

11-2009

# Smooth potential energy surface for cavitation, dispersion, and repulsion free energies in polarizable continuum model

Yali Wang

*University of Nebraska-Lincoln*, [yali.wang@huskers.unl.edu](mailto:yali.wang@huskers.unl.edu)

Hui Li

*University of Nebraska - Lincoln*, [hli4@unl.edu](mailto:hli4@unl.edu)

Follow this and additional works at: <http://digitalcommons.unl.edu/chemistryli>

 Part of the [Chemistry Commons](#)

---

Wang, Yali and Li, Hui, "Smooth potential energy surface for cavitation, dispersion, and repulsion free energies in polarizable continuum model" (2009). *Hui Li Publications*. 1.  
<http://digitalcommons.unl.edu/chemistryli/1>

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 Hui Li Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Smooth potential energy surface for cavitation, dispersion, and repulsion free energies in polarizable continuum model

Yali Wang and Hui Li<sup>(a)</sup>

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

(Received 13 July 2009; accepted 7 November 2009; published online 24 November 2009)

[doi:10.1063/1.3268921]

The fixed points with variable areas [FIXPVA (Ref. 1)] tessellation scheme is used to obtain smooth potential energy surfaces for the cavitation  $G_{\text{cav}}$ , dispersion  $G_{\text{dis}}$ , and repulsion  $G_{\text{rep}}$  free energies in the polarizable continuum model (PCM).<sup>2,3</sup> It is shown that FIXPVA can reproduce the standard GEPOL (Ref. 4) results to within 1 kcal/mol.

Here the method developed by Pierotti<sup>5</sup> and Langlet *et al.*<sup>6</sup> for evaluating  $G_{\text{cav}}$  in PCM is considered

$$G_{\text{cav}} = \sum_i \frac{a_i}{4\pi R_i^2} [K_0 + K_1(R_i + R_S) + K_2(R_i + R_S)^2], \quad (1)$$

where  $i$  runs over all the tesserae (solute surface elements),  $a_i$  is the area of tessera  $i$ ,  $R_i$  is the radius of the sphere associated with tessera  $i$ , and  $R_S$  is the assumed radius of the solvent molecule. The meaning of the parameters  $K_0$ ,  $K_1$ , and  $K_2$  can be found in the literature.<sup>6</sup> This method is based on the statistical scaled-particle theory originally developed by Reiss and Tully-Smith,<sup>7</sup> in which the surface tension term ( $K_0$ ) is the leading term.

The force field atomic interaction method developed by Floris *et al.*<sup>8</sup> for evaluating the dispersion and repulsion free energies in PCM is considered

$$G_{\text{dis}} = \rho \sum_N \sum_M \sum_i \left( -\frac{C_{MN}}{|\mathbf{r}_M - \mathbf{r}_i|^6} \cdot \frac{R_i^2 + |\mathbf{r}_M - \mathbf{r}_i|^2 - |\mathbf{r}_M - \mathbf{r}_i|^2}{2R_i} \cdot a_i \right), \quad (2)$$

$$G_{\text{rep}} = \rho \sum_N \sum_M \sum_i \left[ \left( \frac{1}{\alpha_{MN} |\mathbf{r}_M - \mathbf{r}_i|} + \frac{2}{\alpha_{MN}^2 |\mathbf{r}_M - \mathbf{r}_i|^2} + \frac{2}{\alpha_{MN}^3 |\mathbf{r}_M - \mathbf{r}_i|^3} \right) \cdot \beta_{MN} \exp(-\alpha_{MN} |\mathbf{r}_M - \mathbf{r}_i|) \cdot \frac{R_i^2 + |\mathbf{r}_M - \mathbf{r}_i|^2 - |\mathbf{r}_M - \mathbf{r}_i|^2}{2R_i} \cdot a_i \right], \quad (3)$$

where  $N$  runs over all solvent atoms,  $M$  runs over all solute atoms,  $i$  runs over all tesserae;  $\rho$  is the numerical density of the solvent;  $C_{MN}$  is the dispersion coefficient for solute atom  $M$  and solvent atom  $N$ ;  $\alpha_{MN}$  and  $\beta_{MN}$  are the two parameters in the exponential repulsion potential between solute atom  $M$  and solvent atom  $N$ ;  $\mathbf{r}_M$  and  $\mathbf{r}_i$  are the coordinates of solute atom  $M$  and tessera  $i$ ;  $R_i$  and  $\mathbf{r}_i$  are, respectively, the radius and center coordinate of the sphere associated with tessera  $i$ ; and  $a_i$  is the area of tessera  $i$ .

Equations (1)–(3) have been implemented in GAMESS (Ref. 9) using the GEPOL tessellation scheme by Tomasi *et al.*<sup>2,10</sup> for the dielectric PCM and integral equation formalism PCM methods. Due to the intrinsic discontinuity of the tessera coordinates as functions of molecular geometry in GEPOL, the corresponding potential energy surfaces are not smooth, and geometry optimization processes are often difficult to converge.

Using the FIXPVA tessellation scheme, smooth potential energy surfaces and analytic gradients for the electrostatic solvation free energy ( $G_{\text{ele}}$ ) in both conductorlike screening model and conductorlike (PCM), have been obtained.<sup>1,11</sup> In FIXPVA, the area of a tessera is scaled by switching functions of its distances to neighboring spheres. For the calculation of  $G_{\text{ele}}$ , values of 0.02, 0.3, 1.0, and 1.5 Å, respectively, were selected for the four parameters  $m_1$ ,  $m_2$ ,  $n_1$ , and  $n_2$  in the FIXPVA switching functions.<sup>1</sup> Using these parameters, FIXPVA produces solute surface roughly 10% less than the surface area computed from GEPOL. Fortunately, the  $G_{\text{ele}}$  is insensitive to the modifications of the tessera areas. For example, compared with GEPOL, FIXPVA produces solvation energy that is 1.1 kcal/mol smaller in magnitude for acetate anion, which has a solvation free energy around  $-80$  kcal/mol.<sup>1</sup>

However, the  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  in Eqs. (1)–(3) are sensitive to the surface areas. If the total surface area is reduced by  $\sim 10\%$ , these free energies will likely be reduced by  $\sim 10\%$ . In order to reproduce the GEPOL results, it is necessary to adjust the FIXPVA distance parameters  $m_1$ ,  $m_2$ ,  $n_1$ , and  $n_2$ . In addition, it is necessary to redefine the switching function variables  $m$  and  $n$  by using auxiliary spheres that represent neighboring spheres but with smaller radii, so that the area of a tessera on the boundary is scaled by  $\sim 0.5$ , similar to that in GEPOL. Based on extensive tests, it is found that for  $G_{\text{cav}}$  in Eq. (1), best results can be obtained by using 0.02, 0.3, 0.5, and 1.0 Å, respectively, for  $m_1$ ,  $m_2$ ,  $n_1$ , and  $n_2$ , and by using auxiliary spheres with radii reduced by 0.16 Å. For  $G_{\text{dis}}$  and  $G_{\text{rep}}$  in Eqs. (2) and (3), best results can be obtained by using 0.02, 0.3, 1.0, and 1.5 Å, respectively, for  $m_1$ ,  $m_2$ ,  $n_1$ , and  $n_2$ , and by using auxiliary spheres with radii reduced by 0.11 Å. Although these values were optimized to match the standard GEPOL results, they are largely determined by the physical fact that solvent molecules start to be excluded between two solute atoms when the space between them is about 1.0–1.5 Å. This modified FIXPVA scheme was implemented in GAMESS for both energy and analytic gradi-

TABLE I. Cavitation, dispersion, and repulsion free energies (kcal/mol) calculated for 11 molecules.

	$G_{\text{cav}}$		$G_{\text{dis}}$		$G_{\text{rep}}$	
	GEPOL	FIXPVA	GEPOL	FIXPVA	GEPOL	FIXPVA
2,5-diketopiperazine-3-acetate, $\text{C}_6\text{H}_7\text{O}_4\text{N}_2^-$	20.56	20.51	-22.51	-22.32	4.61	4.63
$\text{C}_6\text{H}_5\text{COO}^-$	16.19	16.25	-18.49	-19.24	3.89	4.04
$\text{CH}_3\text{NHCOCH}_2\text{COO}^-$	15.76	15.74	-17.62	-17.19	3.91	3.86
Pyrrolidine-2-formate, $\text{C}_4\text{H}_8\text{NHCOO}^-$	15.24	15.16	-19.49	-19.62	4.83	4.95
$\text{NH}_2\text{COCH}_2\text{COO}^-$	13.92	13.91	-14.93	-14.94	2.97	3.00
$\text{C}_6\text{H}_6$	11.87	12.08	-16.34	-16.90	4.26	4.45
1-H-imidazole, $\text{C}_3\text{H}_4\text{N}_2$	10.05	10.15	-12.99	-13.54	3.22	3.31
$\text{CH}_3\text{COO}^-$	9.46	9.47	-10.82	-10.65	2.63	2.58
$\text{CH}_3\text{S}^-$	8.32	8.27	-8.10	-7.94	2.12	2.11
$\text{CH}_3\text{O}^-$	6.21	6.25	-8.35	-8.42	2.68	2.71
$\text{CH}_3\text{NH}_2$	6.60	6.55	-10.31	-10.43	3.38	3.46
Maximum unsigned deviation		0.21		0.75		0.19
RMSD		0.08		0.37		0.09

ent calculations for the  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  terms in Eqs. (1)–(3).

In the CPCM calculations discussed below, spheres with radii of 0.00, 1.77, 1.68, 1.59, and 2.10 Å were input for H, C, N, O, and S atoms, respectively. For  $G_{\text{ele}}$  calculation the input radii were scaled by 1.2. For  $G_{\text{cav}}$  calculation, they were used as input. For  $G_{\text{dis}}$  and  $G_{\text{rep}}$  calculations, the atomic radii of the solvent (in this work, H and O for water) were added to these input atomic radii to define the solute surface. The solvent was water with  $\epsilon=78.39$ , and the default parameters for  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  for water were used.

The GEPOL and FIXPVA tessellations are not rotationally invariant. Acetate in 20 random orientations was used to test the rotational variances of the computed  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$ . Using GEPOL and 60 initial tesserae per sphere, the maximum rotational variances observed for  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  energies are all less than 0.1 kcal/mol. Using FIXPVA and 60 initial tesserae per sphere, the maximum variances observed for  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  are 0.83, 1.12, and 0.25 kcal/mol, respectively. Using FIXPVA and 240 initial tesserae per sphere, the maximum variances are 0.21, 0.42, and 0.08 kcal/mol, respectively. Therefore, 240 initial tesserae per sphere are recommended for general use.

Table I presents  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  calculated with GEPOL scheme (60 initial tesserae per sphere) and FIXPVA scheme (240 initial tesserae per sphere) for 11 molecules. For  $G_{\text{cav}}$ , the maximum unsigned deviation is 0.21 kcal/mol and root-mean-square deviation (RMSD) is 0.08 kcal/mol, with most of the molecules showing deviations smaller than 0.1 kcal/mol. The maximum unsigned deviations of the FIXPVA  $G_{\text{dis}}$  and  $G_{\text{rep}}$  are 0.75 and 0.19 kcal/mol, respectively, from the GEPOL values, with RMSDs being 0.37 and 0.09 kcal/mol, respectively. Because the differences are caused mainly by the geometric differences in GEPOL and FIXPVA tessellation, the tests performed with so many different molecules are sufficient.

2,5-diketopiperazine-3-acetate anion ( $\text{C}_6\text{H}_7\text{O}_4\text{N}_2^-$ ) was optimized with the CPCM/RHF/6-31G\* method (FIXPVA tessellation for  $G_{\text{ele}}$ ), in which  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  were included using either GEPOL or FIXPVA. Since analytic GE-

POL gradients for  $G_{\text{cav}}$ ,  $G_{\text{dis}}$ , and  $G_{\text{rep}}$  were not available in GAMESS, numerical gradients were obtained via single displacement with a step size of  $10^{-6}$  a.u. Using GEPOL with 60 initial tesserae per sphere, both the energy and root-mean square gradients (RMSG) in the geometry optimization process fluctuate, with no sign of convergence. In fact the geometry optimization could not converge in 250 steps to the criteria that the maximum gradient be smaller than  $5 \times 10^{-5}$  a.u. and the RMSG be smaller than  $1.67 \times 10^{-5}$  a.u. Using FIXPVA with 240 initial tesserae per sphere, the maximum gradient and RMSG smoothly decreases and fall below the convergence criterion at the 36th step.

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

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: hli4@unl.edu.

<sup>1</sup> P. Su and H. Li, *J. Chem. Phys.* **130**, 074109 (2009).

<sup>2</sup> S. Miertès, E. Scrocco, and J. Tomasi, *Chem. Phys.* **55**, 117 (1981); E. Cancès, B. Mennucci, and J. Tomasi, *J. Chem. Phys.* **107**, 3032 (1997); B. Mennucci, E. Cancès, and J. Tomasi, *J. Phys. Chem. B* **101**, 10506 (1997).

<sup>3</sup> V. Barone and M. Cossi, *J. Phys. Chem. A* **102**, 1995 (1998); M. Cossi, N. Rega, G. Scalmani, and V. Barone, *J. Comput. Chem.* **24**, 669 (2003); H. Li and J. H. Jensen, *ibid.* **25**, 1449 (2004).

<sup>4</sup> J. L. Pascual-Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, *J. Comput. Chem.* **8**, 778 (1987); J. L. Pascual-Ahuir, E. Silla, and I. Tunon, *ibid.* **15**, 1127 (1994).

<sup>5</sup> R. A. Pierotti, *Chem. Rev. (Washington, D.C.)* **76**, 717 (1976).

<sup>6</sup> J. Langlet, P. Claverie, J. Caillet, and A. Pullman, *J. Phys. Chem.* **92**, 1617 (1988).

<sup>7</sup> H. Reiss and D. M. Tully-Smith, *J. Chem. Phys.* **55**, 1674 (1971).

<sup>8</sup> F. Floris and J. Tomasi, *J. Comput. Chem.* **10**, 616 (1989); F. M. Floris, J. Tomasi, and J. L. Pascual Ahuir, *ibid.* **12**, 784 (1991).

<sup>9</sup> M. 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).

<sup>10</sup> J. Tomasi, B. Mennucci, and E. Cancès, *J. Mol. Struct.: THEOCHEM* **464**, 211 (1999).

<sup>11</sup> D. Si and H. Li, *J. Chem. Phys.* **131**, 044123 (2009).