent accessible surface (SES) is the union of all possible probes that do not overlap with the molecule. (Tomasi, et. al., 2005)

Computing Solvation Energies

Methods were benchmarked by their ability to compute solvation energies, represented as W below. The calculations were performed using Q-Chem. Solvation energies were benchmarked against the experimental Minnesota Solvation Database (Truhlar et. al., 2009).

$$W = \Delta G = G_{gas} - G_{liq}$$
$$G = G_{elec} + G_{nonelec}$$
$$G_{nonelec} = G_{cav} + G_{disp} + G_{rep}$$

$$\Delta W = |W_{expt} - W_{calc}|$$

Results

Fig. 4 – The many flavors of PCMs. Note the asterisks represent data published in literature. All data presented was computed with Q-Chem's PCM implementations, using Hartree-Fock/6-31+G* correlation function and theory, van der Waals cavity construction, for N=2413 neutral solutes in solution. The reported values from the literature are using an isodensity contour cavity construction and N=264 neutrals. (Pomogaeva et. al., 2011)

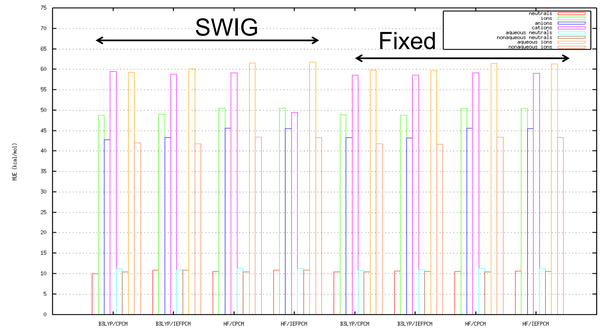

Fig. 6 – Error statistics for variations of correlation functions and PCM with SWIG on and off (fixed). These calculations were performed and averaged over a data set of 383 aqueous solutes with vdW cavities.

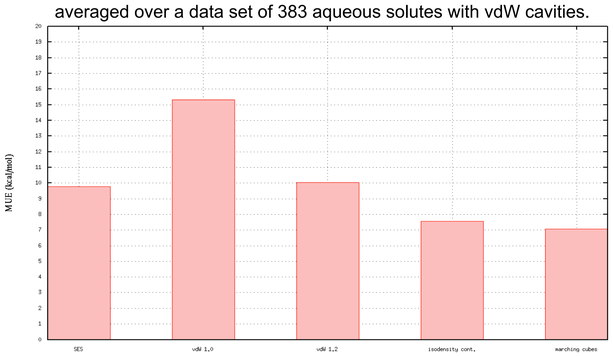

Fig. 6 – Error Statistics for cavity schemes. All calculations were performed using C-PCM / 6-31+G* / B3LYP with 110 Lebedev grid points. The data was averaged for 386 aqueous neutrals, except for the isodensity contour, which was only able to treat 350 of the solutes satisfactorily.

Conclusions & Future Works

-The SWIG approach does not appear to affect the non-electrostatics in any substantial way

-The ion solvation energies are currently flawed
-Short range electrostatics corrections may be necessary.

-The error statistics of neutral solutes are not yet comparable with other leading solvation models, but do have a smooth PES, minimal parameterization, and converge continuously.

Acknowledgements

Special thanks to project advisor, John M. Herbert.

All computations presented were performed on the Ohio Supercomputer Center.

This work was supported by the Undergraduate Research Scholarship through the College of Arts and Sciences.