Non-bonded interactions

Continuum solvation models

Marcus Elstner and Tomáš Kubař

May 25, 2016

Continuum electrostatics methods

Situation up to now

- molecules in an explicit solvent
- all interactions obviously involved
- polarizability / permittivity of the solvent considered implicitly
- for instance, solvation free energy involved implicitly
  - if desired, may be evaluated with special methods

Continuum electrostatics methods

Example – polypeptide in the α-helix and β-sheet conformations.
The free energy difference of the two structures is given by

- the difference of internal energies / enthalpies
- the entropic contributions – above all vibrational entropy
- the difference of free energies of solvation

Continuum electrostatics methods

Example – polypeptide in the α-helix and β-sheet conformations.
The free energy difference of the two structures is given by

- the difference of internal energies / enthalpies
- the entropic contributions – above all vibrational entropy
- the difference of free energies of solvation

α-helix: much larger dipole moment than β-sheet
$\rightarrow$ α-helix is better solvated in a polar medium (H_2O)
$\rightarrow$ crucial effect of solvation on the equilibrium
between conformations of solvated peptide

Continuum electrostatics methods

Example – polypeptide in the α-helix and β-sheet conformations.
The free energy difference of the two structures is given by

- the difference of internal energies / enthalpies
- the entropic contributions – above all vibrational entropy
- the difference of free energies of solvation

α-helix: much larger dipole moment than β-sheet
$\rightarrow$ α-helix is better solvated in a polar medium (H_2O)
$\rightarrow$ crucial effect of solvation on the equilibrium
between conformations of solvated peptide

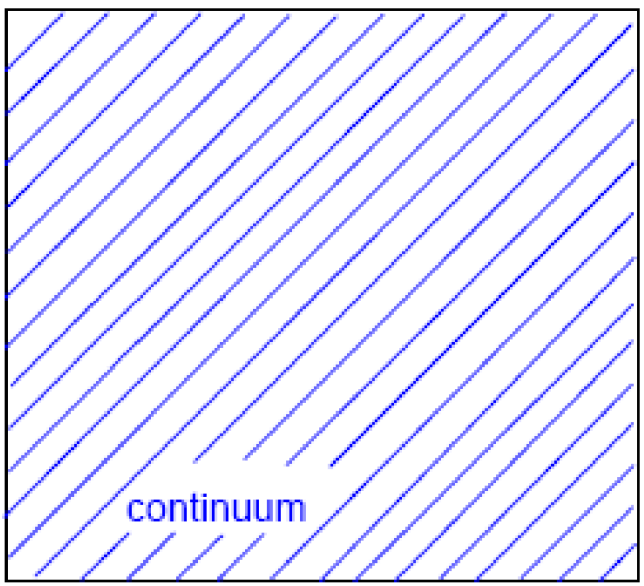
Motivation: the amount of solvent becomes excessive easily, so maybe meaningful – abandon explicit solvent, apply implicit model

Continuum electrostatics methods

Solvation free energy: $\Delta G_{\text{solv}} = \Delta G_{\text{cav}} + \Delta G_{\text{vdW}} + \Delta G_{\text{ele}}$

- A cavity in the solvent is formed
  – rearrangement of the solvent molecules

  ΔG_{cav}: decrease of S and loss of solvent–solvent interactions
- solute–solvent interaction – van der Waals and electrostatic

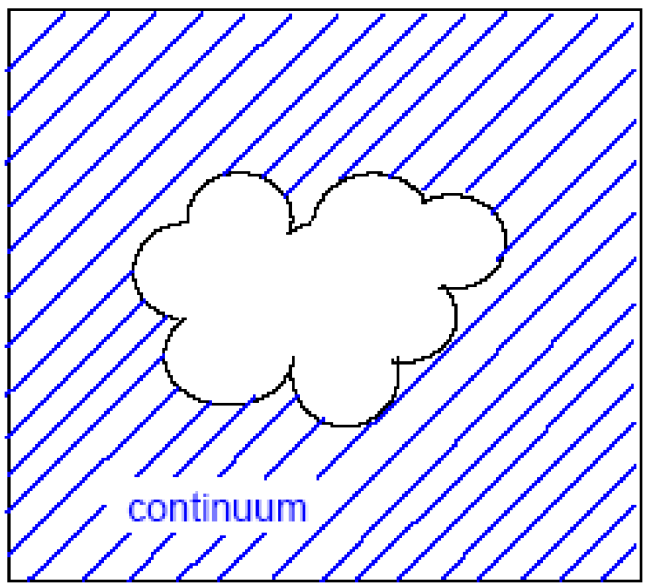


Continuum electrostatics methods

Solvation free energy: $\Delta G_{\text{solv}} = \Delta G_{\text{cav}} + \Delta G_{\text{vdW}} + \Delta G_{\text{ele}}$

- A cavity in the solvent is formed
  – rearrangement of the solvent molecules

  ΔG_{cav}: decrease of S and loss of solvent–solvent interactions
- solute–solvent interaction – van der Waals and electrostatic

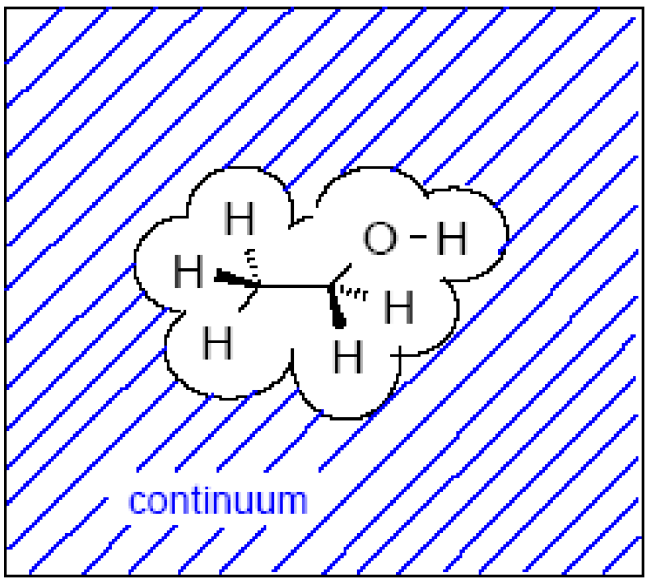


Continuum electrostatics methods

Solvation free energy: $\Delta G_{\text{solv}} = \Delta G_{\text{cav}} + \Delta G_{\text{vdW}} + \Delta G_{\text{ele}}$

- A cavity in the solvent is formed
  – rearrangement of the solvent molecules

  ΔG_{cav}: decrease of S and loss of solvent–solvent interactions
- solute–solvent interaction – van der Waals and electrostatic



Solvent-accessible surface area

SASA – important concept

- solvent-exposed surface of molecule as a solid body
- reasonable approx.: ΔG_{cav} and ΔG_{vdW} proportional to SASA.
- total surface composed from surfaces of individual atoms S_i
- Then: $\Delta G_{\text{cav}} + \Delta G_{\text{vdW}} = \sum_i c_i \cdot S_i$
- alternative: obtain SASA by rolling a ball of a certain diameter (typically 2.8 Å to mimic H_2O) on the molecular surface

Solvent-accessible surface area

When does it work?

- if the electrostatic effect of the surrounding solvent dominates (shielding of solvent-exposed charged side chains of proteins)
- not if there is specific solute–solvent interaction (like hydrogen bonding)

Difficult example: dynamics of small peptides dissolved in water
– competition between various hydrogen-bonding patterns

Continuum electrostatics methods

Big question: how to calculate ΔG_{ele}?

often used is the term "reaction field"

$$\Delta G_{\text{ele}} = q \cdot \Phi_{\text{rf}}(\vec{r})$$

for moving the cavity with the solute from vacuo to the solvent

Born and Onsager models

Born: the work to bring charge q from vacuo into spherical cavity of radius a in solvent with dielectric constant ε:

$$\Delta G_{\text{ele}} = -\frac{q^2}{2a}\left(1 - \frac{1}{\varepsilon}\right)$$

ε: 1 for vacuo (thus $\Delta G_{\text{ele}} = 0$), 80 for water, 2 to 20 for protein

Onsager and Kirkwood: model for dipole μ in cavity

$$\begin{aligned} \Phi_{\text{rf}} &= \frac{2(\varepsilon - 1)}{2\varepsilon + 1} \cdot \frac{1}{a^3} \cdot \mu \\ \Delta G_{\text{ele}} &= -\frac{1}{2}\Phi_{\text{rf}} \cdot \mu \end{aligned}$$

Born and Onsager models

- simple models
- implemented in many standard programs
- quite uneralistis approximations even for small molecules

Born and Onsager models

- simple models
- implemented in many standard programs
- quite uneralistis approximations even for small molecules

Extensions:

- polarizable continuum model (PCM) – arbitrary surfaces constructed with the use of vdW radii of individual atoms
- conductor-like screening models (COSMO) – polarization of the dielectric (insulating) solvent derived from scaled-conductor approximation.

Poisson–Boltzmann equation (PBE)

For big molecules, the simple models may be too simple and inefficient at the same time :-(

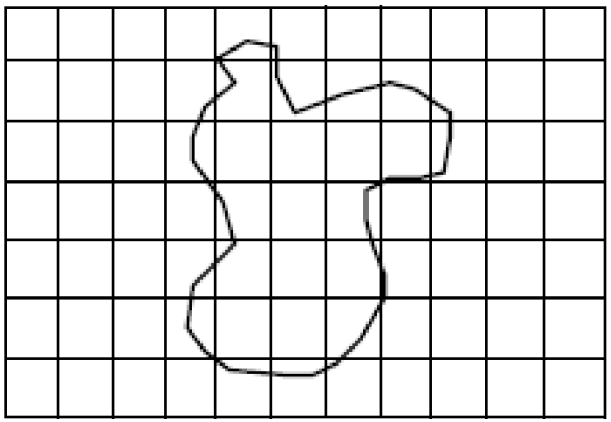
other approximations – starting from Poisson's equation

$$\nabla \varepsilon \nabla \Phi = -4\pi\rho$$

given – charge distribution ρ and dielectric constant ε
to be found – potential Φ

possibility to solve:

- discretize on a 3D grid,
  use finite differences
  calc. Φ on every grid point
  iteratively



Ions in the solvent

ions are very important – counterions compensate charged solute, or salt mimicks physiologic conditions

the position of ions depends on the potential:

$$\rho_{\text{ions}} = \sum_i q_i \cdot c_i \cdot \exp\left[-\frac{q_i \cdot \Phi(r)}{k_B T}\right]$$

or: anions like to be where $\Phi > 0$, and cations like $\Phi < 0$

Ions in the solvent

ions are very important – counterions compensate charged solute,
or salt mimicks physiologic conditions

the position of ions depends on the potential:

$$\rho_{\text{ions}} = \sum_i q_i \cdot c_i \cdot \exp\left[-\frac{q_i \cdot \Phi(r)}{k_B T}\right]$$

or: anions like to be where $\Phi > 0$, and cations like $\Phi < 0$

an additional term appears in Poisson's equation:
linearized Poisson–Boltzmann equation at low ionic strength:

$$\nabla \varepsilon \nabla \Phi = -4\pi\rho + \varepsilon \cdot \kappa^2 \cdot \Phi(r)$$

with the Debye–Hückel parameter $\kappa^2 = \frac{8\pi q^2 I}{\varepsilon \cdot k_B T}$
(ionic strength $I = \frac{1}{2}\sum_i c_i z_i^2$, c_i concentration, z_i charge of ion i)

Ions in the solvent – PBE

- charge distribution on the protein
  polarizes the dielectric outside ("solvent")
  $\rightarrow$ screening of any solvent-exposed charges of the protein
  effectively, charges pointing into the solvent will vanish nearly
- solvent ions will distribute to make
  the overall charge distribution more uniform
  if a negative charge points into the solvent,
  a cation will be located close to it

The solvent around a protein should always be taken into account.

PBE – not efficient enough to be calculated in every MD step
(approximations are necessary)

Generalized Born model (GB)

idea – use the simple Born equation for MM atomic charges

$$\Delta G^1_{\text{ele}} = -\left(1 - \frac{1}{\varepsilon}\right) \sum_i \frac{q_i^2}{2a_i}$$

the interaction of individual charges changes in solution

$$\begin{aligned} E_{\text{ele}} &= \frac{1}{2} \sum_{i \neq j} \frac{1}{\varepsilon} \frac{q_i \cdot q_j}{r_{ij}} = \\ &= \frac{1}{2} \sum_{i \neq j} \frac{q_i \cdot q_j}{r_{ij}} - \frac{1}{2} \left(1 - \frac{1}{\varepsilon}\right) \sum_{i \neq j} \frac{q_i \cdot q_j}{r_{ij}} \end{aligned}$$

Generalized Born model (GB)

idea – use the simple Born equation for MM atomic charges

$$\Delta G^1_{\text{ele}} = -\left(1 - \frac{1}{\varepsilon}\right) \sum_i \frac{q_i^2}{2a_i}$$

the interaction of individual charges changes in solution

$$\begin{aligned} E_{\text{ele}} &= \frac{1}{2} \sum_{i \neq j} \frac{1}{\varepsilon} \frac{q_i \cdot q_j}{r_{ij}} = \\ &= \frac{1}{2} \sum_{i \neq j} \frac{q_i \cdot q_j}{r_{ij}} - \frac{1}{2} \left(1 - \frac{1}{\varepsilon}\right) \sum_{i \neq j} \frac{q_i \cdot q_j}{r_{ij}} \end{aligned}$$

giving another contribution to solvation free energy

$$\Delta G^2_{\text{ele}} = -\frac{1}{2} \left(1 - \frac{1}{\varepsilon}\right) \sum_{i \neq j} \frac{q_i \cdot q_j}{r_{ij}}$$

solvation free energy $= \Delta G^1_{\text{ele}} + \Delta G^2_{\text{ele}}$

Generalized Born model (GB)

problem 1 – Born's formula holds
for interaction of charges located in spherical cavities (with radii a_i)
– only valid for charged bodies of general shapes if $r_{ij} \gg a_i + a_j$
– two extreme cases are covered:

$$E = \begin{cases} \dfrac{q_i^2}{a_i}, & \text{if } i = j \text{ ('self-interaction, i.e. solvation energy)} \\ \dfrac{q_i \cdot q_j}{r_{ij}}, & \text{if } i \neq j \text{ and } r_{ij} \to \infty \end{cases}$$

what to do at intermediate distances (2 Å to 10 Å)?

Generalized Born model (GB)

problem 1 – Born's formula holds
for interaction of charges located in spherical cavities (with radii a_i)
– only valid for charged bodies of general shapes if $r_{ij} \gg a_i + a_j$
– two extreme cases are covered:

$$E = \begin{cases} \dfrac{q_i^2}{a_i}, & \text{if } i = j \text{ ('self-interaction, i.e. solvation energy)} \\ \dfrac{q_i \cdot q_j}{r_{ij}}, & \text{if } i \neq j \text{ and } r_{ij} \to \infty \end{cases}$$

what to do at intermediate distances (2 Å to 10 Å)? interpolate!

$$f(r_{ij}) = \sqrt{r_{ij}^2 + a_i a_j \exp\left[-\frac{r_{ij}^2}{4 a_i a_j}\right]}$$

$$\Delta G_{\text{ele}} = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right) \cdot \sum_{i,j} \frac{q_i \cdot q_j}{f(r_{ij})}$$

Generalized Born model (GB)

Born's equation holds for a charged particle in contact with solvent

problem 2 – many charges are buried deeply inside the protein, far from the solvent!
$\rightarrow$ solvation free energy may be overestimated heavily

possible solution – scale up a_i in a reasonable way!

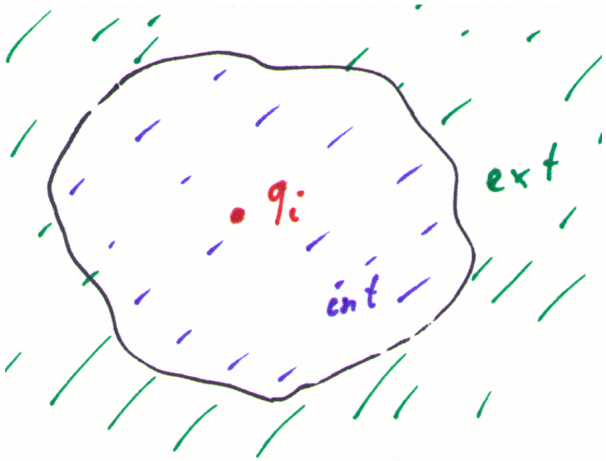
the most important task when using the GB method – to use/calculate reasonable radii a_i

How to get the radii in GB

approximate interaction energy of a charge q_i in the protein interior with the solvent:

$$\Delta G^i_{\text{ele}} = -\frac{1}{8\pi}\left(1 - \frac{1}{\varepsilon_W}\right)\int_{\text{ext}} \frac{q_i^2}{r^4}\,\mathrm{d}V$$

integration runs over
the 'exterior' of the protein

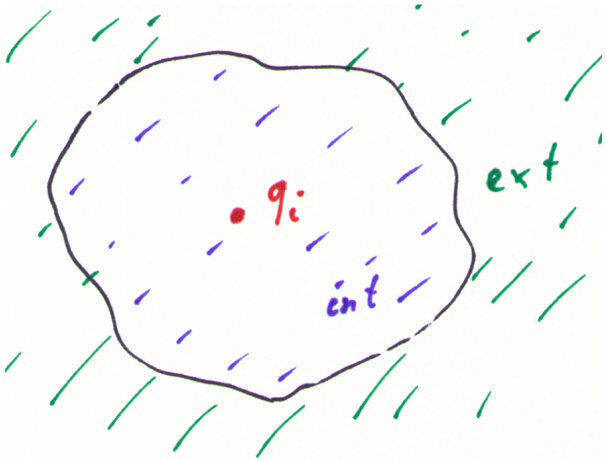


How to get the radii in GB

approximate interaction energy of a charge q_i in the protein interior with the solvent:

$$\Delta G^i_{\text{ele}} = -\frac{1}{8\pi}\left(1-\frac{1}{\varepsilon_W}\right)\int_{\text{ext}}\frac{q_i^2}{r^4}\,\mathrm{d}V$$

integration runs over the 'exterior' of the protein



comparing with the Born formula, we find

$$\frac{1}{a_i} = \frac{1}{4\pi}\int_{\text{ext}}\frac{1}{r^4}\,\mathrm{d}V$$

r – distance from the charge to the 'boundary' of the protein.

How to get the radii in GB

several GB models exist; generally, $\int_{\text{ext}}$ transformed to $\int_{\text{int}}$

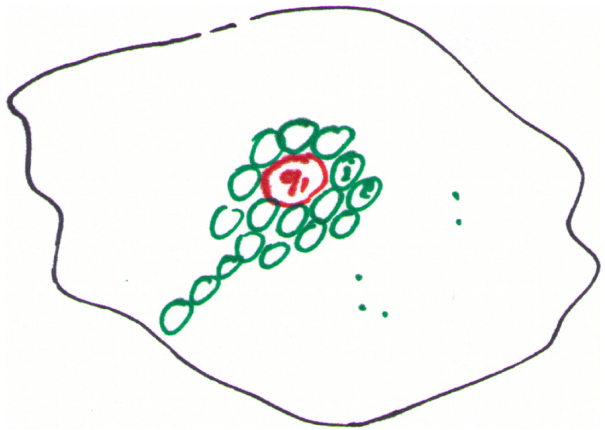
- GB molecular volume – with van der Waals radius α_i:

$$\frac{1}{a_i} = \frac{1}{\alpha_i} - \frac{1}{4\pi} \int_{\text{int}, r>\alpha_i} \frac{1}{r^4} \, \mathrm{d}V$$

  – possibly longish calculation time

- pairwise models – the interior $\approx$ union of atomic spheres

$$\frac{1}{a_i} = \frac{1}{\alpha_i} - \sum_{j \neq i} \frac{1}{4\pi} \int_{\text{sphere } j} \frac{1}{r^4} \, \mathrm{d}V$$



How to get the radii in GB

several GB models exist; generally, $\int_{\text{ext}}$ transformed to $\int_{\text{int}}$

- GB molecular volume – with van der Waals radius α_i:

$$\frac{1}{a_i} = \frac{1}{\alpha_i} - \frac{1}{4\pi} \int_{\text{int}, r>\alpha_i} \frac{1}{r^4} \, \mathrm{d}V$$

  – possibly longish calculation time

- pairwise models – the interior $\approx$ union of atomic spheres
  this is insufficient because of partial overlap / void places
  empirical formula may be used instead:

$$a_i^{-1} = \frac{1}{\lambda \cdot R_{\text{vdW},i}} - P_1 \frac{1}{R_{\text{vdW},i}^2} - \sum_j^{\text{bond}} \frac{P_2 V_j}{r_{ij}^4} - \sum_j^{\text{angle}} \frac{P_3 V_j}{r_{ij}^4}$$

$$- \sum_j^{\text{nonbond}} \frac{P_4 V_j}{r_{ij}^4} \cdot \text{CCF}(P_5, r_{ij})$$

MM-PBSA

- another application of implicit solvent models
- free energies of binding of ligands to biomolecules
- post-processing approach to evaluate free energies
- a normal MD simulation is run,
  and free energies are computed a posteriori

binding free energy obtained component-wise with various methods
solvation free energy – with Poisson–Boltzmann or so
non-polar contribution – SASA-dependent terms
configurational entropy – normal-mode analysis

very approximative, yet may still give results of good quality