Level-Set Variational Implicit-Solvent Modeling of biomolecular Solvation

Bo Li
Department of Mathematics and
NSF Center for Theoretical Biological Physics
University of California, San Diego

Department of Molecular Biophysics and Physiology
Rush University, Chicago
September 21, 2011
Collaborators

J. Andrew McCammon (Biochem & CTBP, UCSD)
Joachim Dzubiella (Phys., Helmholtz Ctr. & Humboldt Univ.)
Jianwei Che (Genomics Inst. of Novartis Res. Found)
Piotr Setny (Biophysics, Munich & Warsaw)
Yang Xie (Mech Eng., UCSD & Georgia Tech)

Li-Tien Cheng (Math, UCSD)
Zhongming Wang (Math, Florida International Univ.)
Shenggao Zhou (Math, Zhejiang Univ. & UCSD)
Hsiao-Bing Cheng (Math, UCSD)
Xiaoliang Cheng (Math, Zhejiang Univ.)
Zhengfang Zhang (Math, Zhejiang Univ. & UCSD)

Funding: NSF, DOE, NIH, CTBP
OUTLINE

1. Introduction
2. Variatioal implicit-solvent models
3. Continuum electrostatics
4. The level-set method
5. Numerical results
6. Conclusions
1. Introduction

\[ \Delta G = ? \]

Explicit solvent

Implicit solvent

- Solvation
- Conformational change
- Binding

Water
Solute
Solute
Receptor
Ligand
Main interactions in implicit solvation

Electrostatic interactions

Fundamental, many-body, long-range.

Coulomb’s law:

\[
\vec{F} = \frac{1}{4\pi\varepsilon_0} \frac{Q_1 Q_2}{r^3} \hat{r}
\]

Poisson’s equation:

\[
\nabla \cdot \varepsilon \varepsilon_0 \nabla \psi = -\rho
\]

- Local dielectric screening by water
- \( \rho = \) fixed charges + mobile ions
- \( \psi = \) electrostatic potential
Excluded volume and van der Waals dispersion

The van der Waals (vdW) equation

\[
\left[ P + a \left( \frac{n}{V} \right)^2 \right] \left( \frac{V}{n} - b \right) = RT
\]

attraction excluded volume

The Lennard-Jones (LJ) potential

\[
U_{LJ}(r) = 4\varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]
\]

Fermi repulsion vdw attraction
Hydrophobic interactions

Water does not like to be at an interface: missing attractive interactions; rearranging.

Curvature effects at small scales

Symbols: MD.
\[ \gamma = \gamma_0 \left(1 - 2\tau H\right) \]

\( \gamma_0 = 73 \text{ mJ/m}^2 \)
\( \tau = 0.9 \ \text{Å} \)

Commonly used implicit-solvent models

PB = Poisson-Boltzmann
GB = Generalized Born

Fixed surface (dielectric boundary)

Possible issues:
- Hydrophobic cavities
- Curvature correction
- Decoupling of polar and nonpolar contributions

Four examples below illustrate these issues.
Capillary evaporation in hydrophobic confinement

A receptor-ligand system


MD: weakly solvated pocket, strong hydrophobic attraction.

SASA/MSA: Onset of attraction is wrong by 2-4 Angstroms!
Evaporation in proteins

MD simulations of the melittin protein tetramer

- Water in hydrophobic core
- Stable nanobubble


More MD simulations

- Electrostatics
- Curvature

Giovambattista et al., PNAS, 105, 2274 (2008).
Voltage-gating in ion channels

Consistent with MD simulations

Narrow hydrophobic region induces evaporation.

Channel is not permeable to ions!

Pore fills with application of electrostatic potential.

Channel conducts!

2. Variational Implicit-Solvent Models (VISM)

Dzubiella, Swanson, & McCammon

Guiding principles

- Solvation structure
  = Solute atomic positions + Solute-solvent interface
- Free-energy minimization → solute-solvent interfaces and solvation free energy
- Free energy couples different contributions: polar, nonpolar, dispersive, etc.
Free-energy functional

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

\[ G_{\text{geom}}[\Gamma] = P_{\text{vol}}(\Omega_m) + \int_{\Gamma} \gamma(\vec{r}) dS \]

- \( P_{\text{vol}}(\Omega_m) \): Creation of a cavity in the solvent
- \( P \): Liquid-vapor pressure difference

\[ \int_{\Gamma} \gamma(\vec{r}) dS \]: Molecular rearrangement near the interface

\[ \gamma = \gamma(\vec{r}) \]: Surface tension

\[ \gamma(\vec{r}) = \gamma_0 \left[ 1 - 2\tau H(\vec{r}) \right] \] (Scaled Particle Theory)

- \( \gamma_0 \): the (planar) surface tension
- \( \tau \): the Tolman length, a fitting parameter
- \( H = H(\vec{r}) \): mean curvature
\[ G_{vdW}[\Gamma] = \rho_w \int_{\Omega_w} U(\mathbf{r})dV \]

solute-solvent interaction

\[ \rho_w = \text{constant solvent density} \]

\[ U(\mathbf{r}) = \sum_i U_{LJ,i}(|\mathbf{r} - \mathbf{r}_i|) \]

\[ G_{elec}[\Gamma] - \text{Electrostatic free energy} \]

- The Poisson-Boltzmann (PB) theory
- The Coulomb-field or Yukawa-field approximation

**Free-energy functional**

\[ G[\Gamma] = Pvol(\Omega_m) + \gamma_0 \int (1 - 2\tau H)dS + \rho_w \int \sum_{\Omega_w} U_{LJ,i}(|\mathbf{r} - \mathbf{r}_i|)dV + G_{elec}[\Gamma] \]
\[ G_{\text{geom}}[\Gamma] = P\text{vol}(\Omega_m) + \gamma_0 \text{area}(\Gamma) - 2\gamma_0 \tau \int_{\Gamma} H dS \left( + c_K \int_{\Gamma} K dS \right) \]

**Hadwiger’s Theorem**

Let \( C = \) the set of all convex bodies,
\[ M = \] the set of finite union of convex bodies.

If \( F : M \to R \) is

- rotational and translational invariant,
- additive:
  \[ F(U \cup V) = F(U) + F(V) - F(U \cap V) \quad \forall U, V \in M, \]
- conditionally continuous:
  \[ U_j, U \in C, U_j \to U \Rightarrow F(U_j) \to F(U), \]

then

\[ F(U) = a \text{Vol}(U) + b \text{Area}(\partial U) + c \int_{\partial U} H dS + d \int_{\partial U} K dS \quad \forall U \in M. \]

**Application to nonpolar solvation**


Coupling solute molecular mechanics with implicit solvent

\[ V[\vec{r}_1, \ldots, \vec{r}_N] = \sum_{i,j} W_{\text{bond}}(\vec{r}_i, \vec{r}_j) + \sum_{i,j,k} W_{\text{bend}}(\vec{r}_i, \vec{r}_j, \vec{r}_k) \]
\[ + \sum_{i,j,k,l} W_{\text{torsion}}(\vec{r}_i, \vec{r}_j, \vec{r}_k, \vec{r}_l) + \sum_{i,j} W_{\text{LJ}}(\vec{r}_i, \vec{r}_j) \]
\[ + \sum_{i,j} W_{\text{Coulomb}}(\vec{r}_i, Q_i; \vec{r}_j, Q_j) \]

An effective total Hamiltonian

\[ H[\Gamma; \vec{r}_1, \ldots, \vec{r}_N] = V[\vec{r}_1, \ldots, \vec{r}_N] + G[\Gamma; \vec{r}_1, \ldots, \vec{r}_N] \]
\[ \min H[\Gamma; \vec{r}_1, \ldots, \vec{r}_N] \quad \Rightarrow \quad \text{Equilibrium conformations} \]
3. Continuum Electrostatics

The (generalized) PBE: \[ \nabla \cdot \varepsilon \varepsilon_0 \nabla \psi - \chi_w B'(\psi) = -\rho_f \]

Poisson’s Eq. \[ \nabla \cdot \varepsilon \varepsilon_0 \nabla \psi = -\rho \]

Charge density \[ \rho = \rho_f + \rho_i \]

Boltzmann dist. \[ \rho_i = -\chi_w B'(\psi) \]

- Nonlinear PB without size effect
  \[ B(\psi) = \beta^{-1} \sum_{j=1}^{M} c_j^\infty \left( e^{-\beta q_j \psi} - 1 \right) \]

- Linearized PBE (Debye-Huckel Eq.)
  \[ B(\psi) = \frac{1}{2} \kappa^2 \psi \]

- Nonlinear PBE with (uniform) size effect
  \[ B(\psi) = -(\beta v)^{-1} \log \left( 1 + v \sum_{j=1}^{M} c_j^\infty e^{-\beta q_j \psi} \right) \]
**PBE:** \[ \nabla \cdot \varepsilon \varepsilon_0 \nabla \psi - \chi_w B'(\psi) = -\rho_f \]

**Electrostatic free energy**

\[
G_{elec}[\Gamma] = \int \left[ -\frac{\varepsilon \varepsilon_0}{2} |\nabla \psi|^2 + \rho_f \psi - \chi_w B(\psi) \right] dV = \max_{\phi} G_{elec}[\Gamma, \phi]
\]

\[
G_{elec}[\Gamma, \phi] = \int \left[ -\frac{\varepsilon \varepsilon_0}{2} |\nabla \phi|^2 + \rho_f \phi - \chi_w B(\phi) \right] dV
\]

**Theorem.** \( G_{elec}[\Gamma, \bullet] \) has a unique maximizer \( \psi \):

- Uniformly bounded in \( H^1 \) and \( L^\infty \);
- The unique solution to the PBE.

**Proof.** Direct methods in the calculus of variations.

- Uniform bounds by comparison.
- Regularity theory and routine calculations.  **Q.E.D.**
Dielectric boundary force: \( F_n = -\delta_\Gamma G_{elec}[\Gamma] \)

A shape derivative approach

Perturbation defined by \( V : \mathbb{R}^3 \to \mathbb{R}^3 \):

\[
\begin{cases}
\dot{x} = V(x) \\
x(0) = X
\end{cases}
\quad \Rightarrow \quad x = x(X,t) = T_t(X)
\]

\( \Gamma \quad \Rightarrow \quad \text{PBE: } \psi \quad \Rightarrow \quad G_{elec}[\Gamma] \)

\( \Gamma_t \quad \Rightarrow \quad \text{PBE: } \psi_t \quad \Rightarrow \quad G_{elec}[\Gamma_t] \)

\( \delta_\Gamma G_{elec}[\Gamma] = \frac{\partial}{\partial t} G_{elec}[\Gamma_t] \)

Theorem.

\[
\delta_\Gamma G_{elec}[\Gamma] = \frac{\varepsilon_0}{2} \left( \frac{1}{\varepsilon_m} - \frac{1}{\varepsilon_w} \right) |\varepsilon \partial_n \psi|^2 + \frac{\varepsilon_0}{2} (\varepsilon_w - \varepsilon_m) |(I - n \otimes n) \nabla \psi|^2 + B(\psi).
\]
**The Coulomb-field approximation (CFA):**  \( \vec{D}_2 \approx \vec{D}_1 \)

**Electric field:**  \( \vec{E} = -\nabla \psi \)  

**Electric displacement:**  \( \vec{D} = \varepsilon \varepsilon_0 \vec{E} \)

**Electrostatic free energy**

\[ G_{elec}[\Gamma] = \int \frac{1}{2} \vec{D}_2 \cdot \vec{E}_2 dV - \int \frac{1}{2} \vec{D}_1 \cdot \vec{E}_1 dV = \frac{1}{32 \pi^2 \varepsilon_0} \left( \frac{1}{\varepsilon_w} - \frac{1}{\varepsilon_m} \right) \int_{\Omega_w} \sum_{i=1}^{N} \frac{Q_i (\vec{r} - \vec{r}_i)}{\left| \vec{r} - \vec{r}_i \right|^3} dV \]

**Dielectric boundary force**

\[ -\delta_\Gamma G_{elec}[\Gamma](\vec{r}) = \frac{1}{32 \pi^2 \varepsilon_0} \left( \frac{1}{\varepsilon_w} - \frac{1}{\varepsilon_m} \right) \left| \sum_{i=1}^{N} \frac{Q_i (\vec{r} - \vec{r}_i)}{\left| \vec{r} - \vec{r}_i \right|^3} \right|^2 \]
The Yukawa-field approximation (YFA)

Electrostatic free energy

\[ G_{elec}[\Gamma] = \frac{1}{32\pi^2\varepsilon_0} \int_{\Omega_w} \left( \frac{1}{\varepsilon_w} \sum_{i=1}^{N} f_i(\mathbf{r},\kappa,\Gamma) \frac{Q_i(\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3} \right)^2 - \frac{1}{\varepsilon_m} \sum_{i=1}^{N} \frac{Q_i(\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3} dV \]

\[ f_i(\mathbf{r},\kappa,\Gamma) = \frac{1 + \kappa |\mathbf{r} - \mathbf{r}_i|}{1 + \kappa |\mathbf{r}_i - P_i(\mathbf{r})|} \exp(-\kappa(\mathbf{r} - P_i(\mathbf{r})) \]

Dielectric boundary force

\[ -\delta_{\Gamma} G_{elec}[\Gamma] : \Gamma \rightarrow R \]
A unified variational theory

Electrostatic free-energy functional of ionic concentrations

\[ F[c] = \int \left[ \frac{1}{2} \rho \psi + \beta^{-1} Q(c) - \sum_{i=1}^{M} \mu_i c_i \right] dV \]

Charge density

\[ \rho = \rho_f + \sum_{i=1}^{M} q_i c_i \]

Poisson’s Eq.

\[ \nabla \cdot \epsilon \epsilon_0 \nabla \psi = -\rho \]

Without size effect

\[ Q(c) = \sum_{i=1}^{M} c_i \ln(\Lambda^3 c_i) \]

With size effect

\[ Q(c) = \sum_{i=0}^{M} c_i \ln(a_i^3 c_i) \]

Sonlvent concentration

\[ a_0^3 c_0 = 1 - \sum_{i=1}^{M} a_i^3 c_i \]
\( \delta_i F[c] = 0 \iff \text{Boltzmann distributions} \)

Without size effect

\[ c_i = c_i^\infty e^{-\beta q_i \psi} \]

With a uniform size

\[ c_i = \frac{c_i^\infty e^{-\beta q_i \psi}}{1 + a^3 \sum_{j=1}^{M} c_j^\infty e^{-\beta q_j \psi}} \]

With nonuniform sizes

\[ \left( \frac{a_i}{a_0} \right)^3 \ln(a_0^3 c_0) - \ln(a_i^3 c_i) = \beta(q_i \psi - \mu_i) \]

No explicit formulas are available!

**A constrained optimization method:** minimize

\[ F[\psi, c] = \int \left[ \frac{1}{2} \rho \psi + \beta^{-1} Q(c) \right] dV \]

Subject to Poisson’s equation and mass conservations.
4. The Level-Set Method

- Interface motion
  \[ V_n = V_n(\vec{r}, t) \quad \text{for} \quad \vec{r} \in \Gamma(t) \]

- Level-set representation
  \[ \Gamma(t) = \{ \vec{r} \in \Omega : \varphi(\vec{r}, t) = 0 \} \]

- The level-set equation
  \[ \varphi_t + V_n | \nabla \varphi | = 0 \]

\[ \left[ \begin{array}{c} \varphi(\vec{r}(t), t) = 0 \\ \nabla \varphi \cdot \vec{r}_t = \left( \frac{\nabla \varphi}{|\nabla \varphi|} \right) \cdot \vec{r}_t \end{array} \right] \Rightarrow \left[ \begin{array}{c} \varphi_t + \nabla \varphi \cdot \vec{r}_t = 0 \\ \left( \frac{\nabla \varphi}{|\nabla \varphi|} \right) \cdot \vec{r}_t = V_n | \nabla \varphi | \end{array} \right] \]
Examples of normal velocity

- Geometrically based motion
  - Motion by mean curvature: $V_n = -H$
  - Surface diffusion: $V_n = \Delta_s H$
  - Willmore flow: $V_n = \Delta_s H + H(H^2 - K)$

- External field

\[
\begin{cases}
  u_t - \Delta u = 0 & \text{in } \Omega_- \cup \Omega_+ \\
  u = -H & \text{on } \Gamma \\
  \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \\
  V_n = \left[\frac{\partial u}{\partial n}\right] & \text{on } \Gamma
\end{cases}
\]
Level-set formulas of geometrical quantities

- **Unit normal**
  \[ \vec{n} = \frac{\nabla \varphi}{|\nabla \varphi|} \]

- **Mean curvature**
  \[ H = \frac{1}{2} \nabla \cdot \vec{n} \]

- **Gaussian curvature**
  \[ K = \vec{n} \cdot \text{adj}(He(\varphi))\vec{n} \]

- **Surface integral**
  \[ \int_{\Gamma} f(\overline{r})dS = \int_{\mathbb{R}^3} f(\overline{r})\delta(\varphi)dV \]

- **Volume integral**
  \[ \int_{\Omega} f(\overline{r})dV = \int_{\mathbb{R}^3} f(\overline{r})[1 - H(\varphi)]dV \]

Topological changes – motion with no emotion!
Application to variational solvation

\[ \varphi_t + V_n \left| \nabla \varphi \right| = 0 \]

\[ \frac{d\vec{r}_i}{dt} = -\nabla_{\vec{r}_i} H[\Gamma; \vec{r}_1, \ldots, \vec{r}_N] = -\nabla_{\vec{r}_i} V[\vec{r}_1, \ldots, \vec{r}_N] - \nabla_{\vec{r}_i} G[\Gamma] \]

\[ V_n = -\delta_\Gamma H[\Gamma; \vec{r}_1, \ldots, \vec{r}_N] = -\delta_\Gamma G[\Gamma] \]

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

\[ \delta_\Gamma \int_{\Omega} dV = 1 \quad \delta_\Gamma \int_{\Gamma} dS = -2H \quad \delta_\Gamma \int_{\Gamma} HdS = -K \]
Discretization of the level-set equation

\[ \varphi_t + V_n | \nabla \varphi | = 0 \]
\[ V_n = -P - 2\gamma_0 [H(\vec{r}) - \tau K(\vec{r})] + \rho_w U(\vec{r}) \]

- Special case: \( \tau = 0 \)

Semi-implicit \( \varphi_t = 2\gamma_0 \Delta \varphi + N(\nabla \varphi, \nabla^2 \varphi) \)

Central differencing + FFT or Cholesky

- General case: \( \tau > 0 \)

Forward Euler \( \varphi^{k+1}(x) - \varphi^k(x) = -\Delta t V^k_n(x) | \nabla \varphi^k(x) | \)

Decomposition \( \varphi_t = A + B \)

Central differencing \( A = 2\gamma_0 [H(\vec{r}) - \tau K(\vec{r})] | \nabla \varphi | \)

Upwinding \( B = [P - \rho_w U(\vec{r})] | \nabla \varphi | \)
Convergence test on a single charged particle

\[ G(R) = 4\pi (R^2 - 2\tau R) + 16\pi \rho_w \left( \frac{\sigma^{12}}{9R^9} - \frac{\sigma^6}{3R^3} \right) - \frac{Q^2}{8\pi \varepsilon_0} \left( \frac{1}{\varepsilon_m} - \frac{1}{\varepsilon_w} \right) \]
New level-set techniques

- Pre-computation of the potential
- Numerical regularization
- Fast numerical integration
- Local level-set method

Efficiency

- 4,000 solute atoms, 50x50x50 grids, a good initial guess → 5 minutes
- 4,000 solute atoms, high resolution, a bad initial guess → 20 – 60 minutes
- Dynamics: a different situation
5. Numerical Results

Parameters

- Pressure: $P$ (bar)
- Surface tension: $\gamma_0$ ($k_B T$)
- Tolman length: $\tau$ ($\text{Å}$)
- Water density: $\rho_w$ ($\text{Å}^{-3}$)
- LJ parameters: $\sigma$ ($\text{Å}$), $\varepsilon$ ($k_B T$)
- Point charges: $Q_i$ ($e$)

Free energy vs. optimization steps
Two xenon atoms

PMF: Level-set (circles) vs. MD (solid line).

Two paraffin plates

PMF: Level-set (circles) vs. MD (line).

Helical alkanes and C60 fullerene
A hydrophobic receptor-ligand system

Each wall consists of 4,242 atoms.

System setup for the level-set VISM calculation.
Different initial surfaces lead to different local minima.
PMF

wall-particle distance
A benzene molecule
Two charged paraffin plates

Plate-plate separation \( d = 10 \text{ Å} \).
Left: no charges. Middle: partial charges (0.2 e, 0.2 e).
Right: partial charges (0.2 e, -0.2 e).
Color code represents mean curvature.
6. Conclusions

**Accomplishments**

- **Theory**
  - Variational implicit-solvent models (VISM)
  - Coupling VISM with solute molecular mechanics
  - PB theory and effective dielectric boundary forces
  - Coulomb-field and Yukawa-field approximations
  - Continuum models of nonuniform ionic size effects

- **Computation**
  - A level-set method for VISM coupled with solute molecular mechanics, including charges with CFA.
  - Capturing hydrophobic cavities, charge effect, multiple states, etc. Agree with MD simulations.
  - Constrained optimization for size effects.
Some Issues

► Models: Fitting parameter $\tau$? Well-posedness?
► Level set: Gaussian curvature? Speed up?
► More chemistry/physics: Hydration shell? Real dynamics?

Current and future work

► Monte Carlo level-set VISM: searching local minima
► Solvent dynamics: hydrodynamics + fluctuation
► Multiscale approach for protein folding: solute MD + solvent dynamics
► Molecular recognition + drug design: host-guest systems
► Fast algorithms, GPU computing, software development
► Mathematics and statistical mechanics of VISM
Thank you!