Supporting information for "Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces"

Adrian W. Lange and John M. Herbert*

Department of Chemistry, The Ohio State University, Columbus, OH 43210

1 Additional information on the SWIG formalism

In Section 1.1 of this Supporting Information, we present a formal derivation of the SWIG methodology and a rigorous proof that this method affords smooth potential energy surfaces. These arguments generalize the S-COSMO model of York and Karplus[1] to more a more general class of PCMs.[2–4] In slightly modified form, these arguments can be used to extend the FIXPVA method[5] to these more general PCMs, although the FIXPVA method as described in Ref. 5 uses a different switching function that the one employed here. Our switching function is described in detail in Section 1.2.

1.1 Derivation

Chipman[2–4] provides a detailed discussion of the integral-equation formalism for apparent surface charge PCMs, which includes the C-PCM/COSMO, IEF-PCM, and SS(V)PE methods, and the reader is referred to Refs. 2–4 for details. For our purposes, it suffices to note that the PCM equations, in the continuous representation, are expressed in terms of a pair of integral operators, $\hat{S}$ and $\hat{D}$. The operator $\hat{S}$ acts on the surface charge density, $\sigma(\mathbf{s})$, to generate the corresponding surface potential at the surface point $\mathbf{s}$,

$$\hat{S}\sigma(\mathbf{s}) \int_{\substack{cavity \\ surface}} d\mathbf{s}' \, \sigma(\mathbf{s}') \frac{1}{|\mathbf{s}-\mathbf{s}'|} \, . \tag{S1}$$

*herbert@chemistry.ohio-state.edu

whereas the action of $\hat{D}$ on $\sigma(\mathbf{s})$ generates the surface dipole potential at $\mathbf{s}$,

$$\hat{D}\sigma(\mathbf{s}) = \int_{\substack{cavity \\ surface}} d\mathbf{s}' \; \sigma(\mathbf{s}') \frac{\partial}{\partial \mathbf{n}_{\mathbf{s}'}} \frac{1}{|\mathbf{s} - \mathbf{s}'|} \; . \tag{S2}$$

Here $\mathbf{n}_{\mathbf{s}'}$ indicates the normalized, outward-pointing surface normal vector at the surface point $\mathbf{s}'$.

The matrices $\mathbf{S}$ and $\mathbf{D}$ are discretized representations of the operators $\hat{S}$ and $\hat{D}$. Within the SWIG approach, we blur the point charge q_i, which is located at the ith Lebedev grid point, $\mathbf{r}_i$, using a normalized gaussian function

$$g_i(\mathbf{r}) = (\zeta_i^2/\pi)^{3/2} \; e^{-\zeta_i^2 |\mathbf{r} - \mathbf{r}_i|^2} \; . \tag{S3}$$

The exponents of these gaussians are treated as fixed parameters that characterize the model.

The matrix element S_{ij} of $\mathbf{S}$ represents the projection of $\hat{S} g_j$ onto g_i. For $i \neq j$, we determine S_{ij} by using Eq. (S1), projecting onto g_i, and evaluating surface integrals over $\mathbf{s}$ and $\mathbf{s}'$. The result is

$$S_{ij} = \frac{\mathrm{erf}(\zeta_{ij} \, r_{ij})}{r_{ij}} \quad \text{for } i \neq j \; , \tag{S4}$$

where $\zeta_{ij} = \zeta_i \zeta_j / (\zeta_i^2 + \zeta_j^2)^{1/2}$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Diagonal elements of $\mathbf{S}$ are obtained according to $S_{ii} = \lim_{r_{ij} \to 0} S_{ij}$. Recognizing that $\zeta_{ii} = \zeta_i / \sqrt{2}$, this limit is evaluated to obtain

$$S_{ii} = \zeta_i \sqrt{2/\pi} \; . \tag{S5}$$

Equations (S4) and (S5) are nearly the same as Eq. (3) in the text, except for a factor of F_i^{-1} that appears in S_{ii} as given in Eq. (3). We next explain the origin of this factor.

According to Eq. (1) in the text, the surface charges q_i are determined according to $\mathbf{q} = \mathbf{K}^{-1}\mathbf{R}\mathbf{v}$. If a certain switching function $F_k = 0$, however, we wish to ensure that $q_k = 0$, so that the kth surface grid point does not contribute to the PCM solvation energy. However, the electrostatic potential v_k is generally nonzero, even if $F_k = 0$, and furthermore $\mathbf{R} \propto \mathbf{I}$ for C-PCM/COSMO. Hence the kth element of $\mathbf{R}\mathbf{v}$ is generally nonzero. As such, the only way to ensure that the "switched-off"

grid points do not contribute to the solvation energy is to require that $\mathbf{K}^{-1}$ has a null space that corresponds precisely to those surface grid points for which $F_k = 0$.

Partitioning $\mathbf{K}$ into diagonal and off-diagonal parts,

$$\mathbf{K} = \mathbf{K}_{\rm diag} + \mathbf{K}_{\rm off} \ , \tag{S6}$$

in which the matrix $\mathbf{K}_{\rm diag}$ is diagonal, we may write

$$\mathbf{K} = \mathbf{K}_{\rm diag}^{1/2} \Big(\mathbf{I} + \mathbf{K}_{\rm diag}^{-1/2} \mathbf{K}_{\rm off} \mathbf{K}_{\rm diag}^{-1/2} \Big) \mathbf{K}_{\rm diag}^{1/2} \ , \tag{S7}$$

whence

$$\mathbf{K}^{-1} = \mathbf{K}_{\rm diag}^{-1/2} \Big(\mathbf{I} + \mathbf{K}_{\rm diag}^{-1/2} \mathbf{K}_{\rm off} \mathbf{K}_{\rm diag}^{-1/2} \Big)^{-1} \mathbf{K}_{\rm diag}^{-1/2} \ . \tag{S8}$$

Now, if the matrix elements of diagonal matrix $\mathbf{K}_{\rm diag}^{-1}$ were zero whenever $F_k = 0$, the the effect of the right and left multiplication by $\mathbf{K}_{\rm diag}^{-1/2}$ in Eq. (S8) would be to annihilate any rows and columns of $\big(\mathbf{I} + \mathbf{K}_{\rm diag}^{-1/2} \mathbf{K}_{\rm off} \mathbf{K}_{\rm diag}^{-1/2} \big)^{-1}$ for which $F_k = 0$, or in other words, to create the desired null space for $\mathbf{K}^{-1}$. Similarly, the matrix $\mathbf{I} + \mathbf{K}_{\rm diag}^{-1/2} \mathbf{K}_{\rm off} \mathbf{K}_{\rm diag}^{-1/2}$ would become block diagonal and equal to the unit matrix within the null space of $\mathbf{K}^{-1}$, and the dimension of $\mathbf{q} = \mathbf{K}^{-1} \mathbf{R} \mathbf{v}$ could be reduced *without approximation* to just the number of grid points whose switching functions are non-zero.

These arguments suggest that we should replace the expression for S_{ii} in Eq. (S5) with

$$S_{ii} = \frac{\zeta_i \sqrt{2/\pi}}{F_i} \ , \tag{S9}$$

so that $S_{ii} \to \infty$ as $F_i \to 0$. This equation, together with Eq. (S4) for S_{ij}, defines the **S**-matrix used in this work. Note that

$$(\mathbf{DAS})_{ii} = \sum_{j \neq i} D_{ij} a_j S_{ji} \ , \tag{S10}$$

owing to our choice $D_{ii} = 0$, as discussed in the text. Since S_{ij} is finite, as is $D_{ij} = -\hat{\mathbf{n}}_j \cdot (\partial S_{ij} / \partial \mathbf{r}_j)$ [Eq. (6)], it follows that $K_{ii} \to \infty$ as $F_i \to 0$, for both C-PCM/COSMO and SS(V)PE. As York and Karplus point out,[1] the quantity S_{ii} is the self-energy of the ith surface tessera, so the definition in

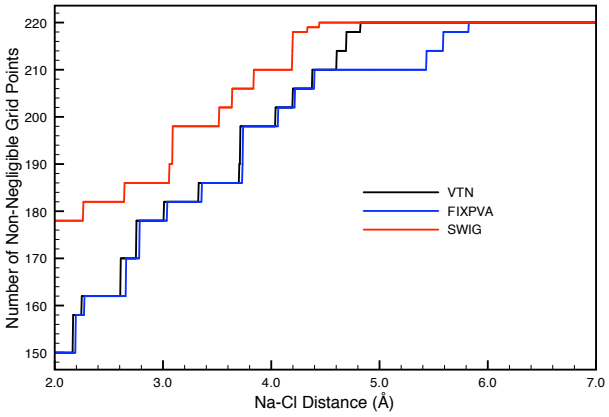

Figure S1: Number of non-negligible ($F_i > 10^{-8}$) Lebedev grid points for NaCl, as a function of the Na–Cl distance.

Eq. (S9) has the effect of imposing an increasing steep penalty for polarizing the ith tessera, as the ith Lebedev point enters and passes through the the switching region from the cavity surface, until finally $S_{ii} = \infty$ within the cavity interior, and the ith tesserae can no longer be polarized.

Although the dimension of the null space of $\mathbf{K}^{-1}$ can and does change as the atoms are moved, by construction (Section 1.2), $F_i \to 0$ smoothly, and therefore the block structure of $\mathbf{K}^{-1}$ changes smoothly as well. This proves that the SWIG potential energy surfaces are rigorously smooth, using exact arithmetic. In practice, of course, we must employ a finite drop tolerance, and in the present work we remove any Lebedev point for which $F_i < 10^{-8}$. (In practice, this is no different than the finite drop tolerances used to discard shell pairs an to screen AO integrals in standard electronic structure calculations. Potential energy surfaces are only as smooth as these finite thresholds allow, but the thresholds can be made as small as machine precision, if desired.) Figure S1 shows how the number of non-negligible points (and thus the dimension of $\mathbf{K}$) changes as a function of bond length, for the NaCl calculation described in the paper. The dimension of $\mathbf{K}$ changes numerous times as the atoms are pulled apart, yet the SWIG potential surface remains smooth.

1.2 Switching function

The switching function F_i used in the current work is equivalent to that used by York and Karplus.[1] For a given surface grid point, $\mathbf{r}_i$, the function F_i is a product of elementary switching functions, f:

$$F_i = \prod_{J, i \notin J}^{atoms} f(\hat{r}_{iJ}) \ . \tag{S11}$$

The notation "$i \notin J$" indicates that the product over J omits the atom associated with the ith Lebedev point. The elementary switching function used in Eq. (S11) is

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^3(10 - 15x + 6x^2) & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases} \tag{S12}$$

The argument of f in Eq. (S11) is a dimensionless number, $\hat{r}_{iJ}$, that quantifies the extent to which the ith discretization point has penetrated into the buffer region surrounding the Jth atom:

$$\hat{r}_{iJ} = \frac{r_{iJ} - R_{\text{in},J}}{R_{\text{sw},J}} \ . \tag{S13}$$

Here, $r_{iJ} = |\mathbf{r}_i - \mathbf{R}_J|$ is the distance from the ith discretization point to the Jth atom, $R_{\text{in},J} < R_J$ is the radius of the inner limit of the switching region for atom J, and $R_{\text{sw},J}$ is the width of this switching region. The parameters $R_{\text{in},J}$ and $R_{\text{sw},J}$ are determined by R_J, the radius of the Jth atomic sphere, and N_J, the number of Lebedev points used to discretize this sphere. Specifically,[1]

$$R_{\text{sw},J} = \gamma_J R_J \tag{S14}$$

and

$$R_{\text{in},J} = R_J - \alpha_J R_{\text{sw},J} \ , \tag{S15}$$

where

$$\alpha_J = \frac{1}{2} + \frac{1}{\gamma_J} - \sqrt{\frac{1}{\gamma_J^2} - \frac{1}{28}} \tag{S16}$$

and

$$\gamma_J = \sqrt{14/N_J}\,. \tag{S17}$$

2 Computational details

Solute cavities were discretized using atom-centered Lebedev grids.[6] Grid points were discarded whenever $F_i < 10^{-8}$; tighter thresholds had little discernible effect on the calculations. The rest of this section provides details regarding the (adenine)$(H_2O)_{52}$ and NaCl calculations discussed in this work.

2.1 (Adenine)$(H_2O)_{52}$ optimization and frequencies

An (adenine)$(H_2O)_{52}$ cluster was carved out of an equilibrated molecular dynamics simulation of adenine in bulk water under ambient conditions, by selecting all H_2O molecules within 5 Å of any adenine atom. This cluster was optimized in the gas phase (using the AMBER99 force field) for 50 optimization cycles, in order to eliminate any artifacts in the gradient caused by extracting the cluster from the bulk. The resulting structure was then used as a starting point for the VTN-, FIXPVA-, and SWIG-COSMO optimizations.

Solute cavities for these optimizations were constructed from atom-centered spheres whose radii were chosen by adding 1.4 Å to the AMBER99 Lennard–Jones radii. This surface was discretized using 50 Lebedev points per atomic sphere.

Q-Chem default convergence criteria were used for the geometry optimizations (maximum step size = 0.3 Å; maximum gradient component = 3.0×10^{-4} a.u.; maximum atomic displacement = 1.2×10^{-3} a.u.; maximum energy change = 1.0×10^{-6} a.u.). Harmonic frequencies were calculated by finite difference of analytic energy gradients. For the SWIG procedure, the finite-difference step size was taken to be the Q-Chem default, 3.0×10^{-4} Å. The step size used for FIXPVA was reduced to 3.0×10^{-5} Å, in order to avoid imaginary frequencies that were encountered when larger displacements

were used.

2.2 NaCl dissociation

Dissociation curves for NaCl [computed at the HF/6-31+G*/SS(V)PE level] consist of single-point calculations spaced 0.005 Å apart. The total energy includes the SS(V)PE electrostatic interaction energy as well as non-electrostatic cavitation, dispersion, and repulsion interactions.

Each non-electrostatic component of the solvation energy is defined by a fairly simple function of the cavity surface area. The cavitation energy was computed using the formula given in Ref. 7, whereas dispersion and repulsion energies were computed as described in Ref. 8 Lennard–Jones parameters for the dispersion/repulsion energy were taken from the AMBER99 force field[9] for Na^+ and Cl^-, and from the TIP3P force field[10] for H_2O.

As suggested in Ref. 7, the non-electrostatic interaction terms utilize a somewhat different cavity surface than that used for the electrostatic interactions. For the electrostatic [SS(V)PE] terms, we use Bondi's atomic van der Waals radii,[11] each scaled by a factor of 1.2. For the dispersion and repulsion interaction energies, a solvent-accessible surface was constructed by adding the Lennard–Jones minimum-energy distance for each solvent atom to the scaled Bondi radius. In either case, each atomic sphere was discretized using 110 Lebedev points.

References

[1] York, D. M.; Karplus, M. Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.

[2] Chipman, D. M. Reaction Field Treatment of Charge Penetration. *J. Chem. Phys.* **2000**, *112*, 5558–5565.

[3] Chipman, D. M. Comparison of Solvent Reaction Field Representations. *Theor. Chem. Acc.* **2002**, *107*, 80–89.

[4] Chipman, D. M.; Dupuis, M. Implementation of Solvent Reaction Fields for Electronic Structure. *Theor. Chem. Acc.* **2002**, *107*, 90–102.

[5] Su, P.; Li, H. Continuous and Smooth Potential Energy Surface for Conductor-Like Screening Solvation Model Using Fixed Points with Variable Areas. *J. Chem. Phys.* **2009**, *130*, 074109.

[6] Lebedev, V. I. Values of the nodes and weights of ninth to seventeenth order Gauss–Markov quadrature formulae invariant under the octahedron group with inversion. *Zh. Vychisl. Mat. Mat. Fiz.* **1975**, *1*, 48–54.

[7] Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. Ab Initio Study of Solvated Molecules: A New Implementation of the Polarizable Continuum Model. *Chem. Phys. Lett.* **1996**, *255*, 327–335.

[8] Cossi, M.; Mennucci, B.; Cammi, R. Analytical first derivatives of molecular surfaces with respect to nuclear coordinates. *J. Comput. Chem.* **1996**, *17*, 57–73.

[9] Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

[10] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Imprey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

[11] Bondi, A. Van der Waals Volumes and Radii. *J. Phys. Chem.*, **68**, 441–451 (1964).