Variational approach for nonpolar solvation analysis

Zhan Chen,[1,a)] Shan Zhao,[2] Jaehun Chun,[3] Dennis G. Thomas,[3] Nathan A. Baker,[3] Peter W. Bates,[1] and G. W. Wei[1,4,b)]
[1]*Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, USA*
[2]*Department of Mathematics, University of Alabama, Tuscaloosa, Alabama 35487, USA*
[3]*Pacific Northwest National Laboratory, Richland, Washington 99352, USA*
[4]*Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan 48824, USA*

(Received 11 May 2012; accepted 27 July 2012; published online 23 August 2012)

Solvation analysis is one of the most important tasks in chemical and biological modeling. Implicit solvent models are some of the most popular approaches. However, commonly used implicit solvent models rely on unphysical definitions of solvent-solute boundaries. Based on differential geometry, the present work defines the solvent-solute boundary via the variation of the nonpolar solvation free energy. The solvation free energy functional of the system is constructed based on a continuum description of the solvent and the discrete description of the solute, which are dynamically coupled by the solvent-solute boundaries via van der Waals interactions. The first variation of the energy functional gives rise to the governing Laplace-Beltrami equation. The present model predictions of the nonpolar solvation energies are in an excellent agreement with experimental data, which supports the validity of the proposed nonpolar solvation model. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4745084]

I. INTRODUCTION

Under physiological conditions, water constitutes 65%–90% of cellular mass. As such, essentially all important biological processes, including signal transduction, transcription, and translation, occur in an aqueous environment. Therefore, an elementary requirement for quantitative modeling and analysis of biological processes is a detailed understanding of the solvation process, in which solute molecules are transferred from their lowest energy state in vacuum to an equilibrium state in a solvent environment.[1–4] The solvation process involves the work of inserting a molecule into the solvent and a range of possible solvent-solute interactions at the interface, including hydrogen bonding, ion-ion, ion-dipole, dipole-dipole and multipole attractions, Debye attractions, and London dispersion attractions. These interactions induce structural reorganization of the solvent near the interface as well as possible solute configurational changes.

Experimental techniques, such as neutron diffraction with isotopic substitution and anomalous x-ray diffraction, have been applied to determine the solvent atomic distribution in the vicinity of the interface. A vast range of computational approaches has also been developed to predict the solvent microstructures. In molecular mechanics, the solvent can be described with a variety of sites, different levels of bond flexibilities, different orders of multipole expansions, and many quantum mechanical treatments.[5–9] For example, solvent microstructure can be extracted from molecular dynamics simulations and described by density distribution and correlation functions.[10–12] These functions can also be described by the density functional theory[13] and/or integral equations.[14,15] In general, both molecular mechanical and statistical mechanical based microstructural theories are capable of predicting solvent radial distribution functions, surface tension, temperature, and pressure.[16–19]

The calculation of solvation free energy has captured a great deal of interest with research developing solvation models ranging from simple phenomenological modifications of Coulomb's law, implicit solvent models that describe the solvent by mean-field approximations,[20–24] explicit solvent models that treat the solvent in molecular or atomic detail,[1] to complex quantum mechanical methods.[2,25–27] Each of these models has its own domain of applicability, merits, and limitations. Explicit and quantum methods are ideal for studying the solvation of relatively small molecules, but involve an excessively large number of degrees of freedom for large molecules, which can become prohibitively expensive. Implicit solvent methods and, especially multiscale methods which are able to reduce the number of degrees of freedom, are indispensable in solvation analysis and the quantitative description of other biological processes in general.

An essential element in all implicit solvent models is the description of the interface that separates the solute from the solvent.[28] Many solvation quantities of physical interest, including electrostatic free energies, surface areas, cavitation volumes, solvation free energies, and pK_a values, are very sensitive to the interface definition. Currently, a number of different surface definitions, including the van der Waals (vdW) surface, the solvent accessible surface,[29] the solvent excluded surface,[30] and electron density isosurface are used in implicit solvent analysis, and have had much success in biomolecular modeling,[31–38] including

a)Current address: School of Mathematics, University of Minnesota–Twin Cities, Minneapolis, Minnesota 55455, USA.
b)Author to whom correspondence should be addressed. Electronic mail: wei@math.msu.edu.

protein folding,[31,32] protein-protein interfaces,[33] protein surface topography,[34] oral drug classification,[35] DNA binding and bending,[36] macromolecular docking,[37] and enzyme catalysis.[38] However, these surface definitions are *ad hoc* partitions to separate solute atoms from the surrounding solvent, and often create geometric singularities.[39] Currently, there is no consensus about which surface should be used to describe the solvent-solute boundary in implicit solvent models.

In the past few years, a series of efforts has been taken to improve the surface description of implicit solvent models. The first effort was the introduction of curvature driven partial differential equations (PDEs), i.e., geometric flows, to construct singularity free protein molecular surfaces in 2005.[40] This has led to the recent surge in PDE based approaches for molecular surface analysis. The next step was the introduction of the minimal molecular surface (MMS), obtained by using variational principles and the mathematical theory of geometric flows, namely, the mean curvature flow.[41,42] The MMSs are able to minimize the surface free energy of a macromolecule in solution. A natural step to improve the minimal molecular surface model was to consider more solvent-solute interactions in the surface construction by using potential-driven geometric flows.[43] Recently, differential geometry based multiscale models have been introduced for the analysis of structure, dynamics, and transport of complex chemical and biological systems.[44] An essential ingredient of such models is the use of the differential geometry of surfaces as a natural means to separate the continuum solvent domain from the atomistic solute domain, while dynamically coupling discrete descriptions and continuum descriptions. The main strategy is to construct a total energy functional of the system to encompass the polar and nonpolar free energies of solvation, and other energies of interest. Intensive investigation has been carried out to practically implement differential geometry based solvation models in the Eulerian representation,[45] Lagrangian representation,[46] and quantum formulation.[47] These models incorporate solvent-solute van der Waals interactions and thus, partially take care of solvent size effects near the interface. They have been extensively validated with experimental data of solvation free energies.[45–47] While these solvation models are similar to those of Dzubiella *et al.* in spirit,[48,49] the treatment of surface tension is different in the two approaches. Most recently, generalized correlations have been considered in differential geometry based multiscale models to account for not only solvent-solute interactions, but also ion-ion and ion-solute interactions.[50] These multiscale models, in principle, should be able to capture the solvent microstructure. However, these differential geometry based multiscale models involve polar and nonpolar solvation components, and chemical potential type of energies, which can contribute to uncertainty in the model validation and quantification.

The objective of the present work is to minimize modeling uncertainty in the differential geometry based multiscale models by considering a relatively isolated situation where electrostatic interactions do not play an important role. A specific physical process is the solvation of nonpolar molecules, in which the electrostatic interactions between solvent and solute are negligible. This simplified situation constitutes a test of whether the differential geometry based framework provides a viable description of solvation free energies. Note that in the present approach, while an atomistic description is utilized for the solute, the solvent domain is treated as a continuum.

The rest of this paper is organized as follows. Section II is devoted to the formulation of the present differential geometry based nonpolar solvation model. We present a variational framework for the solvation free energy contributions. The governing equation is derived by the Euler-Lagrange variation. The solution of the governing equation gives rise to the solvent-solute boundary, which facilitates solvation free energy calculation. We design schemes of second-order numerical accuracy for the construction and evolution of solute characteristic function. Appropriate iterative procedures are provided to ensure convergence of the solution. Section III presents validation and analysis of the proposed new solvation model. The applications of the proposed theories, methods and algorithms are considered to two sets of compounds: alkanes and alkenes. Comparison is given to experimental measurements and results in the literature. Finally, this paper ends with a conclusion.

II. THEORY AND ALGORITHM

A. Theoretical model

In this work, we investigate the physical boundary between solvent and solute by considering a realistic problem with less uncertainty. Specifically, we apply our variational approach to apolar molecules that admit negligible polar solvation effects. Nonpolar solvation processes are generally associated with the insertion of an uncharged solute into solvent. There are many nonpolar solvation models available. The most commonly used model is the scaled particle theory (SPT) (Ref. 22) which includes the energy of the surface tension effect and the mechanical work of immersing a particle into the solvent. Recent work by Levy, Gallicchio, and others[3,4,51] has demonstrated the importance of attractive solute-solvent terms as well as models of solvent-solvent repulsive interactions, in addition to both area and volume contributions.[51] In the present work, we employ the following model for nonpolar solvation free energies:[51]

$$G_{\rm np} = \gamma(\text{Area}) + p(\text{Vol}) + \int_{\Omega_s} \rho_s U^{\rm vdW}(\mathbf{r}) d\mathbf{r}, \qquad (1)$$

where γ is the surface tension, Area is the solvent-excluded surface area of the solute, p is the hydrodynamic pressure, Vol is the solvent-excluded volume of the solute, ρ_s is the solvent density, Ω_s denotes the solvent accessible region, and $U^{\rm vdW}(\mathbf{r})$ is the solvent-solute vdW interaction potential. The first two terms in Eq. (1) are those from the SPT.[22] Since the division of polar and nonpolar solvation energies is somewhat arbitrary, the polar contribution might not be completely excluded in real situations. Nevertheless, the attractive van der Waals interactions can offset possible small polar contributions. This nonpolar solvation model has been shown to provide a good agreement for the solvation forces of proteins[51] and RNA hairpins[23] with explicit solvent models. Levy and

co-workers have demonstrated good performance of a similar nonpolar model.[3,4,52,53]

The free energy functional (1) provides a method for practical nonpolar solvation analysis. Moreover, a variational principle based on (1) leads to a potential-driven geometric flow equation for the surface evolution[46]

$$\frac{\partial \mathbf{X}}{\partial t} = -W_n \mathbf{N}, \tag{2}$$

where $\mathbf{X} \in \Gamma \subset \mathbb{R}^3$ is a position vector on the evolving manifold Γ, $\mathbf{N}$ is the outward unit normal direction at $\mathbf{X}$, and the potential $W_n \equiv -2\gamma H + p - \rho_s U^{\text{vdW}}$ vanishes upon optimization of the free energy functional (1) with respect to the solvent-solute interface Γ. Here, H is the mean curvature of Γ. The generalized geometric flow equation (2) is in the Lagrangian formulation. The numerical solution of (2) involves the evolution of a triangularization mesh representing the solvent-solute interface. The advantage of such a Lagrangian approach is that it evolves only a two-dimensional surface. The Lagrangian approach encounters difficulties in handling topological changes, such as surface breaking or merging, which commonly occur in biomolecular surface constructions and molecular dynamics applications. However, one can overcome these obstacles by using the Eulerian formulation,[40–43] in which the surface is embedded into a higher dimensional space as a level set of higher dimensional surface. Then the topological changes can be easily handled.

B. Governing equations in Lagrangian and Euler formulations

In the Lagrangian formulation, the potential driven geometric flow equation of the solvent-solute interface is given by Eq. (2). The total free energy is known to be decreasing when the surface is evolved by (2) (Ref. 46) and the steady state solution of Eq. (2) yields a solvent-solute interface with $W_n = 0$.

In the Eulerian formulation, the interfacial surface is embedded in a hypersurface and the latter is evolved under prescribed driving forces. Consider a hypersurface function $S(\mathbf{r})$ with $\mathbf{r} \in \mathbb{R}^3$. Then the desired surface can be represented as a set of points with a constant value of function S,

$$\Gamma = \{\mathbf{r} | S(\mathbf{r}) = L\}, \tag{3}$$

where L is an isosurface value. On an isosurface, ∇S vanishes along two tangential directions. Thus, the normal direction at $\mathbf{r}$ is given as

$$\mathbf{N} = \frac{\nabla S}{\|\nabla S\|}. \tag{4}$$

Moreover, the explicit form of mean curvature can be obtained according to the equality $2H = \nabla \cdot \mathbf{N}$,

$$H = \frac{1}{2} \nabla \cdot \left(\frac{\nabla S}{\|\nabla S\|} \right). \tag{5}$$

By the Chain Rule

$$\frac{\partial S}{\partial t} = \frac{\partial S}{\partial \mathbf{X}} \cdot \frac{\partial \mathbf{X}}{\partial t} = \nabla S \cdot \frac{\partial \mathbf{X}}{\partial t} = -W_n \nabla S \cdot \mathbf{N}. \tag{6}$$

Substituting from (4), one has

$$\frac{\partial S}{\partial t} = -\|\nabla S\| W_n = \|\nabla S\| [2\gamma H - p + \rho_s U^{\text{vdW}}]. \tag{7}$$

Finally, by using (5), we arrive at the Eulerian form of the proposed potential driven geometric flow equation for the solute-solvent interface

$$\frac{\partial S}{\partial t} = \|\nabla S\| \left[\gamma \nabla \cdot \left(\frac{\nabla S}{\|\nabla S\|} \right) - p + \rho_s U^{\text{vdW}} \right]. \tag{8}$$

The vdW potential U^{vdW} is computed pairwise for each atom $U^{\text{vdW}}(\mathbf{r}) = \sum_i V_i^{\text{vdW}}(\mathbf{r})$, where $V_i^{\text{vdW}}(|\mathbf{r} - \mathbf{r}_i|) = \epsilon_i [(\frac{\sigma_i + \sigma_s}{|\mathbf{r} - \mathbf{r}_i|})^{12} - 2(\frac{\sigma_i + \sigma_s}{|\mathbf{r} - \mathbf{r}_i|})^6]$ is the standard 12-6 Lennard-Jones potential for the ith particle of the solute at $\mathbf{r}_i$ and the solvent at position $\mathbf{r}$. Here, σ_s represents the solvent radius, σ_i is the radius of the ith particle of the solute, and ϵ_i the well-depth. In this work, $V_i^{\text{vdW}}(\mathbf{r})$ is taken as the attractive part of the Lennard-Jones potential according to the Weeks-Chandler-Anderson (WCA) theory.[45,46,54] σ_i is taken as the solute atomic radius and σ_s is fixed to be 0.65 Å as a calibrated solvent radius.[51] To determine all the well-depth parameters in U^{vdW}, we assume that the Lennard-Jones potential is a constant when sampling on the vdW surface of the ith atom.[45,46] Therefore, we set $V_i^{\text{vdW}}(\sigma_i) = \epsilon_i [(\frac{\sigma_i + \sigma_s}{\sigma_i})^{12} - 2(\frac{\sigma_i + \sigma_s}{\sigma_i})^6] = V_1^{\text{vdW}}(\sigma_1)$. The atom-dependent well depth parameters ϵ_i are determined if $V_1^{\text{vdW}}(\sigma_1)$, which is used as a fitting parameter, is known. In general, each $V_i^{\text{vdW}}(\sigma_i)$ should admit a different value for each type of atom, however, we feel that the loss of generality in taking them equal is small while computational savings are great.

The computation of the proposed nonpolar solvation model is carried out in two steps. At the first step, the potential driven geometric flow equation (7) is numerically solved by using the finite difference method. A brief description of this solution procedure is given here, and we refer to our earlier works[42,45] for more details. We first rewrite equation (8) as

$$\frac{\partial S}{\partial t'} = \sqrt{1 + \|\nabla S\|^2} \left[\nabla \cdot \left(\frac{\nabla S}{\sqrt{1 + \|\nabla S\|^2}} \right) - \frac{p}{\gamma} + \frac{\rho_s}{\gamma} U^{\text{vdW}} \right], \tag{9}$$

where $t' = t\gamma$. Two modifications have been conducted. First, to avoid a vanishing value in the denominator, $\|\nabla S\|$ is replaced by $\sqrt{1 + \|\nabla S\|^2}$. Such a replacement will not affect the optimized solvent-solute interface, because $\|\nabla S\|^2$ is much larger than one near the interface. Second, the temporal variable is scaled by γ. We note that the free parameters of this equation are essentially the surface tension γ, hydrodynamic pressure constant p, and solvent density ρ_s.

Consider a macromolecule with total N_a number of atoms. All geometric structures of compounds considered in this work are taken from the pubchem web (http://pubchem.ncbi.nlm.nih.gov). vdW radii are set as 1.87 Å, which is ZAP-9 radius,[55] for the carbon atom and 1.10 Å for the hydrogen atom. We denote the center and radius of the ith atom to be $\mathbf{r}_i = (x_i, y_i, z_i)$ and r_i, respectively, for $i = 1, 2, \ldots, N_a$. We then define the domain enclosed by the solvent accessible surface to be $D = \bigcup_{i=1}^{N_a} \{\mathbf{r} : |\mathbf{r} - \mathbf{r}_i|$

$< r_i + r_p\}$, where r_p is the probe radius. For the initial value of S, we consider an indicator function

$$S(x, y, z, 0) = \begin{cases} S_0, & (x, y, z) \in D \\ 0, & \text{otherwise} \end{cases}, \tag{10}$$

where $S_0 = 1000$. The explicit Euler method and the second-order central difference scheme are used for temporal and spatial discretization, respectively. We only numerically update the values of $S(x, y, z, t)$ at the points in between the vdW surface and solvent accessible surface, i.e., $(x, y, z) \in \bigcup_{i=1}^{N_a}\{\mathbf{r} : r_i < |\mathbf{r} - \mathbf{r}_i| < (r_i + r_p)\}$. This is for the purpose of protecting the van der Waals surface and making the computation more efficient. After the steady state solution is numerically reached, the solvent-solute interface Γ can be represented as a set of points with a constant value of function S, $\Gamma = \{\mathbf{r}|S(\mathbf{r}) = L\}$. Here, the isosurface value is chosen to be $L = \frac{S_0}{2}$.

In the second step, we calculate the nonpolar solvation free energy based on the solvent-solute interface Γ. For this purpose, we need to compute surface integrals and volume integrals over the solute domain Ω_m, which are defined originally in the Lagrangian formulation. However, the hypersurface function S is defined on a three-dimensional (3D) Cartesian grid in the Eulerian formulation. Consequently, the isosurface Γ is also represented based on the 3D Cartesian grid. Therefore, great care has to be exercised when calculating these integrals.

C. Surface extraction

A stand-alone algorithm based on the marching cubes method[56] is constructed. Consider a grid point (x_i, y_j, z_k) on the 3D Cartesian grid and denote $S_{i,j,k} = S(x_i, y_j, z_k)$. We first compute a normal vector $\mathbf{n} = (n_x, n_y, n_z)$, according to S values nearby,

$$\begin{aligned} n_x(x_i, y_j, z_k) &= \frac{S_{i+1,j,k} - S_{i-1,j,k}}{2\Delta x}, \\ n_y(x_i, y_j, z_k) &= \frac{S_{i,j+1,k} - S_{i,j-1,k}}{2\Delta y}, \\ n_z(x_i, y_j, z_k) &= \frac{S_{i,j,k+1} - S_{i,j,k-1}}{2\Delta z}. \end{aligned} \tag{11}$$

The surface integral of a density function f can be approximated by[57]

$$\begin{aligned} \int_\Gamma f(x, y, z)d\sigma &= \int_\Omega f(x, y, z)\delta(d)d\mathbf{r} \\ &\approx \sum_{i,j,k} f(x_i, y_j, z_k)\tilde{\delta}_{i,j,k}h^3, \end{aligned} \tag{12}$$

where $\delta(d)$ is a delta function and d is distance of a point (x, y, z) defined in Ω from the interface Γ. By means of the delta function, the surface integral becomes a volume integral in the entire domain, and the latter is numerically approximated based on discrete function values $f(x_i, y_j, z_k)$ and a 3D discrete delta function $\tilde{\delta}_{i,j,k}$. Here, for simplicity, we assume a uniform mesh size h along x, y and z directions. Generalization to a nonuniform mesh is possible. Following Smereka, the 3D discrete delta function $\tilde{\delta}_{i,j,k}$ is taken to be

$$\tilde{\delta}_{i,j,k} = \tilde{\delta}^{(+x)}_{i,j,k} + \tilde{\delta}^{(-x)}_{i,j,k} + \tilde{\delta}^{(+y)}_{i,j,k} + \tilde{\delta}^{(-y)}_{i,j,k} + \tilde{\delta}^{(+z)}_{i,j,k} + \tilde{\delta}^{(-z)}_{i,j,k}, \tag{13}$$

where $\tilde{\delta}^{(\pm\alpha)}_{i,j,k}, (\alpha = x, y, z)$ are 1D discrete delta functions.[57] For a node (x_i, y_j, z_k), these six 1D delta functions are never present simultaneously. For a reasonable mesh resolution, at most three of them could be seen.

We restrict our following discussions from one side of interface, say inside the interface. Consider all grid points which are either on the interface, or are inside the interface but whose distance away from the interface is less than h. Denote the set of such grid points to be I. To carry out integration exactly on the interface, we use the following discrete surface integration formula:[57]

$$\begin{aligned} \int_\Gamma f(x, y, z)d\sigma \approx \sum_{(i,j,k)\in I} \Bigg(& f(x_o, y_j, z_k)\frac{|n_x|}{h} \\ & + f(x_i, y_o, z_k)\frac{|n_y|}{h} + f(x_i, y_j, z_o)\frac{|n_z|}{h} \Bigg) h^3, \end{aligned} \tag{14}$$

where (x_o, y_j, z_k) is the intersecting point of the interface and the x meshline that passes through (i, j, k), and n_x is the x component of the unit normal vector at (x_o, y_j, z_k). Similar relations exist between (x_i, y_o, z_k) and n_y, and (x_i, y_j, z_o) and n_z. The interface locations and normal directions are calculated by linear interpolation. For instant, if $(x_i, y_j, z_k) \in I$ and (x_{i+1}, y_j, z_k) is outside the interface, we calculate a scaled distance from the interface to be

$$d = \frac{S_{i,j,k} - L}{S_{i,j,k} - S_{i+1,j,k}}. \tag{15}$$

Then the function value and the normal direction can be computed as

$$\begin{aligned} f(x_o, y_j, z_k) &= (1-d)f(x_i, y_j, z_k) + df(x_{i+1}, y_j, z_k), \\ \mathbf{n}_o &= (1-d)\mathbf{n}(x_i, y_j, z_k) + d\mathbf{n}(x_{i+1}, y_j, z_k). \end{aligned} \tag{16}$$

The unit normal $\mathbf{N}_o$ at the intersecting point can be easily computed as $\mathbf{N}_o = \frac{\mathbf{n}_o}{\|\mathbf{n}_o\|}$, whose x component is n_x in (14). Other terms in (14) can be similarly approximated.

The surface area can be calculated by setting $f = 1$ in Eq. (14). The error of the surface integration depends on the grid resolution and was observed to be approximately second-order in h.[57] The volume integral of the density function f can be simply approximated by[57]

$$\begin{aligned} \int_{\Omega_m} f(x, y, z)d\mathbf{r} \approx \Bigg(& \sum_{(i,j,k)\in J_1} f(x_i, y_j, z_k) \\ & + \frac{1}{2}\sum_{(i,j,k)\in J_2} f(x_i, y_j, z_k) \Bigg) h^3, \end{aligned} \tag{17}$$

where J_1 is the set of grid points inside/on the surface, and J_2 contains the points outside the surface while adjacent to at least one point belonging to J_1.

TABLE I. Decomposition analysis for the calibration set of 11 alkanes. Comparisons between our variational approach and the explicit solvent model in Ref. 4 with respect to repulsive and attractive parts and total solvation free energies are conducted. The error in total free energy is calculated according to the experimental data.[58]

	Rep. part (kcal/mol)		Att. part (kcal/mol)		Total (kcal/mol)		Error (kcal/mol)	
Compound	Present	Explicit	Present	Explicit	Present	Explicit	Present	Explicit
Methane	4.71	5.72	− 2.73	− 3.31	1.98	2.41	− 0.02	0.41
Ethane	6.65	8.07	− 4.75	− 5.44	1.90	2.63	0.07	0.80
Butane	10.30	10.10	− 8.18	− 7.21	2.12	2.89	0.04	0.81
Propane	8.50	12.19	− 6.45	− 8.98	2.04	3.21	0.08	1.25
Pentane	12.19	14.22	− 9.82	− 10.77	2.37	3.45	0.04	1.12
Hexane	14.03	16.17	− 11.54	− 12.38	2.50	3.78	0.01	1.30
Isobutane	10.14	11.91	− 7.97	− 8.88	2.16	3.03	− 0.36	0.51
2-methylbutane	11.73	13.64	− 9.35	− 10.13	2.38	3.51	0.00	1.13
Neopentane	11.81	13.62	− 9.20	− 10.39	2.61	3.23	0.11	0.73
Cyclopentane	10.60	12.79	− 9.43	− 9.99	1.17	2.80	− 0.03	1.60
Cyclohexane	12.05	14.00	− 10.78	− 11.66	1.27	2.34	0.04	1.11

D. Iteration procedure

Essentially, the model involves three optimization parameters: the surface tension γ, hydrodynamic pressure p and the Lennard-Jones potential parameter $V_1^{\mathrm{vdW}}(\sigma_1)$. An iterative procedure is used to optimize three model parameters γ, p, and $V_1^{\mathrm{vdW}}(\sigma_1)$:

(1) Choose a trial set of molecules with given atomic coordinates, radii, and experimental data of solvation free energies. For each molecule, take an initial set of parameters γ, p, and $V_1^{\mathrm{vdW}}(\sigma_1)$.
(2) For jth molecule, solve Eq. (2) to a steady state to compute surface area, molecular volume, and solvation free energy G_{np}^j.
(3) Set up a target function

$$T\left(\gamma, p, V_1^{\mathrm{vdW}}(\sigma_1)\right) = \sum_j \left(G_{\mathrm{np}}^j - G_{\mathrm{np}}^{j,\mathrm{exp}}\right)^2,$$

where $G_{\mathrm{np}}^{j,\mathrm{exp}}$ are experimental data of solvation free energies.
(4) Parameters p, γ, and $V_1^{\mathrm{vdW}}(\sigma_1)$ are updated by resolving a least-squares problem to determine non-negative parameters. The iterative procedure continues until the target function T reaches within a pre-set tolerance for the above three parameters.

Numerically, we have found that the above iterative procedure is not sensitive to the initial values of γ, p, and $V_1^{\mathrm{vdW}}(\sigma_1)$. The same procedure can be applied to allow each $V_i^{\mathrm{vdW}}(\sigma_i)$ to vary independently, which may offer better results. However, it would require a large training set to determine many extra parameters in the nonpolar model.

III. RESULTS AND DISCUSSIONS

We validate the proposed nonpolar variational approach by conducting both numerical calibration and prediction studies for 11 alkanes, see Table I. All geometric structures of compounds considered in this work are taken from the pubchem web (http://pubchem.ncbi.nlm.nih.gov). The solvent density is set to $\rho_s = 0.033428$ Å^{-3}.[51] The potential driven geometric flow equation is numerically solved by using the explicit Euler method for time and the second-order central finite difference scheme for space. In all calculations, we choose a uniform grid spacing $h = 0.25$ Å in all Cartesian directions and a time stepping $\tau = h^2/4.5$. Based on the resulting surface Γ, the solvation free energy G_{np} can finally be determined by calculating volume and surface integrals according to (1).

Based on different initial values, the optimized fitting parameters for alkanes are consistently found to be $\gamma = 0.0806$ kcal/(mol Å^2), $p = 0.0191$ kcal/(mol Å^3), and $V_1^{\mathrm{vdW}}(\sigma_1) = 11.272$ kcal/mol. Note that when $V_1^{\mathrm{vdW}}(\sigma_1) = 11.272$ kcal/mol the depth of potential well ϵ_H of hydrogen is 0.046 kcal/mol, which is of the same order of magnitude as that used in the literature.[59,60]

In the calculation of alkenes, we take parameters p and $V_1^{\mathrm{vdW}}(\sigma_1)$ values from the alkane calculation by assuming the same solvent behavior. Therefore, only the surface tension γ is regarded as a global fitting parameter here. The parameter optimization based on seven alkene compounds converges to $\gamma = 0.0775$ kcal/(mol Å^2).

An example plot of the surface evolution for a nonpolar compound is illustrated in Fig. 1. It is seen that the final surface is smooth and free of geometric singularity.

The solvent-solute dispersion term is crucial in the present nonpolar model. The Lennard-Jones potential according to the WCA theory[46,54] is used for the vdW potential. In Fig. 2, vdW potentials $U^{\mathrm{vdW}}(\mathbf{r})$ are plotted on the final surfaces for three nonpolar compounds. Clearly, surface vdW potentials differ much from surface electrostatic potentials computed by the Poisson model.

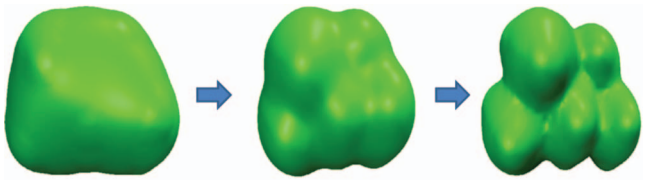

FIG. 1. The schematic plot of the surface evolution of a nonpolar compound (2,2,4-trimethylpentane) with respect to the time.

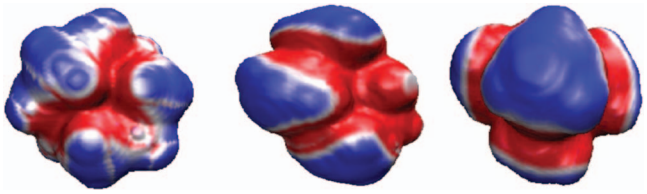

FIG. 2. The final isosurfaces of four nonpolar compounds projected with the corresponding vdW potential. Involved compounds are 1,1-diethoxyethane, bis-2-chloroethyl ether and dimethoxymethane, from left to right.

To benchmark our model, we consider nonpolar molecules which have high quality experimental data of solvation free energies. We first consider a calibration study of alkane compounds. A set of 11 small alkane compounds is chosen which includes linear, branched, and cyclic apolar compounds so that subtle differences are expected in the solvation free energy for the compounds with the same molecular formula. This test set has been frequently used in the literature for the hydrophobic solute test[4,51] because experimental data is available.[58] Moreover, the work of Gallicchio *et al.*[4] already gave a detailed decomposition of hydrophobic free energies into attractive and repulsive components by using an explicit solvent model and an OPLS all-atom force field. This allows a detailed comparison of our variational approach with other computational models. Using the optimized parameters, the results of 11 alkane compounds obtained from the present variational approach are shown in top section of Table II. It is evident that our model reproduces the total solvation free energies of 11 alkanes very well. The present model accurately catches the subtle differences between linear, branched, and cyclic apolar compounds. The root mean square (RMS) error is found to be as small as 0.12 kcal/mol.

For a comparison, the results reported by Gallicchio *et al.*[4] and Wagoner and Baker[51] are also depicted in Fig. 3. Clearly, their models overestimate the solvation free energies of 11 alkane compounds. What is considered as a difficult task is the prediction of the check mark shape of the solvation free energies of the first six alkanes, i.e., a drop of the solvation free energy at ethane C_2H_6 before a monotonic increase of the solvation free energies as the alkanes increase in mass. It is noteworthy that the present

TABLE II. Numerical and experimental total solvation free energies for 30 alkane compounds. The numerical energy is the sum of the repulsive part and attractive part. The free parameters are chosen as $\gamma = 0.0806$ kcal/(mol Å^2), $p = 0.0191$ kcal/(mol Å^3), and $V_1^{\mathrm{vdW}}(\sigma_1) = 11.272$ kcal/mol.

Compound	Area (Å^2)	Volume (Å^3)	(kcal/mol)			
			Rep. part	Att. part	Numerical	Experimental[a]
Methane	50.78	32.27	4.71	− 2.73	1.98	2.00
Ethane	70.62	50.44	6.65	− 4.75	1.90	1.83
Propane	89.18	68.75	8.50	− 6.45	2.04	1.96
Butane	107.30	86.75	10.30	− 8.18	2.12	2.08
Pentane	126.49	105.01	12.19	− 9.82	2.37	2.33
Hexane	144.94	123.23	14.03	− 11.54	2.50	2.49
Isobutane	105.28	86.70	10.14	− 7.97	2.16	2.52
2-methylbutane	120.68	104.89	11.73	− 9.35	2.38	2.38
Neopentane	121.83	104.48	11.81	− 9.20	2.61	2.50
Cyclopentane	109.70	92.55	10.60	− 9.43	1.17	1.20
Cyclohexane	123.64	109.69	12.05	− 10.78	1.27	1.23
RMS of calibration set					0.12	
Octane	181.89	159.30	17.70	− 14.94	2.77	2.89
Heptane	163.14	141.27	15.85	− 13.22	2.63	2.62
Nonane	200.46	177.25	19.54	− 16.59	2.95	3.14
Decane	218.96	195.28	21.38	− 18.41	2.97	3.16
2-methylpentane	140.98	123.52	13.72	− 11.13	2.59	2.52
3-methylpropane	139.79	123.33	13.62	− 11.01	2.61	2.51
3-methylhexane	157.13	141.47	15.37	− 12.48	2.89	2.71
2-methylhexane	157.84	140.78	15.41	− 12.72	2.69	2.93
3-methylpentane	139.79	123.33	13.62	− 11.01	2.61	2.51
224-trimethylpentane	165.05	157.86	16.32	− 13.20	3.12	2.87
22-dimethylbutane	136.64	123.47	13.37	− 10.54	2.83	2.57
24-dimethylpentane	150.90	139.16	14.82	− 12.13	2.70	2.87
22-dimethylpentane	155.61	140.95	15.23	− 12.28	2.96	2.88
23-dimethylpentane	153.36	141.02	15.05	− 12.10	2.96	2.52
Cyclopropane	81.50	62.11	7.76	− 6.38	1.38	0.75
Cycloheptane	137.19	127.16	13.49	− 12.10	1.39	0.80
Cyclooctane	150.46	144.47	14.89	− 13.28	1.61	0.86
Methylcyclopentane	126.98	110.56	12.35	− 10.91	1.44	1.60
Methylcyclohexane	140.72	127.39	13.78	− 12.20	1.57	1.61
RMS of prediction set					0.31	

[a]Reference 58.

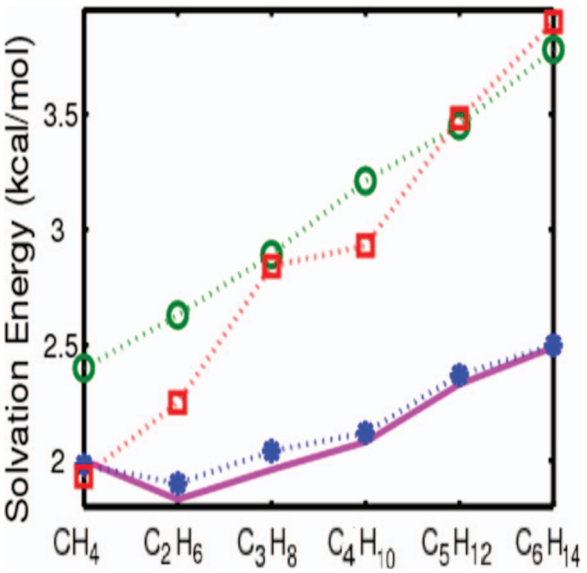

FIG. 3. Comparison of model predictions and experimental data[58] of solvation free energies of six alkanesmethane, ethane, propane, butane, pentane, and hexane. Solid lines are for experimental data. Stars are for the numerical data obtained by the present variational approach. Squares are for the numerical data reported in Ref. 51. Circles are for the numerical data reported in Ref. 4.

variational approach correctly captures the check mark shape.

The repulsive and attractive parts of solvation free energies can be calculated separately in the present model. Such solvation decomposition results are also listed in Table II. A comparison of the energies between the present variational approach and the explicit solvent model in Ref. 4 is given in Fig. 4. The present studies indicate that our variational approach performs better than an explicit solvent model[4] in predicting the total solvation energy.

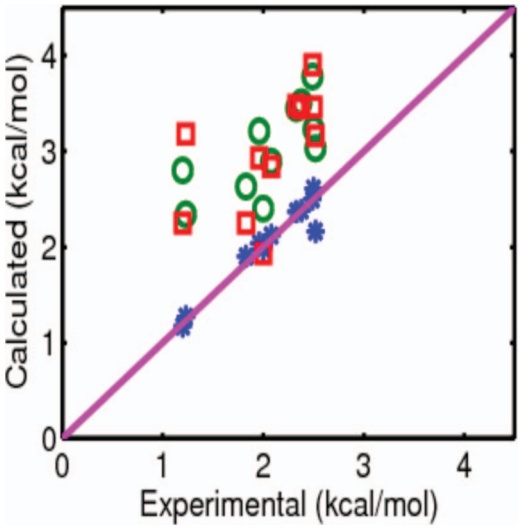

FIG. 4. Comparison of model predictions and experimental data[58] of solvation free energies of eleven alkanes. Solid lines are for experimental data. Stars are for the numerical data obtained by the present variational approach. Squares are for the numerical data reported in Ref. 51. Circles are for the numerical data reported in Ref. 4.

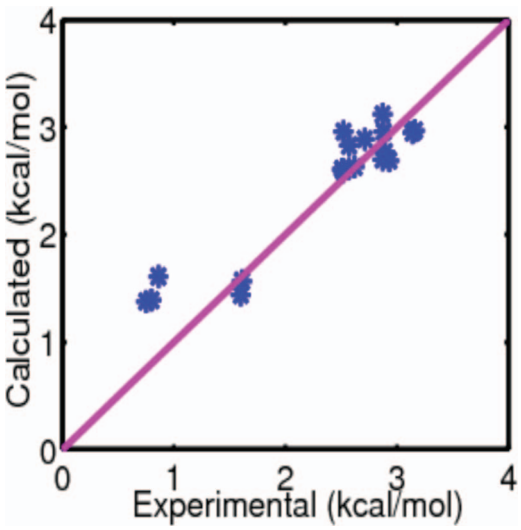

FIG. 5. Comparison of model predictions and experimental data of solvation free energies of nineteen alkanes.

Having demonstrated the accuracy and reliability of our variational approach, we next conduct a predictive study of 19 other alkane compounds. This test set includes linear, branched, and cyclic apolar compounds. The same optimized parameters and computational procedure used for the above mentioned 11 alkane molecules are employed. Results are shown in the bottom section of Table II, in which the repulsive and attractive decomposition is demonstrated as well. Our variational approach reproduces experimental results very well. The RMS error is 0.31 kcal/mol. Excellent agreement can be clearly seen in Fig. 5.

To further demonstrate the accuracy and reliability of the proposed variational approach, we finally consider calibration and prediction studies for a different type of apolar molecules, namely, alkenes. We employ a set of 11 alkene

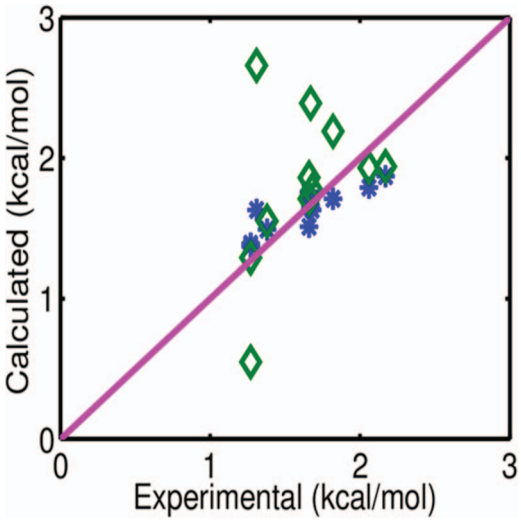

FIG. 6. Comparison of model predictions and experimental data[58] of solvation free energies of eleven alkenes. In all charts, solid lines are for experimental data. Stars are for the numerical data obtained by the present variational approach. Diamonds are for the numerical data reported in Ref. 61.

TABLE III. Numerical and experimental total solvation free energies for 11 alkenes. The numerical energy is the sum of the repulsive part and attractive part. The free parameters are chosen as $\gamma = 0.0775$ kcal/(mol Å^2), $p = 0.0192$ kcal/(mol Å^3), and $V_1^{\mathrm{vdW}}(\sigma_1) = 11.272$ kcal/mol.

Compound	Area (Å^2)	Volume (Å^3)	(kcal/mol)			
			Rep. part	Att. part	Numerical	Experimental[a]
3-methyl-1-butene	121.37	101.51	11.34	− 9.63	1.71	1.82
1-butene	103.99	83.30	9.65	− 8.16	1.49	1.38
Ethene	66.70	46.94	6.07	− 4.69	1.37	1.27
1-heptene	160.72	137.95	15.09	− 13.33	1.76	1.66
1-hexene	141.46	119.22	13.24	− 11.61	1.63	1.68
1-nonene	197.53	173.23	18.62	− 16.82	1.79	2.06
2-methyl-2-butene	120.55	101.33	11.28	− 9.65	1.63	1.31
RMS of calibration set					0.177	
1-octene	179.23	156.17	16.87	− 15.00	1.87	2.17
1-pentene	122.52	100.86	11.42	− 9.91	1.51	1.66
1-propene	85.59	65.55	7.89	− 6.49	1.39	1.27
Trans-2-heptene	160.84	137.00	15.08	− 13.43	1.65	1.66
RMS of prediction set					0.180	

[a]References 58 and 61.

compounds which have been studied by Ratkova *et al.*[61] using integral equation techniques. In the present study, we randomly choose a set of seven alkene compounds to train our model parameters, and then use them for the predictive study of the remaining four alkene molecules. Solvation free energies of 11 alkene compounds are shown in Table III. The RMS errors for the calibration set and the prediction set are found to be, respectively, 0.177 kcal/mol and 0.180 kcal/mol. These RMS errors are significantly better than that reported in Ref. 61, which is about 0.462 kcal/mol, obtained by using the integral equation approach. This finding can be observed clearly from Fig. 6. In general, the present variational approach is simpler in theory and easier to implement than the integral equation approach.[61]

IV. CONCLUSION

In conclusion, this work provides a differential geometry based variational approach to improve the solvent-solute interface description in implicit solvent theory. This variational model offers increased accuracy and reliability in solvation free energy prediction, indicating the power of differential geometric methods for the analysis of solvation free energies. An implication of the present result is that the differential geometry based definition of solvent-solute boundaries is perhaps sufficient for solvation free energy predictions, although generalized correlations[50] are needed in order to reproduce correct solvent microstructures.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation (NSF) Grant Nos. CCF-0936830 and DMS-1160352, and National of Institutes of Health (NIH) Grant Nos. R01GM-090208 and GM069702. We thank the anonymous reviewers for useful suggestions.

[1]J. W. Ponder and D. A. Case, "Force fields for protein simulations," Adv. Protein Chem. **66**, 27–85 (2003).
[2]S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, "Progress in *ab initio* QM/MM free-energy simulations of electrostatic energies in proteins: Accelerated QM/MM studies of pK_a, redox reactions and solvation free energies," J. Phys. Chem. B **113**, 1253–1272 (2009).
[3]R. M. Levy, L. Y. Zhang, E. Gallicchio, and A. K. Felts, "On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute-solvent interaction energy," J. Am. Chem. Soc. **125**(31), 9523–9530 (2003).
[4]E. Gallicchio, M. M. Kubo, and R. M. Levy, "Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation," J. Phys. Chem. B **104**(26), 6271–6285 (2000).
[5]R. Ishizuka, S. H. Chong, and F. Hirata, "An integral equation theory for inhomogeneous molecular fluids: The reference interaction site model approach," J. Chem. Phys. **128**(3), 34504–34504 (2008).
[6]C. S. Hsu and D. Chandler, "Rism calculation of the structure of liquid acetonitrile," Mol. Phys. **36**(1), 215–224 (1978).
[7]A. Kovalenko and F. Hirata, "Self-consistent description of a metal–water interface by the kohn–Sham density functional theory and the three-dimensional reference interaction site model," J. Chem. Phys. **110**, 10095 (1999).
[8]J. L. Lebowitz and J. K. Percus, "Mean spherical model for lattice gases with extended hard cores and continuum fluids," Phys. Rev. **144**(1), 251 (1966).
[9]R. Triolo, J. R. Grigera, and L. Blum, "Simple electrolytes in the mean spherical approximation," J. Phys. Chem. **80**(17), 1858–1861 (1976).
[10]J. M. J. Van Leeuwen, J. Groeneveld, and J. De Boer, "New method for the calculation of the pair correlation function. I," Physica **25**(7–12), 792–808 (1959).
[11]L. S. Ornstein and F. Zernike, "Accidental deviations of density and opalescence at the critical point of a single substance," Proc. R. Acad. Sci. Amsterdam **17**, 793 (1914).
[12]J. K. Percus, H. L. Frisch, and J. L. Lebowitz, *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz (Benjamin, 1964), p. 1133.
[13]J. P. Donley, J. G. Curro, and J. D. McCoy, "A density functional theory for pair correlation functions in molecular liquids," J. Chem. Phys. **101**, 3205 (1994).
[14]J. K. Percus and G. J. Yevick, "Analysis of classical statistical mechanics by means of collective coordinates," Phys. Rev. **110**(1), 1 (1958).
[15]M. S. Wertheim, "Exact solution of the Percus-Yevick integral equation for hard spheres," Phys. Rev. Lett. **10**(8), 321–323 (1963).
[16]P. Koehl, H. Orland, and M. Delarue, "Beyond Poisson-Boltzmann: Modeling biomolecule-water and water-water interactions," Phys. Rev. Lett. **102**(8), 087801 (2009).
[17]S. A. Hassan, "Liquid-structure forces and electrostatic modulation of biomolecular interactions in solution," J. Phys. Chem. B **111**(1), 227–241 (2007).

[18]A. Rubinstein and S. Sherman, "Evaluation of the influence of the internal aqueous solvent structure on electrostatic interactions at the protein-solvent interface by nonlocal continuum electrostatic approach," Biopolymers **87**(2-3), 149–164 (2007).
[19]P. Madden and D. Kivelson, *A Consistent Molecular Treatment of Dielectric Phenomena*, Advances in Chemical Physics Vol. 56, (1984), pp. 467–566.
[20]K. A. Sharp and B. Honig, "Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equatlon," J. Phys. Chem. **94**, 7684–7692 (1990).
[21]M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon, "Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation," J. Phys. Chem. **97**(14), 3591–3600 (1993).
[22]F. H. Stillinger, "Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory," J. Solution Chem. **2**, 141–158 (1973).
[23]F. Dong, J. A. Wagoner, and N. A. Baker, "Assessing the performance of implicit solvation models at a nucleic acid surface," Phys. Chem. Chem. Phys. **10**, 4889–4902 (2008).
[24]C. J. Cramer and D. G. Truhlar, "Implicit solvation models: Equilibria, structure, spectra, and dynamics," Chem. Rev. **99**, 2161–2200 (1999).
[25]B. Husowitz and V. Talanquer, "Solvent density inhomogeneities and solvation free energies in supercritical diatomic fluids: A density functional approach," J. Chem. Phys. **126**(5), 054508 (2007).
[26]M. R. Reddy, U. C. Singh, and M. D. Erion, "*Ab initio* quantum mechanics-based free energy perturbation method for calculating relative solvation free energies," J. Comput. Chem. **28**(2), 491–494 (2007).
[27]A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Perspective on foundations of solvation modeling: The electrostatic contribution to the free energy of solvation," J. Chem. Theory Comput. **4**(6), 877–887 (2008).
[28]W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, and B. Honig, "Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects," J. Comput. Chem. **23**, 128–137 (2002).
[29]B. Lee and F. M. Richards, "The interpretation of protein structures: Estimation of static accessibility," J. Mol. Biol. **55**(3), 379–400 (1971).
[30]F. M. Richards, "Areas, volumes, packing, and protein structure," Annu. Rev. Biophys. Bioeng. **6**(1), 151–176 (1977).
[31]R. S. Spolar, J. H. Ha, and M. T. Record, Jr., "Hydrophobic effect in protein folding and other noncovalent processes involving proteins," Proc. Natl. Acad. Sci. U.S.A. **86**(21), 8382–8385 (1989).
[32]J. R. Livingstone, R. S. Spolar, and M. T. Record, Jr., "Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area," Biochemistry **30**(17), 4237–4244 (1991).
[33]P. B. Crowley and A. Golovin, "Cation-π interactions in protein-protein interfaces," Proteins: Struct., Funct., Bioinf. **59**(2), 231–239 (2005).
[34]L. A. Kuhn, M. A. Siani, M. E. Pique, C. L. Fisher, E. D. Getzoff, and J. A. Tainer, "The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures," J. Mol. Biol. **228**(1), 13–22 (1992).
[35]C. A. S. Bergstrom, M. Strafford, L. Lazorova, A. Avdeef, K. Luthman, and P. Artursson, "Absorption classification of oral drugs based on molecular surface properties," J. Med. Chem. **46**(4), 558–570 (2003).
[36]A. I. Dragan, C. M. Read, E. N. Makeyeva, E. I. Milgotina, M. E. Churchill, C. Crane-Robinson, and P. L. Privalov, "DNA binding and bending by HMG boxes: Energetic determinants of specificity," J. Mol. Biol. **343**(2), 371–393 (2004).
[37]R. M. Jackson and M. J. Sternberg, "A continuum model for protein-protein interactions: Application to the docking problem," J. Mol. Biol. **250**(2), 258–275 (1995).
[38]V. J. Licata and N. M. Allewell, "Functionally linked hydration changes in *Escherichia coli* aspartate transcarbamylase and its catalytic subunit," Biochemistry **36**(33), 10161–10167 (1997).
[39]M. F. Sanner, A. J. Olson, and J. C. Spehner, "Reduced surface: An efficient way to compute molecular surfaces," Biopolymers **38**, 305–320 (1996).
[40]G. W. Wei, Y. H. Sun, Y. C. Zhou, and M. Feig, "Molecular multiresolution surfaces," e-print arXiv:math-ph/0511001v1.
[41]P. W. Bates, G. W. Wei, and S. Zhao, "The minimal molecular surface," e-print arXiv:q-bio/0610038v1 [q-bio.BM].
[42]P. W. Bates, G. W. Wei, and S. Zhao, "Minimal molecular surfaces and their applications," J. Comput. Chem. **29**(3), 380–391 (2008).
[43]P. W. Bates, Z. Chen, Y. H. Sun, G. W. Wei, and S. Zhao, "Geometric and potential driving formation and evolution of biomolecular surfaces," J. Math. Biol. **59**, 193–231 (2009).
[44]G. W. Wei, "Differential geometry based multiscale models," Bull. Math. Biol. **72**, 1562–1622 (2010).
[45]Z. Chen, N. A. Baker, and G. W. Wei, "Differential geometry based solvation models I: Eulerian formulation," J. Comput. Phys. **229**, 8231–8258 (2010).
[46]Z. Chen, N. A. Baker, and G. W. Wei, "Differential geometry based solvation models II: Lagrangian formulation," J. Math. Biol. **63**, 1139–1200 (2011).
[47]Z. Chen and G. W. Wei, "Differential geometry based solvation model. III. Quantum formulation," J. Chem. Phys. **135**, 194108 (2011).
[48]J. Dzubiella, J. M. J. Swanson, and J. A. McCammon, "Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models," Phys. Rev. Lett. **96**, 087802 (2006).
[49]L. T. Cheng, J. Dzubiella, J. A. McCammon, and B. Li, "Application of the level-set method to the implicit solvation of nonpolar molecules," J. Chem. Phys. **127**(8), 084503 (2007).
[50]G.-W. Wei, Q. Zheng, Z. Chen, and K. Xia, "Variational multiscale models for charge transport," SIAM Rev. (to be published).
[51]J. A. Wagoner and N. A. Baker, "Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms," Proc. Natl. Acad. Sci. U.S.A. **103**(22), 8331–8336 (2006).
[52]E. Gallicchio, L. Y. Zhang, and R. M. Levy, "The SGB/NP hydration free energy model based on the surface generalized Born solvent reaction field and novel nonpolar hydration free energy estimators," J. Comput. Chem. **23**(5), 517–529 (2002).
[53]E. Gallicchio and R. M. Levy, "AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling," J. Comput. Chem. **25**(4), 479–499 (2004).
[54]J. D. Weeks, D. Chandler, and H. C. Andersen, "Role of repulsive forces in determining the equilibrium structure of simple liquids," J. Chem. Phys. **54**(12), 5237–5247 (1971).
[55]A. Nicholls, D. L. Mobley, P. J. Guthrie, J. D. Chodera, and V. S. Pande, "Predicting small-molecule solvation free energies: An informal blind test for computational chemistry," J. Med. Chem. **51**(4), 769–779 (2008).
[56]W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface reconstruction algorithm," Comput. Graphics **21**, 163–169 (1987).
[57]W. H. Geng and G. W. Wei, "Multiscale molecular dynamics via the matched interface and boundary (mib) method," J. Comput. Phys. **230**, 435–457 (2011).
[58]S. Cabani, P. Gianni, V. Mollica, and L. Lepori, "Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution," J. Solution Chem. **10**(8), 563–595 (1981).
[59]D. A. Perlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, and P. Kollman, "AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules," Comput. Phys. Commun. **91**, 1–41 (1995).
[60]D. Veenstra, D. Ferguson, and P. Kollman, "How transferable are hydrogen parameters in molecular mechanics calculations?," J. Comput. Chem. **13**(8), 971–978 (1992).
[61]E. L. Ratkova, G. N. Chuev, V. P. Sergiievskyi, and M. V. Fedorov, "An accurate prediction of hydration free energies by combination of molecular integral equations theory with structural descriptors," J. Phys. Chem. B **114**(37), 12068–12079 (2010).