

A Level-Set Variational Implicit-Solvent Approach to Hydrophobic Interactions

Zhongming Wang

Biochemistry and Mathematics, UCSD

Joint work with L.-T. Cheng, P. Setny, J. Dzubiella,
B. Li and J. A. McCammon

UCSD Informal Seminars on Mathematics and
Biochemistry-Biophysics, Jan 27, 2009

Introduction

Variational implicit or continuum solvent models

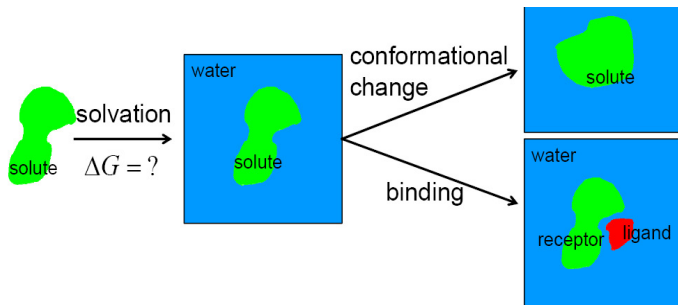
Level set method

Hydrophobic pocket

Conclusions and outlook

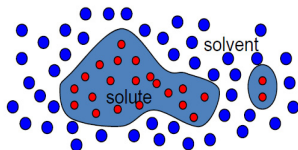
Basic facts

- Biological processes, such as molecular recognition and protein folding, occur in solution
- Solute-solvent interactions are crucial in determining biomolecular structures and solvation free energies



Explicit solvent vs implicit solvent

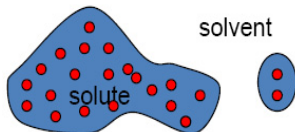
- Explicit solvent: solvent atoms treated explicitly (e.g., MD, MC)



- Advantages
 - First principle, accurate
- Disadvantages
 - Non-efficient, small portion of solutes (5-10%), boundary effects, uncertain convergence, statistical information

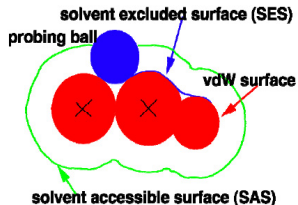
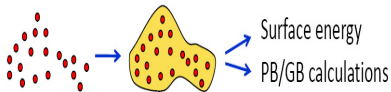
Explicit solvent vs implicit solvent cont'd

- Implicit solvent: solvent atoms are treated implicitly and solvent effects are coarse grained



- Advantages
 - Efficient, large systems, thermodynamical variables
- Disadvantages
 - Mean field ion distribution, ignores details

Existing implicit solvent models

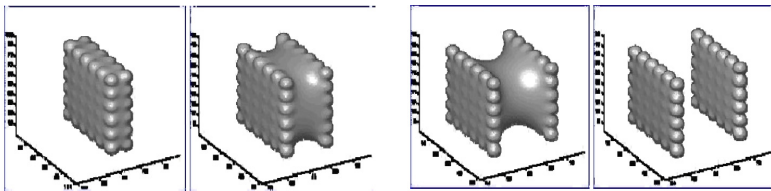


- Get atom positions
- Generate molecular surface(vdW surface, SES/SAS)
- Calculate surface energy
- Calculate electrostatic energy by PB/GB with the surface as dielectric boundary

Existing implicit solvent models cont'd

Issues of vdW surface, SAS/SES models

- Pre-defined dielectric interface
- Curvature correction
- Coupling of polar and nonpolar contributions
- Hydrophobicity ¹

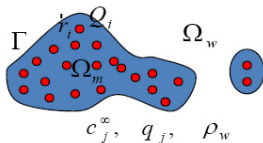


¹From Cheng, Dzubiella, McCammon, & Li, J. Chem. Phys., 2007

Variational implicit solvent model facts

- Variational implicit solvent model (VISM)
 - Dzubiella, Swanson, & McCammon, Phys. Rev. Lett. 96, 087802 (2006); J. Chem. Phys. 124, 084905 (2006)
 - Che, Dzubiella, Li, & McCammon, J. Phys. Chem. B, 112, 3058 (2008)
- VISM ideas
 - Solvation structure = solutes position + solute-solvent interface
 - Free energy minimization determines solute-solvent interface
 - Free energy couples with different interactions: polar, nonpolar, dispersive, etc.

VISM geometry and parameters



- $\Omega_m, \Omega_w, \Gamma$: solute region, solvent region, interface
- r_i, Q_i : center and fixed charge of i^{th} solute atom
- c_j^∞, q_j : bulk density and bulk charge of j^{th} ionic specie
- ρ_w : solvent density

Free energy functional

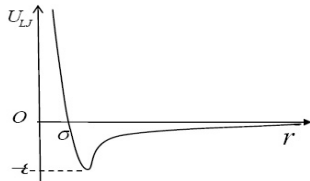
$$G[\Gamma] = G_{geom}[\Gamma] + G_{vdW}[\Gamma] + G_{elec}[\Gamma]$$

- $G_{geom}[\Gamma] = Pvol(\Omega_m) + \int_{\Gamma} \gamma(r) dS$
 $\gamma(r) = \gamma_0[1 - 2\tau H(r)]$
 τ : Tolman length, H : mean curvature

$$G_{vdw}[\Gamma] = \rho_w \int_{\Omega_s} U(r) dV$$

- $U(r) = \sum_i U_i(|r - r_i|)$

$$U_i(|r|) = 4\epsilon \left(\left(\frac{\sigma_i}{|r|} \right)^{12} + \left(\frac{\sigma_i}{|r|} \right)^6 \right)$$



- $G_{elec}[\Gamma]$ use PB/GB, Coulomb approximation, etc (not presented in this pocket project)

Energy minimization

- A Necessary condition for Γ to be a minimizer

$$\delta_{\Gamma} G[\Gamma] = 0, \forall r \in \Gamma$$

- The energy minimizer Γ_{min} satisfies a PDE

$$P + 2\gamma_0(H(r) - \tau K(r)) - \rho_w U(r) + \left(\delta_{\Gamma} G_{elec}[\Gamma] \Big|_{\Gamma_{min}} \right) = 0, \quad \forall r \in \Gamma_{min}$$

- How to find the minimizer Γ_{min} ? Level set method

Level set method

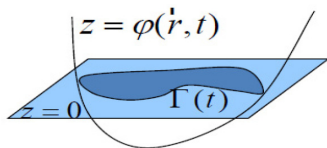
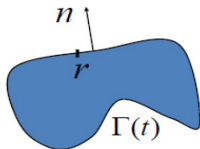
- Proposed by Sethian and Osher (1988); applied to image processing, computer vision, computational physics, etc.
- Interfaces (line, curve, surface, etc) are represented implicitly,

$$\Gamma(t) = \{r \in \Omega : \Phi(r, t) = 0\}$$

- Motion of the interface is realized by evolution of the auxiliary function $\Phi(r, t)$

$$\Phi_t + V_n |\nabla \Phi| = 0, \quad V_n = V_n(r, t) \text{ for } r \in \Gamma$$

- V_n is the normal velocity, which may be given or depend on curvatures

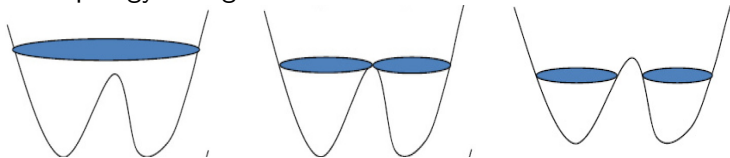


Geometric quantities in level set functions

- Unit normal vector $n = \frac{\nabla\Phi}{|\nabla\Phi|}$
- Mean curvature $H = \frac{1}{2}\nabla \cdot n$
- Gaussian curvature $K = n \cdot \text{adj}(\text{He}(\Phi))n$
 $\text{adj}(\cdot)$: adjoint matrix
 $\text{He}(\cdot)$: Hessian matrix
- Surface integral $\int_{\Gamma} f(r)dS = \int_{\mathbb{R}^3} f(r)\delta(\Phi)dV$
- Volume integral $\int_{\Omega} f(r)dV = \int_{\mathbb{R}^3} f(r)(1 - \text{Heav}(\Phi))dV$

Facts about level set method

- Uniform mesh grids may be used
- Automatically handles topological change: merging, breaking, etc., unlike ray tracing method which needs to relabel marks when topology changes



- Local level set method can be used to reduce computational cost
- Accuracy issue: interface approximation, conservation of mass, etc.

Application to VISM

- Cheng, Dzubiella, McCammon, & Li, J. Chem. Phys., 2007
- Cheng, Xie, Dzubiella, McCammon, Che, & Li, J. Chem. Theory Comput., (2008) accepted
- Cheng, Wang, Setny, Dzubiella, Li, & McCammon, 2008 (preprint)

Level set equation of relaxation:

$$\Phi_t + V_n |\nabla \Phi| = 0, V_n = -\delta_\Gamma G[\Gamma]$$

$$\delta_\Gamma G[\Gamma] = P + 2\gamma_0(H(r) - \tau K(r)) - \rho_w U(r) + \delta_\Gamma G_{elec}[\Gamma]$$

Motion of solute atoms due to interactions:

$$\frac{dx_i}{dt} = -\nabla_{x_i} G[\Gamma; x_1, \dots, x_n]$$

Application to VISM cont'd

Decay of the free energy:

$$\begin{aligned}\frac{dG[V_n(t)]}{dt} &= \int_{\Gamma(t)} \frac{\delta G[V_n(t)]}{\delta V_n} \left[\frac{dr(t)}{dt} \cdot n \right] dS \\ &= - \int_{\Gamma(t)} \left| \frac{\delta G[V_n(t)]}{\delta V_n} \right|^2 dS < 0\end{aligned}$$

- $\Gamma = \Gamma(t)$ and t is not the real dynamical time, but represents the minimization iteration
- The normal velocity does not represent the interface evolution in real dynamics of the system.

Algorithm

- Step 1. Input parameters and initialize level-set function–signed distance function
- Step 2. Calculate the normal and curvatures
- Step 3. Calculate and extend the normal velocity
- Step 4. Solve the level-set equation
- Step 5. Reinitialize the level-set function
- Step 6. Solve ODEs for the motion of solute particles
- Step 7. Set $t = t + \Delta t$ and go to Step 2

Solving the level set equation numerically²

$$\Phi_t + V_n(r)|\nabla\Phi| = 0 \quad (1)$$

$$V_n(r) = -P - 2\gamma_0(H(r) - \tau K(r)) + \rho_w U(r)$$

- The mean curvature H and Gaussian curvature K are calculated from Φ
- Equation (1) is parabolic and may degenerate in certain parameter regimes. Computationally the degeneracy is removed by adding a small constant to the eigenvalues of the matrix that defines the parabolicity

²See in Cheng, Dzubiella, McCammon, & Li, J. Chem. Phys., 2007

Solving the level set equation numerically cont'd

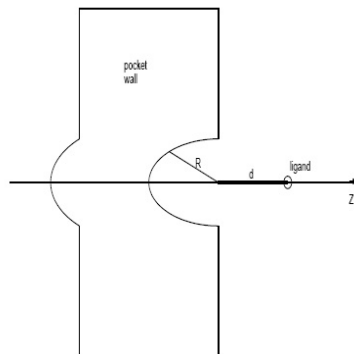
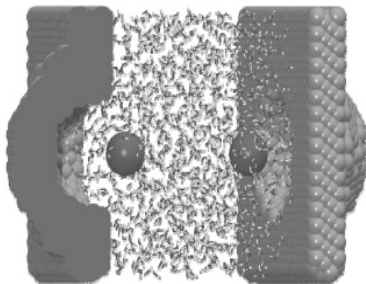
- Reinitialization is used to keep $\nabla\Phi$ away from zero while the interface Γ is unchanged
- The normal velocity V_n is only defined on Γ theoretically, and extended in normal direction both sides numerically
- For large system, $U(r) = \sum_i U_i(|r - r_i|)$ is pre-computed and stored for each grid point (in the case of fixed solutes); a linear approximation is employed for any $r \in \Gamma$

Solving the level set equation numerically cont'd

- When distance is small, LJ potential changes rapidly. To improve the accuracy, we can compute LJ on a fine grid while the level set function is evolved on a coarse grid
- Make proper initial guess based on a priori knowledge of the interface shape

These will drastically speed up the program. For the system of 4,243 solute atoms, with a $50 \times 50 \times 50$ grid size and a good initial guess, it takes 5 minutes. Two hours with a bad initial guess and high resolution.

Pocket system³

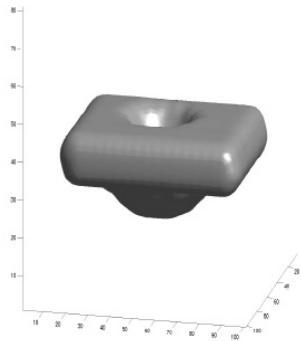
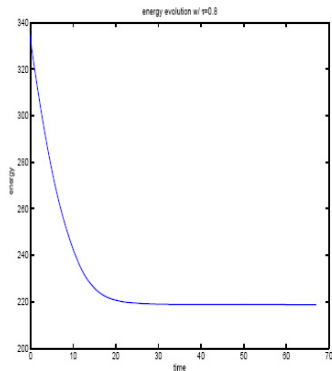


Each wall consists of 4,242 atoms.

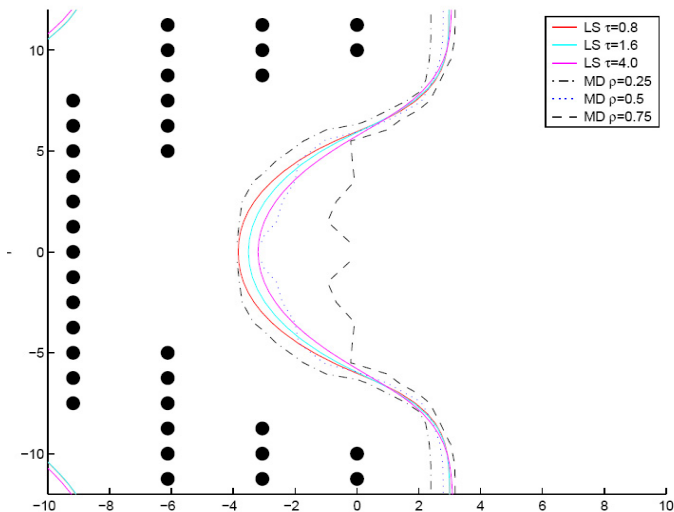
³System studied in P. Setny, J. Chem. Phys. 127, 054505 (2007); P. Setny and M. Geller, 125, 144717 (2006) and P. Setny, J. Chem. Phys. 128, 125105 (2008)

Wall ($R8$) without ligand

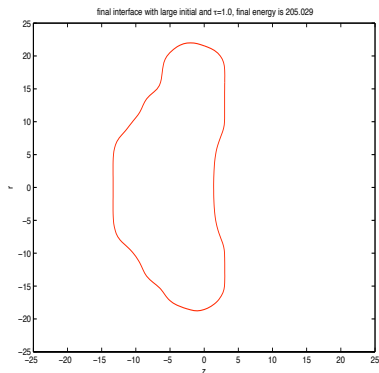
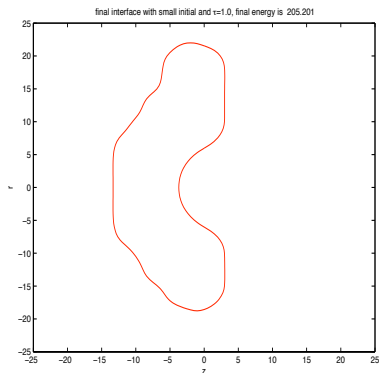
We first study the case of wall without ligand. We focus on the energy relaxation and interface with the effect of Tolman correction.



Wall ($R8$) without ligand: effect of Tolman correction



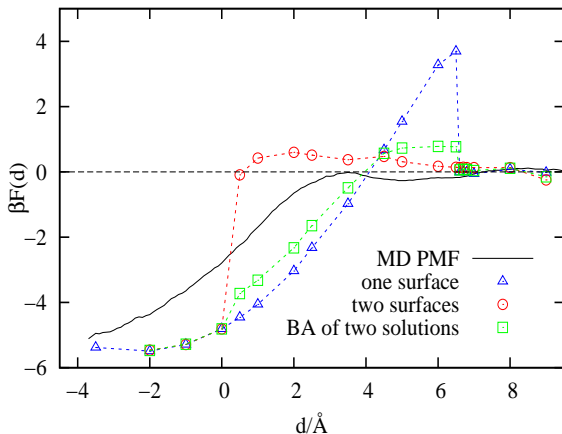
Wall ($R8$) without ligand: effect of initial interface



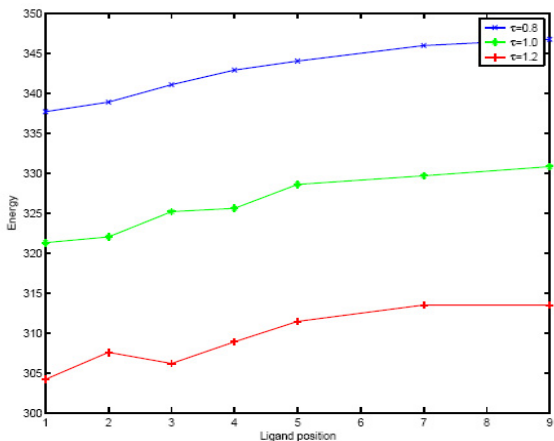
Left: final interface with tight initial; right: final interface with loose initial

Wall ($R8$) with ligand

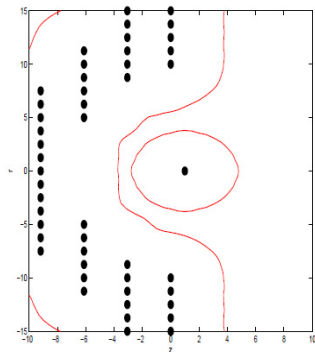
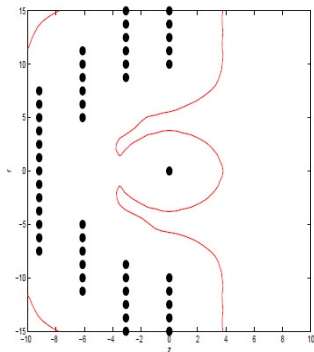
In this case, we study the interface and energy with respect to the ligand position. Bi-mode is observed when $d \in [0, 7]$.



Wall (R_8) with ligand: effect of Tolman correction

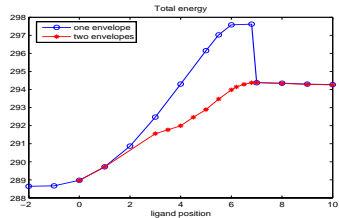
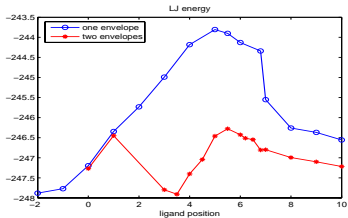
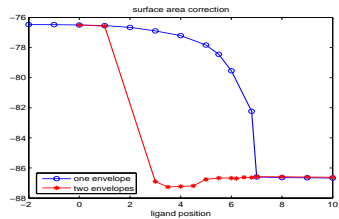
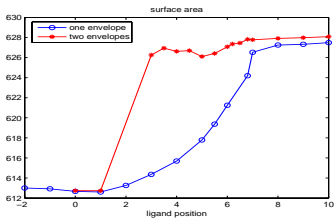


Wall ($R8$) with ligand: effect of initial interface ($d = 0$)

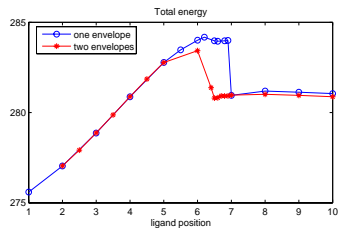
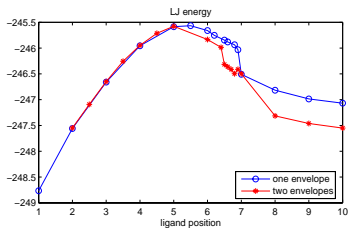
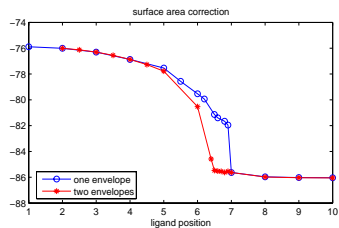
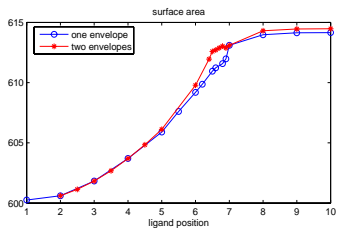


Left: final interface with loose initial; right: final interface with tight initial

Wall ($R5$) with ligand: energy profile

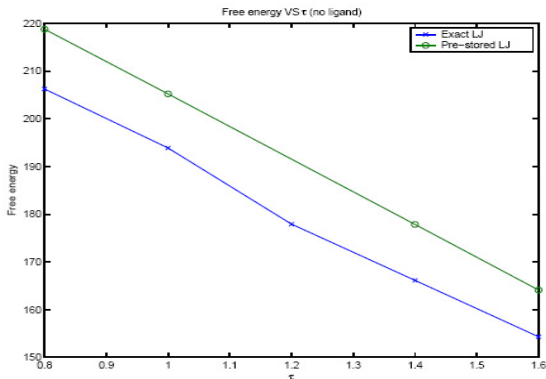


Wall (R_0) with ligand: energy profile



Efficiency vs accuracy

Efficiency is achieved at the cost of accuracy. Here is a figure showing the energy difference of the wall without ligand, computed with and without pre-computation of LJ potential.



Conclusion and outlook

- Conclusions
 - Tested the theory and method, and find that they work well to capture hydrophobic interactions
 - Curvature correction has a stronger influence to the concave region than the convex region of solute-solvent interfaces
 - Describe the bimodal behavior very well
 - Very accurate free energy calculations
- Discussion
 - In terms of the model, how to choose the Tolman length
 - Comparison with SAS/SES type of implicit-solvent models
 - How to include the electrostatics efficiently?
 - The efficiency of the method compared with other implicit models?

Thank you!

- Reinitialization: solve

$$\Phi_t + \text{sgn}(\Phi)(\nabla\Phi - 1) = 0$$

to a steady state, where

$$\text{sgn}(\Phi) = \frac{\Phi}{\sqrt{\Phi^2 + \epsilon}}.$$

- Parabolicity: curvatures H and K are calculated from the second fundamental form

$$H = \frac{1}{|\nabla\Phi|} P_{\nabla\Phi} \nabla^2 \Phi P_{\nabla\Phi}, \quad P_{\nabla\Phi} = I - \frac{\nabla\Phi \otimes \nabla\Phi}{|\nabla|^2}.$$

H is used to make sure the parabolicity