

THE UNIVERSITY of TEXAS

HEALTH SCIENCE CENTER AT HOUSTON SCHOOL of HEALTH INFORMATION SCIENCES

Continuum Electrostatics

For students of HI 6001-100 "Biomolecular Modeling"

Willy Wriggers, Ph.D. School of Health Information Sciences

http://biomachina.org/courses/modeling/08.html

Effect of Charges in Biology

- Mg2+ binding to RNA or DNA
- Zn2+ binding in gene regulation
- Ca2+ binding in signal transduction (calmodulin etc.)
- signal transduction through phosphorylation (Tyr, Ser, His)
- ions form organizing centers for protein folding
- steering of protein assembly
- formation of lipid bilayers (membranes)
- etc...

Electrostatic Potential Contours of Mouse Acetylcholinesterase



McCammon Group - UCSD

Electrostatic Potentials and Fields

- Electrostatic interactions are very long-ranged (recall the 1/r dependence of the Coulombic term in the MM energy function).
- The electrostatic potential is a scalar quantity, i.e. it has no direction.
- Suppose we place a charged particle into an electric field. The electrostatic potential is the quantity that when multiplied by the charge on the particle tells us the energy required to place the particle in the field.
- The electrostatic field is a vector that tells us the gradient of the electrostatic potential. When multiplied by the charge on the particle it tells us the force acting on the particle.

Coulomb Potential



• Electrostatic Potential:

$$\phi(\mathbf{r}) = \sum_{i=1}^{charges} rac{q}{arepsilon \mid \mathbf{r} - \mathbf{r}'_i \mid}$$

• Electrostatic work required or gained to bring a charge q^\prime to point ${\bf r}_{q^\prime}$ in the potential

$$W = q' \phi(\mathbf{r}_{q'})$$

Dieelectric Screening

Permanent Molecular Dipoles

For instance water: 1.9 Debye

Dipolar molecules orient in an electrostatic field



© G. Matthias Ullmann

Contributions Inside Molecule



A Molecule In Solution

 In MD, screening is sometimes modeled implicitly by distancedependent dielectric (1/r dependence of ε, 1/r² term in the MM energy function). See earlier notes.

Continuum electrostatics:

- Inside protein $\varepsilon_p \sim 2-4$.
- Outside molecule $\varepsilon_{w} \sim 60-80$ (solvent)



© G. Matthias Ullmann

Continuum Electrostatics

Conceptual Model:

Protein: Low dielectric region With fixed partial charges

Solvent: High dielectric region with unlocalized (mobile) charges



A continuum electrostatic model describes molecules at atomic detail using a macroscopic description.

Poisson Equation

One of the fundamental equations of classical electrostatics

Electrostatic Field: $\mathbf{E}(\mathbf{r}) = -\nabla \phi(\mathbf{r})$

 ∇ – differential operator: $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$

Poisson Equation equation in vacuum (Gauß Theorem):

$$\nabla \mathbf{E}(\mathbf{r}) = 4\pi \rho(\mathbf{r})$$

A dielectric medium screens the field.

$$\nabla \left[\varepsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}) \right] = 4\pi \rho(\mathbf{r})$$

- $\phi(\mathbf{r})$: electrostatic potential
- $\varepsilon(\mathbf{r})$: relative permittity ("dielectric constant")

Poisson Equation:

$$\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] = -4\pi \rho(\mathbf{r})$$

 $\rho(\mathbf{r})$: charge density

Ionic Distribution

Mobile Ions are distributed according to a Boltzmann statistic. Mean Concentration at ${\bf r}$



Poisson-Boltzmann Equation

Poisson Equation:

 $\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] = -4\pi \rho(\mathbf{r})$

Charge Distribution

$$\rho(\mathbf{r}) = \rho_{prot}(\mathbf{r}) + \rho_{ions}(\mathbf{r})$$

$$\rho_{ions}(\mathbf{r}) = \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{-Z_i e_0 \phi(\mathbf{r})}{RT}\right)$$

Poisson-Boltzmann Equation

$$\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] = -4\pi \left(\rho_{prot}(\mathbf{r}) + \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{-Z_i e_0 \phi(\mathbf{r})}{RT}\right) \right)$$

© G. Matthias Ullmann

Poisson-Boltzmann Equation

Poisson Equation:

 $\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] = -4\pi \rho(\mathbf{r})$

Charge Distribution

$$\rho(\mathbf{r}) = \rho_{prot}(\mathbf{r}) + \rho_{ions}(\mathbf{r})$$

$$\rho_{ions}(\mathbf{r}) = \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{-Z_i e_0 \phi(\mathbf{r})}{RT}\right)$$

Poisson-Boltzmann Equation

$$\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] = -4\pi \left(\rho_{prot}(\mathbf{r}) + \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{-Z_i e_0 \phi(\mathbf{r})}{RT}\right) \right)$$

© G. Matthias Ullmann

Linearized Poisson-Boltzmann Equation

Linearization for $(\phi(\mathbf{r})/RT < 1)$:

$$\sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{-Z_i e_0 \phi(\mathbf{r})}{RT}\right) \approx \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 - \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2 e_0^2 \frac{\phi(\mathbf{r})}{RT}$$
$$\sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_0 = 0 \quad \text{Charge Balance}$$
$$\nabla [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] = -4\pi \left(\rho_{prot}(\mathbf{r}) - \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2 e_0^2 \frac{\phi(\mathbf{r})}{RT}\right)$$
$$\text{Define}: \quad I = \frac{1}{2} \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2; \qquad \bar{\kappa}^2 = \frac{8\pi N_A e_0^2 I}{k_B T}$$
$$\nabla [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] = -4\pi \rho_{prot}(\mathbf{r}) + \bar{\kappa}^2(\mathbf{r}) \phi(\mathbf{r})$$

Properties of Solutions of Linear PBE

If $\rho_1 + \rho_2 = \rho$ then $\phi(\mathbf{r}, \rho_1) + \phi(\mathbf{r}, \rho_2) = \phi(\mathbf{r}, \rho)$



Discrete Model

The boundary molecule/solvent is smoothed.

Charges, dielectic constant, and ionic strength are mapped to the grid.



Assigning Charges to a Grid



Tri-Linear Interpolation

$$q_{grid} = q(1-\frac{a}{h})(1-\frac{b}{h})(1-\frac{c}{h})$$

Problem!! – The splitted charges interact with each other. Grid Artefact!

The Grid Artefact cancels in energy differences!

Numerical Solution (Finite Difference)

 $\nabla \left[\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] - \bar{\kappa}^2(\mathbf{r}) \phi(\mathbf{r}) + 4\pi \rho_{prot}(\mathbf{r}) = 0$

Integration over grid voxels:

$$\int \nabla [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] \, d\mathbf{r} - \int \bar{\kappa}^2(\mathbf{r}) \phi(\mathbf{r}) \, d\mathbf{r} + 4\pi \int \rho_{prot}(\mathbf{r}) \, d\mathbf{r} = 0$$
(Gauss theorem) $\longrightarrow \int [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] \, d\mathbf{A} - h^3 \bar{\kappa}_0^2 \phi_0 + 4\pi q_0 = 0$
Surface Integral: $\int [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] \, d\mathbf{A} = \sum_{i=1}^6 h \varepsilon_i (\phi_i - \phi_0)$

$$\sum_{i=1}^6 h \varepsilon_i (\phi_i - \phi_0) - h^3 \bar{\kappa}_0^2 \phi_0 + 4\pi q_0 = 0$$

$$\phi_0 = \frac{\left(\sum_{i=1}^6 h \varepsilon_i \phi_i\right) + 4\pi q_0}{\left(\sum_{i=1}^6 h \varepsilon_i\right) + h^3 \bar{\kappa}_0^2}$$

© G. Matthias Ullmann

Numerical Solution (Finite Difference)



In the nth iteration:



Iteration until selfconsistency.

Focusing/Boundary Conditions



- outer grid boundary condition from an analytical solution (Debye-Hückel theory, Kirkwood - spherical molecules with charges, Born model)
- inner grids initialized from grid one level up



PBE FlowChart

- an analytical solution is needed for the potential at the outer boundary
- alternatively, periodic boundary conditions could be used
- different numerical procedures exist (for instance boundary element method)
- different programs exist: Delphi,
 MEAD, UHBD, Charmm Module,
 GRASP

GRASP

- http://trantor.bioc.columbia.edu/grasp/
- free for academic use
- currently runs on SGI only





DelPhi

- Developed, along with Grasp, by Barry Honig's group, now at Columbia
- Difference between DelPhi and GRASP:
 - Grasp was intended to be an interactive molecular graphics program with a very rough PDB solver. Uses a 32x32x32 grid size.
 - DelPhi is intended for quantitative analysis, and therefore is more robust. Uses a 65x65x65 grid size.
- For the most accurate figures, use DelPhi to solve the PB-equation then use a visualization program to create images.

DelPhi Input File

```
gsize=165
scale=2.5
in(pdb,file="bb_cmplx_h.pdb")
in(siz,file="charm22.siz")
in(crg,file="charm22.crg")
acenter(28.114,40.477,9.909)
indi=2.0
exdi=80.0
prbrad=1.4
salt=0.10
ionrad=2.0
bndcon=4
maxc=0.0001
linit=400
!nonit=800
energy(s,c,g)
```

© John Kerrigan

Radius File

!my siz based on PARSE ! (value for P taken from Pauling, ! for Mg from Biophys J 2001, 80, 1151) atom res radius 1.4 Ο 1.0 Η С 1.7 1.5 Ν 1.85 S Ρ 1.90 0.99 Mg

© John Kerrigan

DelPhi Charge File

!										
!	! Delphi charge file generated from CHARMM									
!	! top22.pro									
!	(C) 1995 Andreas Windemuth								
atom	atom resnumbc charge									
N	ALA	-0.470								
HN	ALA	0.310								
CA	ALA	0.070								
HA	ALA	0.090								
CB	ALA	-0.270								

Example DelPhi Output File

DELPHI SITE POTENTIAL FILE
grid size,percent fill:5980.00000inner,outer dielectric:2.00000080.00000ionic strength (M):0.0000000E+001.400000linear, nolinear iterations:1210boundary condition:2Data Output:COORDINATES CHARGE POTENTIALS REACTION COULOMBICtitle:qdiffxas:qdiffxs4 with an improved surfacing routine

$\begin{array}{cccccccccccccccccccccccccccccccccccc$	ATOM COO	PDTNATES	(X X 7)	CHARGE	CRID PT	PEAC PT	COUL POT
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$							
$\begin{array}{cccccccccccccccccccccccccccccccccccc$						21.8624	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	25,6800	41.3000	41.0700	0.0000	-0.7988	21.0879	-21.5502
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	25.6900	40.5000	40.1300	0.0000	-0.8230	21.5642	-22.0923
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	26.6700	42.1400	41.3400	0.0000	-0.7973	20.2470	-20.5840
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	26.7700	42.6600	42.1900	0.0000	-1.0643	22.3860	-23.0180
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	27.7900	42.2600	40.4000	0.0000	-0.3994	16.9163	-16.7950
27.9000 46.2000 39.8700 0.0000 5.9398 -19.7567 28.0624 28.8400 46.2500 38.7500 0.0000 20.0697 -37.0564 62.6851 29.7600 46.0900 39.1200 0.0000 7.3198 -25.0434 34.7500	28.4300	43.6500	40.3300	0.0000	0.0680	7.8928	-6.8582
28.8400 46.2500 38.7500 0.0000 20.0697 -37.0564 62.6851 29.7600 46.0900 39.1200 0.0000 7.3198 -25.0434 34.7500	27.4400	44.7400	39.8900	0.0000	1.5067	-7.7255	11.4788
29.7600 46.0900 39.1200 0.0000 7.3198 -25.0434 34.7500	27.9000	46.2000	39.8700	0.0000	5.9398	-19.7567	28.0624
	28.8400	46.2500	38.7500	0.0000	20.0697	-37.0564	62.6851
	29.7600	46.0900	39.1200	0.0000	7.3198	-25.0434	34.7500
28 6500 46 7600 37 5200 0 0000 81 6300 -64 7456 157 5255	28 6500	46 7600	37 5200	0 0000	81 6300	-64 7456	157 5255

Visualization of DelPhi Electrostatic Potentials

SPDV (Expasy):





http://au.expasy.org/spdbv/text/epot.htm

Visualization of DelPhi Electrostatic Potentials



http://agave.wustl.edu/apbs/doc/html/tutorial/x265.html

Visualization of DelPhi Electrostatic Potentials (cont.)

UCSF Chimera



http://www.cgl.ucsf.edu/chimera/

Questions (PBE)?

Brownian Dynamics

Brownian Dynamics = Newtonian Dynamics + Random Terms

$$\frac{d^2 \vec{r}_i(t)}{dt^2} = m_i^{-1} \vec{F}_i + m_i^{-1} \vec{R}_i - \beta_i \frac{d\vec{r}(t)}{dt}$$

In Biomolecular Simulations:

- Diffusion of Macromolecules
- Simulation of Association Processes

Molecules are treated as rigid or only semirigid macroscopic objects.



BD Simulation

• Brownian dynamics (BD) simulations can be used to simulate the diffusion and association of molecules in solution.

• Brownian motion is the random movement of solute molecules in dilute solution that results from repeated collisions with solvent molecules.

• The basic principle involved in BD simulations is similar to that involved in molecular dynamics simulations, but introduces a few new approximations that allow us to perform simulations on the microsecond timescale (remember that MD of proteins is limited to around 10 nanoseconds).

• The technique has been used to calculate the association rates of enzymes with their substrates (e.g. acetylcholinesterase with its substrate acetylcholine). For diffusion-limited enzymes, this association of the enzyme and substrate is the rate-limiting step of the reaction.

• The simulations allow us to understand how association rates are affected by mutations in the protein, and by the presence of dissolved ions such as Na+ and CI- in the solution.

Example: Fasciculin - AchE





Electrostatically Accelerated Protein-Protein Association



McCammon Group - UCSD
Overview of BD Method

• When we replace explicit solvent by an implicit representation, we must make sure that we don't neglect any important properties of the solvent.

• We have already discussed the effects of water on the electrostatic properties of molecules (i.e. its screening behaviour). We have seen how these effects can be approximated in a simplified solvent model by setting the dielectric constant appropriately (see the Electrostatics pages).

• Now, we must also take account of water's effects on the dynamic behaviour of solute molecules in solution. Water has two main effects:

• It is a viscous solvent: it exerts a frictional force on a diffusing solute, slowing it down.

• Collisions with water molecules add a random component to a solute's motion.

• By incorporating both of these effects, BD techniques allow realistic simulation of the diffusion of molecules in solution without the need to include any explicit solvent molecules.

Theory (I)

• The basic algorithm used in BD is similar to that in MD: we use the positions of our particles at time *t*, together with the forces acting on them, to estimate their positions at some later time $t + \Delta t$. However, in BD we typically use much larger time-steps (>1ps) since we don't have to worry about bond stretching etc.

• The algorithm that we use in BD is due to Ermak & McCammon. The translational behaviour of a particle is dictated by:

$r(t+\Delta t)=r(t)+DF\Delta t/kT+R$

where D is the translational diffusion constant of the particle, F is the force acting on the particle, and R is a random displacement added in to mimic the effects of collisions with solvent molecules.

Theory (II)

• The translational diffusion constant of a particle is a measure of the speed with which it diffuses through solution: the higher the diffusion constant, the faster it diffuses.

•Translation diffusion constants can be estimated using the Stokes-Einstein relationship:

 $D = kT/6\pi\eta a$

where η is the solvent viscosity and *a* is the radius of the particle, i.e. bigger particles diffuse slower. A similar expression can be used to estimate the rotational diffusion constant.

Theory (III)

• *R*, the random displacement, is dependent on *D*. *R* is obtained using a random number generator, and is required to have the following statistical properties:

 $<\!\!R\!\!> = 0$ $<\!\!R^*\!R\!\!> = 6D\Delta t$

The first expression says that the average value of the random displacement is zero. This has to be true, otherwise, even with no other forces acting on the particle, it would gradually drift in one direction, which would make no sense. The second expression ensures that the diffusive behavior of the particle is correctly reproduced (Einstein diffusion equation).

Theory (IV)

•In BD simulations, *F*, the force acting on the particles, is generally assumed to be purely electrostatic and is computed from solving the PBE numerically.

•We reject any step that causes overlap of the particles, i.e. we ask the program to pick another random number that doesn't cause overlap.

Brownian Dynamics Simulation of Enzyme-Substrate Encounter



b: start surface

q: quit surface

Calculating Association Rate Constants

We can use BD to calculate the association rate constant for an enzyme binding its substrate using the following relation:

$$k = k(b) * \beta$$

k is the association rate constant, i.e. the quantity we wish to compute and k(b) is the steady state rate at which a diffusing substrate molecule first comes within distance *b* of the enzyme. β is the probability that having come within this distance *b*, the substrate proceeds to associate with the enzyme.

The Smoluchowski Equation

The rate at which two particles come within a given separation *b* can be calculated analytically using the result obtained by Smoluchowski:

 $k(b) = 4\pi Db$

where D is the relative diffusion constant of the two particles. This is simply the sum of the diffusion constants of the enzyme and substrate - note that because the diffusion constant of the substrate is much larger than that of the enzyme, it dominates D.

Calculating k(b) is therefore easy.

Obtaining the Association Probability β

•To calculate β , we perform many separate BD trajectories.

•Each simulation starts with the substrate at a distance *b* from the enzyme. The electrostatic potential should be approximately constant over the *b*-surface.

•In principle, all we have to do now is simulate the motion of the substrate until it either binds or escapes (quit or q-surface).

•Note that some substrate molecules that pass through the *q*-surface might return and bind to the enzyme if we continued the simulation, i.e. they may not actually go on to fully escape. To account for this possibility, we have to correct our calculated value of β (see references at end of class notes).

•We define binding using a set of reaction criteria. We monitor the distance between an atom of the substrate and a point on the enzyme that defines the entrance to the active site.

•In order to obtain statistically meaningful estimates of β , we may have to carry out thousands of trajectories. β is simply the fraction of successful trajectories.



UHBD

• UHBD is a free, well-documented program developed by J. Andrew McCammon's group (originally at UH, now at UCSD) for carrying out Brownian dynamics simulations of protein-ligand association events.

- Local development at UH continued by Prof. Jim Briggs.
- Available at http://adrik.bchs.uh.edu/uhbd.html

Summary

• Electrostatic forces are the most important forces in chemistry and biology.

- The electrostatics of a macromolecule can be approximated by continuum electrostatics (Poisson-Boltzmann Equation).
- The Poisson-Boltzmann Equation can be solved numerically for arbitrarily shaped molecules.
- Brownian Dynamics simulations mimick protein-ligand association and allow calculation of binding rate constants.

Pros/Cons: Continuum Electrostatics

Pros:

- simple model that describes electrostatics aspects of biomolecules very well
- computationally fast
- suitable for binding energy calculations (see following class)

Cons:

- limited conformational flexibility (requires modification of model)
- model may break down when ions from solvent become localized

Resources and Further Reading

WWW:

http://mccammon.ucsd.edu/~chem215 http://trantor.bioc.columbia.edu/programs.html (GRASP, DelPhi) http://adrik.bchs.uh.edu/uhbd.html

Textbooks: Bourne & Weissig, Chapter 21

Papers:Davis and McCammon, *Chem. Rev.* 1990. 90:509-521.Warshel and Papazyan, *Current Opinion Struct. Biol.* 1998. 8:211-217.Davis et al. *Comput. Phys. Commun.* 1991. 62:187-197.

Acknowledgement: G. Matthias Ullmann, Uni Heidelberg