

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications -- Chemistry Department

Published Research - Department of Chemistry

2009

Heterogeneous conductorlike solvation model

Hui Li

University of Nebraska - Lincoln, hli4@unl.edu

Dejun Si

University of Nebraska - Lincoln

Follow this and additional works at: <https://digitalcommons.unl.edu/chemfacpub>

 Part of the [Chemistry Commons](#)

Li, Hui and Si, Dejun, "Heterogeneous conductorlike solvation model" (2009). *Faculty Publications -- Chemistry Department*. 29.

<https://digitalcommons.unl.edu/chemfacpub/29>

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications -- Chemistry Department by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Heterogeneous conductorlike solvation model

Dejun Si and Hui Li^{a)}

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA

(Received 15 May 2009; accepted 7 July 2009; published online 27 July 2009)

A heterogeneous conductorlike solvation model (conductorlike screening model/conductorlike polarizable continuum model) that uses different local effective dielectrics for different portions of the solute cavity surface is implemented for quantum chemical Hartree–Fock and Kohn–Sham methods. A variational treatment is used to form the heterogeneous solvation operator, so a simple analytic expression of the energy gradients, which are vital for geometry optimization and molecular dynamics simulation, is derived and implemented. Using the new Fixed Points with Variable Areas surface tessellation scheme, continuous and smooth potential energy surfaces as well as analytic gradients are obtained for this heterogeneous model. Application of the heterogeneous solvation model to a realistic quantum model consisting of 101 atoms for the type-1 Cu center in rusticyanin shows that the desolvation due to protein burial can likely raise the reduction potential by ~ 200 mV and, including the heterogeneity in geometry optimization, can likely affect the results by ~ 2 kcal/mol or ~ 70 mV. © 2009 American Institute of Physics. [DOI: 10.1063/1.3187527]

I. INTRODUCTION

Continuum solvation models have been widely used to study molecules in solutions.¹ In such models, the solvent is represented by a polarizable medium characterized by a dielectric constant, and the electrostatic and polarization interactions between the solute molecule and the solvent molecules are determined by numerically solving the classic electrostatic Poisson equations. The polarizable continuum models [the earlier DPCM (Ref. 2) and the recent integral equation formalism PCM, or IEF-PCM^{3,4}], the conductorlike screening models [COSMO (Ref. 5) and GCOSMO (Ref. 6) or conductorlike PCM (CPCM^{7,8})], the surface and simulation of volume polarization for electrostatics (Ref. 9) models, as well as the SMx models,¹⁰ are popular continuum solvation models. In DPCM, IEF-PCM, and CPCM, the cavitation, exchange-repulsion, and dispersion interactions follow different functional forms and are treated separately.

Although in many cases the solvent can be treated as a homogeneous and isotropic polarizable medium with a dielectric constant, there are cases in which the heterogeneity or anisotropy of the environment surrounding a molecule must be considered. Typical examples include solute molecules at the interface between two phases, proteins, and other molecules embedded in lipid bilayers, solvated guest-host complexes, and protein active sites or cofactors in protein matrices.

Using DPCM, Tomasi's group performed pioneering studies on various heterogeneous solvation problems, for example, the energy changes in deformations of long DNA fragments and that in the opening of a DNA double helix,¹¹ the partial solvation effect in molecular recognition and docking,¹² and polar solutes placed near the surface of two immiscible liquids or at a liquid/vacuum separation.¹³ Hoshi

*et al.*¹⁴ extended the DPCM to treat anisotropic polarization effects in their guest-host complex calculations.

Cances *et al.*^{3,4} developed a general IEF-PCM for treating anisotropic solvation problems. Frediani *et al.*¹⁵ later extended IEF-PCM to study molecules at diffuse interfaces between two fluid phases by introducing position dependent permittivities. By using different local effective dielectrics for different portions of the solute cavity surface, Iozzi *et al.*¹⁶ further extended IEF-PCM to study the pK_a of a solvent exposed histidine residue in prion protein and a small molecule interacting with a biological membrane. A detailed discussion on heterogeneous solvation models can be found in a recent review by Tomasi *et al.*¹

To the best of the authors' knowledge, no analytic nuclear gradients for heterogeneous continuum solvation models have been derived and implemented, and no geometry optimizations have been reported. In this work, a heterogeneous conductorlike solvation model (COSMO and CPCM, high dielectric versions of the more general IEF-PCM^{3,4}) is implemented. Following the heterogeneous IEF-PCM developed by Iozzi *et al.*,¹⁶ different local effective dielectrics are used for different portions of a solute cavity surface.

By variationally formulating the solvation operators for Hartree–Fock and Kohn–Sham methods, using the Fixed Points with Variable Areas (FIXPVA) tessellation scheme developed by Su and Li,¹⁷ and assuming that the local effective dielectrics are constants in a geometry optimization process, continuous and smooth potential energy surfaces and analytic nuclear gradients can be obtained for the heterogeneous solvation model.

Application of the heterogeneous solvation model to a realistic quantum model consisting of 101 atoms for the type-1 Cu center in rusticyanin shows that the desolvation due to protein burial is likely to raise the reduction potential

^{a)}Author to whom correspondence should be addressed. Electronic mail: hli4@unl.edu.

by ~ 200 mV and, including the heterogeneity in geometry optimization, is likely to affect the results by ~ 2 kcal/mol or ~ 70 mV.

II. THEORY

A. Heterogeneous COSMO and CPCM

The COSMO was originally developed by Klamt and Schüürmann.⁵ The CPCM, similar to COSMO, was implemented in the frame of the more general IEF-PCM method.^{7,8} In the following, COSMO and CPCM are discussed together.

Although previous studies^{7,18} demonstrated that acceptable results could be obtained with COSMO and CPCM for rather low dielectric solvents, it is important to note that both COSMO and CPCM are approximations of IEF-PCM, which is more rigorous from an electrostatic interaction point of view and can be used with equal accuracy for both high and low dielectric solvents.^{3,4} In this work, only COSMO and CPCM are discussed because their simplicity allows for an easier treatment in the development of a heterogeneous model.

In addition, the readers must keep in mind that the rigorous variational treatment presented in this work for heterogeneous COSMO and CPCM solely means that the mathematical structure of these methods in combination with Hartree–Fock and Kohn–Sham methods is analyzed exactly, rather than that the models are electrostatically rigorous.

COSMO and CPCM describe the solvent reaction potential with induced surface charges distributed on the solute cavity surface. By using boundary element method, the continuous distribution of the induced surface charge is represented by a set of induced point charges located at the surface tesserae. These point charges, written as a vector \mathbf{q} , satisfy the following matrix equation:

$$\mathbf{C}\mathbf{q} = -k\mathbf{V}, \quad (1)$$

where the vector \mathbf{V} collects the electrostatic potentials created by the solute at the surface tesserae. The elements of symmetric matrix \mathbf{C} are

$$C_{ii} = 1.07 \sqrt{\frac{4\pi}{a_i}}, \quad (2)$$

$$C_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (3)$$

with a_i being the area and \mathbf{r}_i being the center coordinates of tessera i .

For homogeneous solvents, k in Eq. (1) is a pure number,

$$k = (\varepsilon - 1)/\varepsilon, \quad (4)$$

with ε being the dielectric constant of the solvent. Other values for k have been used in the literature. For example, Klamt and Schüürmann originally suggested $k = (\varepsilon - 1)/(\varepsilon + 0.5)$.⁵

In this work, the heterogeneity of the environment surrounding the solute is modeled by using a local effective dielectric ε_i for each surface tessera, similar to that in the heterogeneous IEF-PCM developed by Iozzi *et al.*¹⁶ Math-

ematically, this corresponds to replacing the pure number k in Eq. (1) with a diagonal matrix \mathbf{K} with the following elements:

$$\mathbf{K}_{ii} = (\varepsilon_i - 1)/\varepsilon_i, \quad (5)$$

$$\mathbf{K}_{ij} = 0. \quad (6)$$

The local effective dielectrics for the tesserae must be determined according to the specific problems and can be difficult in some cases. For example, it is useful to define the local dielectric for a tessera as a function of its position in the study of biomolecules embedded in lipid membranes. In the current work, local dielectrics are assigned to surface tesserae according to the spheres (i.e., atoms) they belong to; the tesserae on the same sphere have the same local effective dielectric given by the user from the input deck.

B. Hartree–Fock and Kohn–Sham methods

In Hartree–Fock and Kohn–Sham calculations, the heterogeneous COSMO and CPCM induced surface charges can be determined separately for solute nuclei and electrons,

$$\mathbf{q}_N = -\mathbf{C}^{-1}\mathbf{K}\mathbf{V}_N, \quad (7)$$

$$\mathbf{q}_e = -\mathbf{C}^{-1}\mathbf{K}\mathbf{V}_e, \quad (8)$$

where \mathbf{V}_N and \mathbf{V}_e are the nuclear and electronic potential, respectively, at the tesserae. Later it will be shown that actually a more efficient “total charge” scheme can be adapted.

Equation (8) can be equivalently written as the basis set induced surface charges $\mathbf{q}_{\rho\sigma}$ contracted by the density matrix

$$\mathbf{q}_e(i) = \sum_{\rho\sigma} P_{\rho\sigma} \mathbf{q}_{\rho\sigma}(i), \quad (9)$$

where ρ and σ , as well as μ and ν used later, are Gaussian type basis functions; $P_{\rho\sigma}$ is the density matrix. $\mathbf{q}_{\rho\sigma}$ is obtained by solving Eq. (10) for the basis set potentials $\mathbf{V}_{\rho\sigma}$ at the tesserae,

$$\mathbf{q}_{\rho\sigma} = -\mathbf{C}^{-1}\mathbf{K}\mathbf{V}_{\rho\sigma}, \quad (10)$$

with

$$\mathbf{V}_{\rho\sigma}(i) = -\langle \rho | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \sigma \rangle. \quad (11)$$

where \mathbf{r}_e is the electronic coordinate and \mathbf{r}_i is the tessera coordinate.

The electrostatic potentials generated by the induced surface charges are then incorporated into the Hartree–Fock or Kohn–Sham equation to variationally determine the total molecular energy E_{total} over a finite basis set,

$$\begin{aligned}
E_{\text{total}} = & \sum_{\mu\nu} P_{\mu\nu}(T_{\mu\nu} + Z_{\mu\nu}) + \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}P_{\rho\sigma}(\langle\boldsymbol{\mu}\boldsymbol{\nu}|\boldsymbol{\rho}\boldsymbol{\sigma}\rangle) \\
& + \sum_{\mu\nu} P_{\mu\nu}\langle\boldsymbol{\mu}|V_{\text{xc}}(\mathbf{r}_e)|\boldsymbol{\nu}\rangle + E_{\text{ZZ}} \\
& - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \\
& - \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}P_{\rho\sigma} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle + \frac{1}{2} \sum_{\alpha} \sum_i \frac{Z_{\alpha}\mathbf{q}_N(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|} \\
& + \frac{1}{2} \sum_{\rho\sigma} P_{\rho\sigma} \sum_{\alpha} \sum_i \frac{Z_{\alpha}\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|}, \quad (12)
\end{aligned}$$

where $T_{\mu\nu}$ and $Z_{\mu\nu}$ are the basis set kinetic energy and electron-nuclei potential energy integrals, respectively; $V_{\text{xc}}(\mathbf{r}_e)$ is the Hartree-Fock exchange or the density functional theory (DFT) exchange-correlation potential; E_{ZZ} is the nuclear repulsion energy; and Z_{α} and \mathbf{r}_{α} are the nuclear charge and coordinates of atom α .

The first four terms in Eq. (12) have the same forms as those in the gas phase Hartree-Fock and Kohn-Sham methods but evaluated with the solvent perturbed density matrix. They can be denoted as E_{gas} . The last four terms in Eq. (12) are due to induced surface charges and can be denoted as E_{sol} ,

$$\begin{aligned}
E_{\text{total}} = & E_{\text{gas}} + \frac{1}{2}\mathbf{V}_e^T\mathbf{q}_N + \frac{1}{2}\mathbf{V}_e^T\mathbf{q}_e + \frac{1}{2}\mathbf{V}_N^T\mathbf{q}_N + \frac{1}{2}\mathbf{V}_N^T\mathbf{q}_e \\
= & E_{\text{gas}} + \frac{1}{2}\mathbf{V}^T\mathbf{q} \\
= & E_{\text{gas}} + E_{\text{sol}}. \quad (13)
\end{aligned}$$

The solvation operator must be determined based on a variational treatment of Eq. (12). The fourth and seventh terms in Eq. (12) do not contain electronic coordinates and thus are irrelevant. Variational treatments of the first three terms in Eq. (12) lead to the gas phase Fock or Kohn-Sham operators. A variational treatment of the fifth term in Eq. (12) leads to an operator representing the potential due to nucleus induced surface charge,

$$\frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \Rightarrow \frac{1}{2} \sum_i \frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|}. \quad (14)$$

A variational treatment of the sixth term in Eq. (12) with respect to the density matrix gives

$$\begin{aligned}
& \delta \left(-\frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}P_{\rho\sigma} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \right) \\
= & -\frac{1}{2} \sum_{\mu\nu\rho\sigma} \delta P_{\mu\nu}P_{\rho\sigma} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \\
& - \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}\delta P_{\rho\sigma} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \\
= & -\frac{1}{2} \sum_{\mu\nu} \delta P_{\mu\nu} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_e(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \\
& + \frac{1}{2} \sum_{\rho\sigma} \delta P_{\rho\sigma} \sum_i \sum_j \mathbf{V}_e(i) \mathbf{C}_{ij}^{-1} \mathbf{K}_{jj} \mathbf{V}_{\rho\sigma}(j)
\end{aligned}$$

$$\begin{aligned}
= & -\frac{1}{2} \sum_{\mu\nu} \delta P_{\mu\nu} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_e(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \\
& - \frac{1}{2} \sum_{\rho\sigma} \delta P_{\rho\sigma} \sum_j \langle\boldsymbol{\rho}|\frac{\tilde{\mathbf{q}}_e(j)}{|\mathbf{r}_e - \mathbf{r}_j}| |\boldsymbol{\sigma}\rangle. \quad (15)
\end{aligned}$$

Therefore, the sixth term leads to an operator representing two sets of potentials due to electron induced surface charges,

$$-\frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}P_{\rho\sigma} \sum_i \langle\boldsymbol{\mu}|\frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i}| |\boldsymbol{\nu}\rangle \Rightarrow \frac{1}{2} \frac{\mathbf{q}_e + \tilde{\mathbf{q}}_e}{|\mathbf{r}_e - \mathbf{r}_i|}, \quad (16)$$

where $\tilde{\mathbf{q}}_e$ is

$$\tilde{\mathbf{q}}_e = -\mathbf{K}\mathbf{C}^{-1}\mathbf{V}_e. \quad (17)$$

Because of the following inequality:

$$\mathbf{C}^{-1}\mathbf{K} \neq \mathbf{K}\mathbf{C}^{-1}, \quad (18)$$

$\tilde{\mathbf{q}}_e$ is different from \mathbf{q}_e .

The eighth term in Eq. (12) leads to an operator representing a potential due to nucleus induced surface charges,

$$\frac{1}{2} \sum_{\rho\sigma} P_{\rho\sigma} \sum_{\alpha} \sum_i \frac{Z_{\alpha}\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|} \Rightarrow \frac{1}{2} \sum_i \frac{\tilde{\mathbf{q}}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|}, \quad (19)$$

where $\tilde{\mathbf{q}}_N$ is

$$\tilde{\mathbf{q}}_N = -\mathbf{K}\mathbf{C}^{-1}\mathbf{V}_N. \quad (20)$$

Therefore, in heterogeneous COSMO and CPCM calculations, the following solvation operator \mathbf{B} shall be added to the gas phase Fock or Kohn-Sham operators:

$$\mathbf{B} = \frac{1}{2} \frac{\mathbf{q}_e + \tilde{\mathbf{q}}_e + \mathbf{q}_N + \tilde{\mathbf{q}}_N}{|\mathbf{r}_e - \mathbf{r}_i|} = \frac{1}{2} \frac{\mathbf{q} + \tilde{\mathbf{q}}}{|\mathbf{r}_e - \mathbf{r}_i|}. \quad (21)$$

Equation (21) shows that the \mathbf{B} operator can be constructed with the total induced charge $\mathbf{q} = \mathbf{q}_e + \mathbf{q}_N$ and $\tilde{\mathbf{q}} = \tilde{\mathbf{q}}_e + \tilde{\mathbf{q}}_N$. This is more efficient for iterative solution of the surface charges because only one solution is necessary. For homogeneous COSMO and CPCM, because $\tilde{\mathbf{q}}_e = \mathbf{q}_e$ and $\tilde{\mathbf{q}}_N = \mathbf{q}_N$, the \mathbf{B} operator is simpler,

$$\mathbf{B} = \frac{\mathbf{q}_e + \mathbf{q}_N}{|\mathbf{r}_e - \mathbf{r}_i|} = \frac{\mathbf{q}}{|\mathbf{r}_e - \mathbf{r}_i|}. \quad (22)$$

Equation (22) shows that for homogeneous COSMO and CPCM, it is possible to use only one set of induced charges because the \mathbf{C} matrix in Eq. (1) is symmetric, while for heterogeneous COSMO and CPCM at least two sets of induced charges are necessary because of the inequality shown by Eq. (18). It is noted that for DPCM and IEF-PCM, the corresponding matrices are asymmetric and two sets of induced charges shall be used for rigorously.

C. Nuclear gradients

A direct differentiation of Eq. (12) with respect to a nuclear coordinate x produces

$$\begin{aligned}
E_{\text{total}}^x = & \sum_{\mu\nu} P_{\mu\nu} (T_{\mu\nu}^x + V_{\mu\nu,N}^x) + \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu} P_{\rho\sigma} \langle \mu\nu | \rho\sigma \rangle^x + \sum_{\mu\nu} P_{\mu\nu} \langle \mu | V_{\text{xc}}(\mathbf{r}_e) | \nu \rangle^x + E_{NN}^x - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \\
& - \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu} P_{\rho\sigma} \sum_i \langle \mu | \frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x + \frac{1}{2} \sum_{\alpha} \sum_i \left(\frac{Z_{\alpha} \mathbf{q}_N(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|} \right)^x + \frac{1}{2} \sum_{\rho\sigma} P_{\rho\sigma} \left(\sum_{\alpha} \sum_i \frac{Z_{\alpha} \mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|} \right)^x + \sum_{\mu\nu} P_{\mu\nu}^x (T_{\mu\nu} + V_{\mu\nu,N}) \\
& + \sum_{\mu\nu\rho\sigma} P_{\mu\nu}^x P_{\rho\sigma} \langle \mu\nu | \rho\sigma \rangle + \sum_{\mu\nu} P_{\mu\nu}^x \langle \mu | V_{\text{xc}}(\mathbf{r}_e) | \nu \rangle - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu}^x \sum_i \langle \mu | \frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle - \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu}^x P_{\rho\sigma} \sum_i \langle \mu | \frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle \\
& - \frac{1}{2} \sum_{\mu\nu\rho\sigma} P_{\mu\nu} P_{\rho\sigma}^x \sum_i \langle \mu | \frac{\mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle + \frac{1}{2} \sum_{\rho\sigma} P_{\rho\sigma}^x \sum_{\alpha} \sum_i \frac{Z_{\alpha} \mathbf{q}_{\rho\sigma}(i)}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|}. \tag{23}
\end{aligned}$$

The last seven terms in Eq. (23) contain the derivatives of the density matrix and can be written as

$$\begin{aligned}
& \sum_{\mu\nu} P_{\mu\nu}^x (T_{\mu\nu} + V_{\mu\nu,N}) + \sum_{\mu\nu\rho\sigma} P_{\mu\nu}^x P_{\rho\sigma} \langle \mu\nu | \rho\sigma \rangle \\
& + \sum_{\mu\nu} P_{\mu\nu}^x \langle \mu | \frac{V_{\text{xc}}(\mathbf{r})}{|\mathbf{r}_e - \mathbf{r}|} | \nu \rangle + \sum_{\mu\nu} P_{\mu\nu}^x \mathbf{B}_{\mu\nu} \\
& = - \sum_{\mu\nu} W_{\mu\nu} S_{\mu\nu}^x, \tag{24}
\end{aligned}$$

where W is the energy-weighted density matrix. Clearly, the explicit evaluation of the density matrix derivatives can be avoided, as originally derived by Pulay for the gas phase Hartree–Fock methods.¹⁹

The first four and the last seven terms in Eq. (23) have the same forms as those in the gas phase Hartree–Fock and Kohn–Sham methods and can be evaluated using the same techniques with the solvent perturbed \mathbf{P} and \mathbf{W} . They are denoted as E_{gas}^x . The remaining terms (fifth, sixth, seventh, and eighth) in Eq. (23) contain the derivatives of the solvation terms and are denoted as E_{sol}^x ,

$$E_{\text{total}}^x = E_{\text{gas}}^x + E_{\text{sol}}^x. \tag{25}$$

The fifth term in Eq. (23) can be written as

$$\begin{aligned}
& - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \\
& = - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \mathbf{q}_N(i) \\
& - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle \mathbf{q}_N^x(i). \tag{26}
\end{aligned}$$

The derivative of the induced charge \mathbf{q}^x in Eq. (26) can be

avoided by converting the induced surface charge back into the potential

$$\begin{aligned}
& - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle \mathbf{q}_N^x(i) \\
& = \frac{1}{2} \sum_i \mathbf{V}_e(i) \mathbf{q}_N^x(i) \\
& = \frac{1}{2} \mathbf{V}_e \mathbf{q}_N^x \\
& = \frac{1}{2} \mathbf{V}_e (-\mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N)^x \\
& = - \frac{1}{2} \mathbf{V}_e [(\mathbf{C}^{-1})^x \mathbf{K} \mathbf{V}_N + \mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N^x] \\
& = \frac{1}{2} (\mathbf{V}_e \mathbf{C}^{-1}) \mathbf{C}^x (\mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N) - \frac{1}{2} \mathbf{V}_e \mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N^x \\
& = \frac{1}{2} (\mathbf{V}_e \mathbf{C}^{-1} \mathbf{K}) \mathbf{K}^{-1} \mathbf{C}^x (\mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N) - \frac{1}{2} \mathbf{V}_e \mathbf{C}^{-1} \mathbf{K} \mathbf{V}_N^x \\
& = \frac{1}{2} \tilde{\mathbf{q}}_e^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_N + \frac{1}{2} \tilde{\mathbf{q}}_e^T \mathbf{V}_N^x. \tag{27}
\end{aligned}$$

Note that \mathbf{K} is a constant matrix, so $\mathbf{K}^x = \mathbf{0}$ because the local dielectrics are not subject to change with respect to x , as discussed at the end of Sec. II A. So the fifth term in Eq. (23) can be written as

$$\begin{aligned}
& - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{\mathbf{q}_N(i)}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \\
& = - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \mathbf{q}_N(i) \\
& + \frac{1}{2} \tilde{\mathbf{q}}_e^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_N + \frac{1}{2} \tilde{\mathbf{q}}_e^T \mathbf{V}_N^x. \tag{28}
\end{aligned}$$

Similarly, the sixth, seventh, and eighth terms can also be simplified, and E_{sol}^x can be written as

$$\begin{aligned}
E_{\text{sol}}^x &= -\frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \mathbf{q}(i) \\
&\quad - \frac{1}{2} \sum_{\rho\sigma} P_{\rho\sigma} \sum_i \langle \rho | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \sigma \rangle^x \tilde{\mathbf{q}}(i) + \frac{1}{2} \tilde{\mathbf{q}}^T \mathbf{V}_N^x \\
&\quad + \frac{1}{2} (\mathbf{V}_N^x)^T \mathbf{q} + \frac{1}{2} \tilde{\mathbf{q}}^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_N + \frac{1}{2} \tilde{\mathbf{q}}^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_e \\
&\quad + \frac{1}{2} \tilde{\mathbf{q}}_N^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_N + \frac{1}{2} \tilde{\mathbf{q}}_N^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}_e \\
&= - \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \left[\frac{\tilde{\mathbf{q}}(i) + \mathbf{q}(i)}{2} \right] \\
&\quad + (\mathbf{V}_N^x)^T \left(\frac{\tilde{\mathbf{q}} + \mathbf{q}}{2} \right) - \frac{1}{2} \tilde{\mathbf{q}}^T \mathbf{K}^{-1} \mathbf{C}^x \mathbf{q}. \tag{29}
\end{aligned}$$

Equation (29) shows that the derivatives of the solvation terms can be evaluated with the induced surface charges \mathbf{q} and $\tilde{\mathbf{q}}$.

In homogeneous COSMO and CPCM method, $\tilde{\mathbf{q}} = \mathbf{q}$, so E_{sol}^x becomes

$$\begin{aligned}
E_{\text{sol}}^x &= - \sum_{\mu\nu} P_{\mu\nu} \sum_i \langle \mu | \frac{1}{|\mathbf{r}_e - \mathbf{r}_i|} | \nu \rangle^x \mathbf{q}(i) + (\mathbf{V}_N^x)^T \mathbf{q} \\
&\quad + \frac{1}{2} \left(\frac{\varepsilon}{\varepsilon - 1} \right) \mathbf{q}^T \mathbf{C}^x \mathbf{q}. \tag{30}
\end{aligned}$$

The first and second terms in Eqs. (29) or (30) are the electrostatic fields at the surface tesserae due to solute electrons and nuclei and can be evaluated using standard techniques.

The third term in Eqs. (29) or (30) is the interactions between the induced surface charges and can be evaluated with the derivatives of the \mathbf{C} matrix, which involves the derivatives of the areas and coordinates of the tesserae with respect to the atomic coordinate x . The generating-polyhedra (GEPOL)^{20,21} scheme has been implemented for DPCM,² IEF-PCM,^{3,4} and CPCM^{7,8} for both energy and gradients calculations, and exact analytic gradients, as well as smooth potential energy surfaces,²¹ have been obtained. Recently, Su and Li¹⁷ implemented a tessellation scheme called FIXPVA for COSMO and CPCM. In FIXPVA the tessera areas are smooth functions of their distance to neighboring spheres, so rigorously continuous and smooth potential energy surfaces, as well as exact analytic gradients, can be obtained for COSMO and CPCM calculations with Hartree-Fock, multi-configuration self-consistent field, and DFT methods.

III. COMPUTATIONAL METHODOLOGY

The heterogeneous CPCM code was implemented by the authors in the GAMESS (Ref. 22) package as a new option of the CPCM code previously implemented by Li and Jensen⁸ on the basis of the IEF-PCM program originally implemented by Cancès *et al.*^{3,23} and Li *et al.*²⁴

In the current implementation, local dielectric constants are assigned to surface tesserae according to the spheres (i.e., atoms in the FIXPVA tessellation scheme) they belong to;

TABLE I. Solvation energy (kcal/mol) calculated with heterogeneous CPCM/RHF/aug-cc-pVTZ for NaCl using different local effective dielectrics for Na and Cl.

	Na ($\varepsilon=1$)	Na ($\varepsilon=2$)	Na ($\varepsilon=4$)	Na ($\varepsilon=20$)	Na ($\varepsilon=78.39$)
Cl ($\varepsilon=1$)	0.00	-10.46	-15.98	-20.53	-21.39
Cl ($\varepsilon=2$)	-3.84	-14.83	-20.62	-25.39	-26.29
Cl ($\varepsilon=4$)	-5.93	-17.20	-23.14	-28.03	-28.95
Cl ($\varepsilon=20$)	-7.68	-19.19	-25.26	-30.25	-31.20
Cl ($\varepsilon=78.39$)	-8.02	-19.58	-25.66	-30.68	-31.63

the tesserae on the same sphere have the same local dielectric constant. Only the electrostatic interaction was considered; heterogeneous cavitation, dispersion, and repulsion terms were not considered.

In the CPCM calculations, spheres with radii of 0, 2.124, 2.016, 1.908, 1.800, 2.520, 2.760, and 2.760 Å were used for H, C, N, O, Na, S, Cl, and Cu atoms, respectively, to define the molecular cavity; no additional spheres were used. Using zero radii for H atoms means that they do not contribute to form the surface. The tessellation scheme FIXPVA was used with 60 initial tesserae per sphere.¹⁷ The induced surface charges were determined by a semi-iterative direct inversion in the iterative subspace procedure^{24,25} with no charge renormalization. Geometry optimization was performed in internal coordinates generated by the automatic delocalized coordinates algorithm.²⁶

IV. RESULTS AND DISCUSSION

A. Energy

The simplest way to establish a quantitative picture of the heterogeneous COSMO and CPCM method is probably to use an ion pair. Table I presents the solvation energies [E_{sol} in Eq. (13)] calculated with the heterogeneous CPCM/RHF/aug-cc-pVTZ (Ref. 27) method for NaCl using various dielectrics for the tesserae on the Na sphere and the Cl sphere. The NaCl distance is 2.397 Å as optimized with the gas phase RHF/aug-cc-pVTZ method. Spheres with radii of 1.800 and 2.760 Å were used for Na and Cl, respectively, to define the molecular cavity. It is emphasized that these calculations were performed for illustrative purposes rather than to reproduce any experimental or theoretical results.

The diagonal data in Table I are actually homogeneous CPCM results because the same dielectric was used for Na and Cl. Ongoing from low dielectric 2.00 to high dielectric 78.39, the solvation energy changes from -14.83 to -31.63 kcal/mol, roughly doubled. This rough doubling can be easily explained by the scaling factor $(\varepsilon-1)/\varepsilon$, which changes from 0.5 to ~ 1.0 . However, due to the changes in the polarization of the NaCl electron density ongoing from low to high dielectric solvents, the actual solvation energy is more than doubled. If the Na⁺ and Cl⁺ are represented by point charges with no polarizability, the solvation energy will be exactly doubled ongoing from $\varepsilon=2$ to $\varepsilon=\infty$.

Holding $\varepsilon=1$ for Cl but varying ε from 1 to 2, 4, 20, and 78.39 for Na, the solvation energy changes from 0.00 to -10.46, -15.98, -20.53, and -21.39 kcal/mol. Holding

TABLE II. Gradients (au) calculated for acetate CH_3COO^- using local effective $\epsilon=4.0$ for the CH_3 group and $\epsilon=78.39$ for the COO^- group.

Coordinates	Analytic	Numerical
C1X	0.000 997 39	0.000 997 45
C1Y	-0.001 462 86	-0.001 462 85
C1Z	0.007 292 40	0.007 292 45
C2X	-0.001 143 79	-0.001 144 10
C2Y	-0.001 225 23	-0.001 225 50
C2Z	0.000 551 69	0.000 551 25
O3X	-0.007 635 35	-0.007 635 05
O3Y	0.000 021 60	0.000 021 60
O3Z	-0.001 177 25	-0.001 177 20
O4X	0.002 315 35	0.002 315 25
O4Y	-0.006 632 41	-0.006 632 20
O4Z	-0.002 709 82	-0.002 709 60
H5X	-0.010 203 54	-0.010 204 05
H5Y	-0.016 417 03	-0.016 417 15
H5Z	0.007 275 99	0.007 276 05
H6X	-0.020 577 57	-0.020 577 60
H6Y	0.026 713 68	0.026 713 80
H6Z	-0.017 705 58	-0.017 705 60
H7X	0.036 247 51	0.036 247 85
H7Y	-0.000 997 75	-0.000 997 75
H7Z	0.006 472 57	0.006 472 60
Max error		0.000 000 44
rms error		0.000 000 22

$\epsilon=1$ for Na but varying ϵ from 1 to 2, 4, 20, and 78.39 for Cl, the solvation energy changes from 0.00 to -3.84 , -5.93 , -7.68 , and -8.02 kcal/mol. Clearly, the Na exhibits a greater ϵ dependence because of its smaller radius (1.8 Å for Na and 2.76 Å for Cl).

B. Gradients

Table II presents the analytic and numerical gradients obtained with heterogeneous CPCM/B3LYP²⁸/6-31G* methods for acetate in the gas phase HF/6-31G* optimized geometry. The numerical gradients were computed with double displacements (forward and backward) using a step size of 0.001 a.u. for each step. A local effective $\epsilon=78.39$ was used for the carboxylate group and $\epsilon=4.0$ was used for the methyl group. Again, this is for illustrative purposes.

The maximum difference between the numerical and analytic gradients is 4.4×10^{-7} a.u., and the root mean square difference is 2.2×10^{-7} a.u. The default settings in GAMESS produce the gas phase gradients good to about 10^{-7} – 10^{-6} a.u. In this sense, the exact gradients are obtained. It is noted that such exact gradients are partially due to the use of the FIXPVA tessellation scheme developed by the author's group.¹⁷ Different tessellation schemes may lead to different solvation energies, total molecular energies, and molecular potential energy surfaces. For example, compared to GEPOL-GB (with no additional spheres), the FIXPVA scheme produces $\sim 10\%$ less surface area and ~ 1 kcal/mol smaller CPCM solvation energies for some typical solutes such as acetate anion.¹⁷ Such small differences in solvation energies have also been observed for the current heteroge-

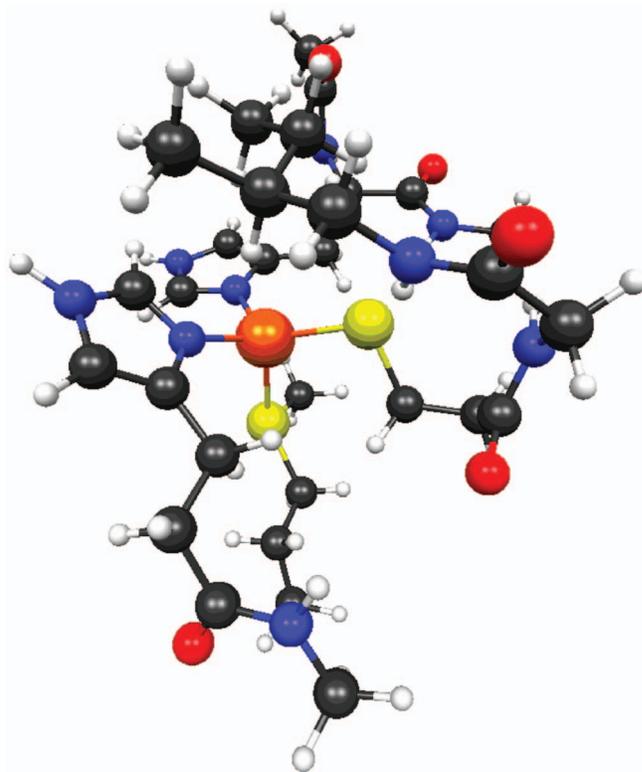


FIG. 1. A 101-atom model molecule extracted from the x-ray crystal structures 2CAK for the type-1 Cu center of rusticyanin.

neous CPCM method. The expression of the analytic gradients for the heterogeneous CPCM method [i.e., Eq. (29)] is general and can be implemented by using different tessellation schemes such as GEPOL-GB as long as the derivatives of the tessera areas and coordinates are available.

C. Type-1 Cu center

Here a realistic example, the type-1 Cu center of rusticyanin is presented to show the application of the heterogeneous COSMO and CPCM methods in quantum chemical study of protein active sites. This example is taken from a separate study of the reduction potential of five type-1 Cu centers performed by the authors. The details will be published in a forthcoming paper.

With 155 residues and ~ 2310 atoms, rusticyanin is a small blue copper protein found in iron-oxidizing bacterium *Thiobacillus ferrooxidans*. This protein displays an unusual acid stability and its type-1 Cu center exhibits a very high reduction potential of $E^0=680$ mV, which is vital for the bacteria to survive.

An active site model molecule consisting of 101 atoms (Figs. 1 and 2) was extracted from the x-ray structures 2CAK in the protein data bank for rusticyanin.²⁹ A comparison between the x-ray structure and the model molecule shows that only the C-terminal His143 imidazole ring in the model molecule is solvent exposed. With 26 atoms fixed, the coordinates of 75 atoms were first optimized with the homogeneous CPCM($\epsilon=78.39$)/B3LYP²⁸/6-31G* method and then with the heterogeneous CPCM/B3LYP/6-31G* method (Fig. 2).

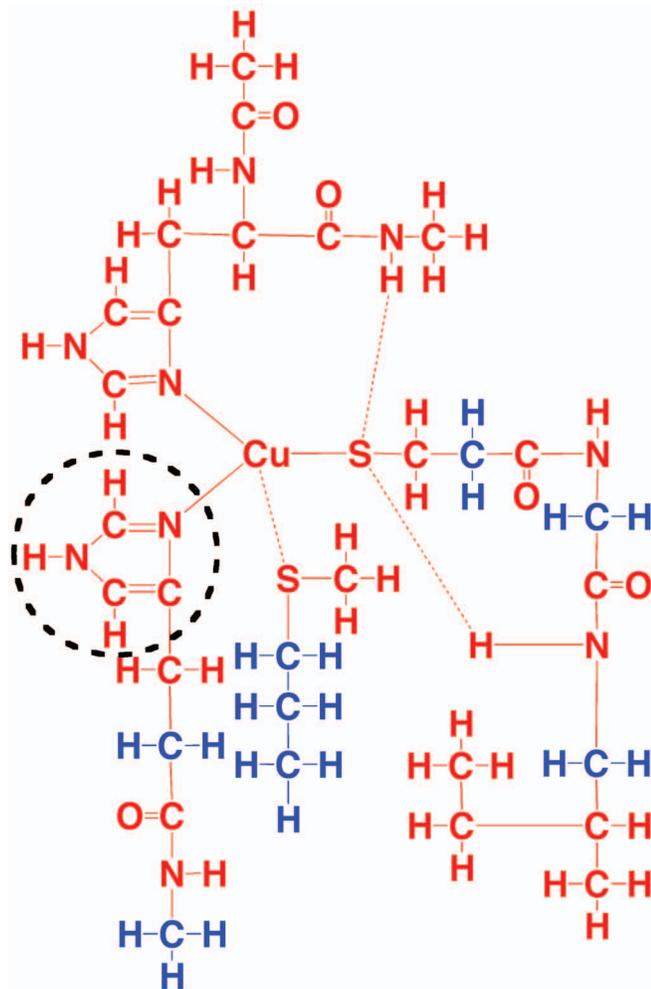


FIG. 2. With 26 atoms (in blue) fixed, 75 atoms (in red) in the 101-atom model molecule are optimized using the heterogeneous CPCM/B3LYP/6-31G* method in which the solvent-exposed C-terminal histidine imidazole group (cycled) uses $\epsilon=78.39$, while all the other atoms use $\epsilon=4$.

The R-B3LYP and RO-B3LYP type of wave function for Cu⁺ and Cu²⁺ oxidation states, respectively, were used.

In the heterogeneous CPCM calculations, the tesseræ associated with the solvent exposed imidazole ring were assigned with a local effective $\epsilon=78.39$, while all the other atoms were assigned with a local effective $\epsilon=4$ as they are buried inside of the protein matrix. Using 4 is reasonable because effective dielectrics from 4 to 20 have been commonly used for protein interiors. A molecular dynamics simulation suggests that the protein matrix around the rusticyanin type-1 Cu center is highly hydrophobic and rigid, which corresponds to a low effective dielectric.³⁰ Of course,

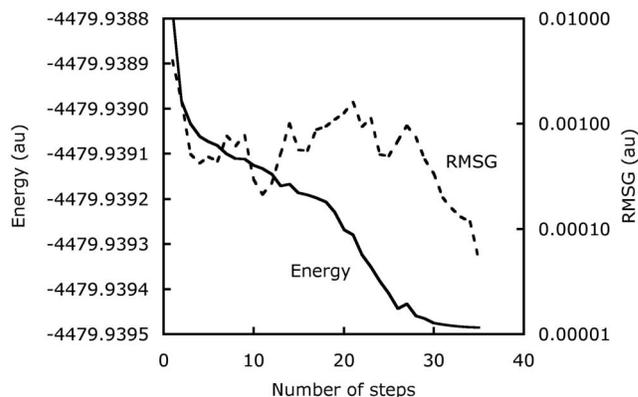


FIG. 3. Energy and RMSG profiles in the heterogeneous PCM/B3LYP/6-31G* geometry optimization of the reduced form of the 101-atom model molecule.

the accurate effective dielectric is unknown and could be significantly different at different portions around the type-1 center. In general, determining protein interior dielectric is a difficult issue.³¹

Figure 3 shows the energy and root mean square gradient (RMSG) profiles along the heterogeneous CPCM/B3LYP/6-31G* geometry optimization course for the model molecule extracted from 2CAK. As mentioned above, the geometry has already been optimized using homogeneous CPCM/B3LYP/6-31G* method so only a small energy change, ~ 0.4 kcal/mol, is seen in Fig. 3. The energy decreases rapidly in the first few steps, which is typical in geometry optimization calculations for such systems. Then the energy keeps decreasing monotonically in the following steps before convergence. The convergence threshold is the default in GAMESS, i.e., maximum gradient is smaller than 1.0×10^{-4} a.u. and the RMSG is smaller than 0.333×10^{-4} a.u. Similar profiles were observed in many other cases. In general, the geometry optimized using the heterogeneous CPCM method is as robust as the homogeneous CPCM method and the gas phase methods.

It is of interest to know how the bulk water and protein matrix affect the reduction potential of the type-1 Cu center in rusticyanin. Table III lists the CPCM/B3LYP/6-31G* energies of the 101-atom model molecule in reduced and oxidized forms. Using the homogeneous CPCM($\epsilon=78.39$)/B3LYP/6-31G* optimized structure and energy, the energy difference is 81.29 kcal/mol or 3525 mV. The subsequent heterogeneous CPCM/B3LYP/6-31G* single point energy calculations show that the energy difference is 87.51 kcal/mol or 3794 mV. Using the heterogeneous CPCM/B3LYP/6-31G* optimized structure and energy, the energy difference

TABLE III. CPCM/B3LYP/6-31G* energies of the 101-atom model molecule extracted from the x-ray structure 2CAK for rusticyanin.

Methods	Reduced (a.u.)	Oxidized (a.u.)	Δ (kcal/mol)	Δ (mV)
Homo-CPCM optimized	-4479.957 144	-4479.827 595	81.29	3525
Het CPCM on homo-CPCM optimized	-4479.938 784	-4479.799 328	87.51	3794
Het CPCM optimized	-4479.939 485	-4479.802 762	85.79	3720

is 85.79 kcal/mol or 3720 mV. Therefore, in rusticyanin the protein burial can likely raise the reduction potential by ~ 200 mV. It is also important to include the heterogeneity in geometry optimization as it can likely affect the reduction potential by ~ 70 mV.

V. CONCLUSION

A heterogeneous conductorlike solvation model that uses different local effective dielectrics for different portions of a solute cavity surface was derived and implemented in the CPCM code in GAMESS. By variationally formulating the solvation operators for Hartree–Fock and Kohn–Sham methods, using the FIXPVA tessellation scheme, and assuming that the local effective dielectrics are constants in a geometry optimization process, continuous and smooth potential energy surfaces and analytic nuclear gradients (accurate to 10^{-7} – 10^{-6} a.u.) have been obtained for the heterogeneous solvation model. Application of the heterogeneous solvation model to a realistic quantum model consisting of 101 atoms for the type-1 Cu center in rusticyanin shows that the desolvation due to protein burial can likely raise the reduction potential by ~ 200 mV and, including the heterogeneity in geometry optimization, can likely affect the results by ~ 2 kcal/mol or ~ 70 mV.

Finally, it is emphasized that although the local effective dielectrics are physically meaningful, their exact values can be difficult to obtain. Future studies in this direction may require force field molecular dynamics simulations or empirical calibrations. For quantum calculations of protein active sites, it might be possible to develop an automatic and physics motivated procedure to assign local effective dielectrics for the heterogeneous solvation model according to the nature and positions of residues near the active sites.

ACKNOWLEDGMENTS

This work was supported by start-up funds from the University of Nebraska-Lincoln.

¹J. Tomasi, B. Mennucci, and R. Cammi, *Chem. Rev. (Washington, D.C.)* **105**, 2999 (2005).

²S. Miertus, E. Scrocco, and J. Tomasi, *Chem. Phys.* **55**, 117 (1981).

³E. Cancès, B. Mennucci, and J. Tomasi, *J. Chem. Phys.* **107**, 3032 (1997).

⁴B. Mennucci, E. Cancès, and J. Tomasi, *J. Phys. Chem. B* **101**, 10506 (1997).

⁵A. Klamt and G. Schuurmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.

⁶T. N. Truong and E. V. Stefanovich, *Chem. Phys. Lett.* **240**, 253 (1995).

⁷V. Barone and M. Cossi, *J. Phys. Chem. A* **102**, 1995 (1998).

⁸H. Li and J. H. Jensen, *J. Comput. Chem.* **25**, 1449 (2004).

⁹D. M. Chipman, *Theor. Chem. Acc.* **107**, 80 (2002); D. M. Chipman and M. Dupuis, *ibid.* **107**, 90 (2002).

¹⁰A. V. Marenich, R. M. Olson, C. P. Kelly, C. J. Cramer, and D. G. Truhlar, *J. Chem. Theory Comput.* **3**, 2011 (2007).

¹¹R. Bonaccorsi, E. Scrocco, and J. Tomasi, *Int. J. Quantum Chem.* **29**, 717 (1986).

¹²R. Bonaccorsi, M. Hodoseck, and J. Tomasi, *J. Mol. Struct.: THEOCHEM* **164**, 105 (1988).

¹³R. Bonaccorsi, E. Ojalvo, and J. Tomasi, *Collect. Czech. Chem. Commun.* **53**, 2320 (1988); R. Bonaccorsi, F. Floris, P. Palla, and J. Tomasi (1990) *Thermochimica Acta* **162**, 213-222 (1990); R. Bonaccorsi, E. Ojalvo, P. Palla, and J. Tomasi, *Chem. Phys.* **143**, 245 (1990).

¹⁴H. Hoshi, M. Sakurai, Y. Inoue, and R. Chujo, *J. Chem. Phys.* **87**, 1107 (1987).

¹⁵L. Frediani, R. Cammi, S. Corni, and J. Tomasi, *J. Chem. Phys.* **120**, 3893 (2004).

¹⁶M. F. Iozzi, M. Cossi, R. Improta, N. Rega, and V. Barone, *J. Chem. Phys.* **124**, 184103 (2006).

¹⁷P. Su and H. Li, *J. Chem. Phys.* **130**, 074109 (2009).

¹⁸M. Cossi, N. Rega, G. Scalmani, and V. Barone, *J. Comput. Chem.* **24**, 669 (2003).

¹⁹P. Pulay, *Mol. Phys.* **17**, 197 (1969).

²⁰J. L. Pascualahir, E. Silla, and I. Tunon, *J. Comput. Chem.* **15**, 1127 (1994); E. Silla, I. Tunon, and J. L. Pascualahir, *ibid.* **12**, 1077 (1991); E. Silla, F. Villar, O. Nilsson, J. L. Pascualahir, and O. Tapia, *J. Mol. Graphics* **8**, 168 (1990); J. L. Pascualahir and E. Silla, *J. Comput. Chem.* **11**, 1047 (1990).

²¹C. S. Pomelli, *J. Comput. Chem.* **25**, 1532 (2004).

²²M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993); M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005).

²³J. Tomasi, B. Mennucci, and E. Cancès, *J. Mol. Struct.: THEOCHEM* **464**, 211 (1999).

²⁴H. Li, C. S. Pomelli, and J. H. Jensen, *Theor. Chem. Acc.* **109**, 71 (2003).

²⁵C. S. Pomelli, J. Tomasi, and V. Barone, *Theor. Chem. Acc.* **105**, 446 (2001).

²⁶J. Baker, A. Kessi, and B. Delley, *J. Chem. Phys.* **105**, 192 (1996).

²⁷T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989); N. B. Balabanov and K. A. Peterson, *ibid.* **123**, 064107 (2005).

²⁸R. H. Hertwig and W. Koch, *Chem. Phys. Lett.* **268**, 345 (1997).

²⁹M. L. Barrett, I. Harvey, M. Sundararajan, R. Surendran, J. F. Hall, M. J. Ellis, M. A. Hough, R. W. Strange, I. H. Hillier, and S. S. Hasnain, *Biochemistry* **45**, 2927 (2006); H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, *Nucleic Acids Res.* **28**, 235 (2000).

³⁰B. Jimenez, M. Piccioli, J. M. Moratal, and A. Donaire, *Biochemistry* **42**, 10396 (2003).

³¹C. N. Schutz and A. Warshel, *Proteins* **44**, 400 (2001).