Computational electrostatics for biomolecular systems

Nathan Baker
Washington University in St. Louis
Dept. of Biochemistry and Molecular Biophysics
Center for Computational Biology
http://agave.wustl.edu/
baker@ccb.wustl.edu

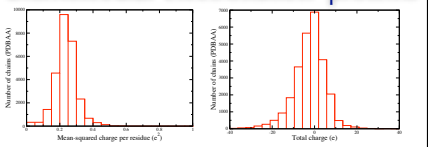
Overview

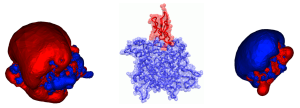
- Sessions:
  - Lecture: "Basic electrostatics and solvation"
  - Lab: "Using APBS and PDB2PQR" and laptop setup (if desired)
  - Lecture: "Advanced electrostatics and solvation"
  - Lab: "Advanced solvation topics"
  - Throughout the day: discussion of your own projects
- Workshop materials available at http://tinyurl.com/ccpb-apbs-workshop

Basic electrostatics and solvation

Electrostatics and solvation in biomolecular systems

Biomolecular electrostatics: proteins

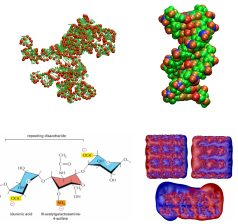




5

Biomolecular electrostatics: other molecules

- dsDNA
  - Approx. linear form
  - Close phosphate spacing
  - 2 e⁻ per 3.4 Å
- RNA
  - Structural diversity
  - Dense phosphate spacing
- Sugars
- Lipids

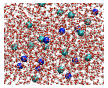


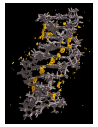
Dermatan sulfate picture from Alberts et al

6

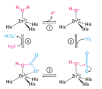
How solvent interacts with biomolecules

- Water properties
  - Dipolar solvent (1.8 D)
  - Hydrogen bond donor and acceptor
  - Polarizable
- Functional behavior:
  - Bulk polarization
  - Site binding or specific solvation
  - Preferential hydration
  - Acid/base chemistry
  - ...





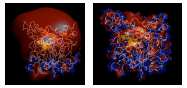
Spine of hydration in DNA minor groove (Kollman, et al.)

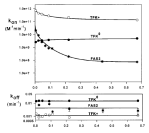


Carbonic anhydrase reaction mechanism (Stryer, et al.)

How ions interact with biomolecules

- Non-specific screening effects
  - Depends only on ionic strength (not species)
  - Results of damped electrostatic potential
  - Described by Debye-Hückel or Poisson-Boltzmann theories for low ionic strengths
- Functional behavior:
  - Described throughout lectures
  - Binding constants
  - Rates



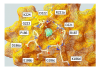


Electrostatic potential of AChE at 0 mM and 150 mM NaCl. Rate and binding affinity decrease with [NaCl] has been attributed to screening effects... although species-dependent influences have been observed. Radic Z, et al. 1997 J Biol Chem 272 (37): 23265-77.

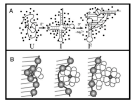
8

How ions interact with biomolecules

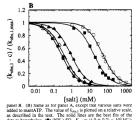
- Site-specific binding
  - Ion-specific
  - Site geometry, electrostatics, coordination, etc. enables favorable binding
  - Functional behavior: co-factors, allosteric activation, folding, etc.



Site of sodium-specific binding in thrombin. Sodium binding converts thrombin to a procoagulant form by allosterically enhancing the rate and changing substrate specificity. Pineda AO, et al. 2004. *J Biol Chem.* **279** (30): 31842-53.



Draper DE, et al. 2005. *Annu. Rev. Biophys. Biomol. Struct.* **34**: 221-43.



Rep + ATP kinetics influenced by specific interactions of divalent anions with ATP binding site. Moore KJM, Lohman TM. 1994. *Biochemistry* **33** (48): 14565-78.

9

How ions interact with biomolecules

- Hofmeister effects (preferential hydration)
  - How much salt is required to precipitate a protein? *It depends on the salt...*
  - Partitioning of ions between water and nonspecific sites on biomolecule
  - Dependent on ion type (solvation energy, etc.)
  - Dominate at high salt concentrations
  - Functional behavior: protein stability, membrane structure and surface potentials, protein-protein interactions

Friedrich Hofmeister

most stabilizing — strongly solvated anions

most destabilizing — weakly solvated anions

$citrate^{2-} > SO_4^{2-} > PO_4^{2-} > F^- > Cl^- > Br^- > I^- > NO_3^- > ClO_4^-$

$N(Me)_4^+ > NH_4^+ > Cs^+ > Rb^+ > K^+ > Na^+ > H^+ > Ca^{2+} > Mg^{2+} > Al^{3+}$

weakly solvated cations — strongly solvated cations

Adapted from http://www.lsbu.ac.uk/water/hofmeist.html

10

Computational methods for biomolecular electrostatics and solvation

11

Modeling biomolecule-solvent interactions

Increasing detail, cost

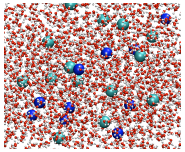
- **Solvent models**
- Quantum
- Explicit
  - Polarizable
  - Fixed charge
- Integral equation
  - RISM
  - 3D methods
  - DFT
- Primitive
  - Poisson equation
- Phenomenological
  - Generalized Born, et al
  - Modified Coulomb's law

- **Ion models**
- Quantum
- Explicit
  - Polarizable
  - Fixed charge
- Integral equation
  - RISM
  - 3D methods
  - DFT
- Field-theoretic
  - Extended models
  - Poisson-Boltzmann equation
- Phenomenological
  - Generalized Born, et al
  - Modified Debye-Hückel

12

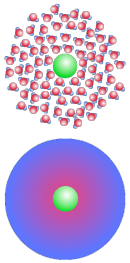
Explicit solvent simulations

- Sample the configuration space of the system: ions, atomically-detailed water, solute
- Sample with respect to a particular ensemble: NpT, NVT, NVE, etc.
- Molecular dynamics or Monte Carlo
- Advantages:
  - High levels of detail
  - Additional degrees of freedom readily included
  - All interactions are explicit
- Disadvantages
  - Slow and uncertain convergence
  - Boundary effects
  - Poor scaling
  - Some effects still not considered in may force fields...



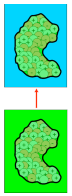
Implicit solvent models

- Solute typically only accounts for 5-10% of atoms in explicit solvent simulation...
- ...so treat solvent effects implicitly:
  - Solvent as polarization density
  - Ions as "mobile" charge density
- Linear and local solvent response
- "Mean field" ion behavior
- Uncertain treatment of "apolar" effects



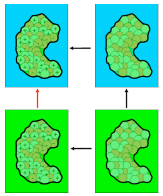
Solvation free energies (and mean forces)

- "Potentials of mean force" (PMF) and solvation free energies
  - Function of conformation
  - Integration over explicit degrees of freedom yields free energy
  - *Global information*
- Mean forces
  - Derivatives of PMFs for atom positions
  - Integration yields PMFs
  - *Local information*



Polar solvation (implicit)

- Charging free energies
  - Solvent: dielectric effects through Poisson equation
  - Ions: mean-field screening effects through Poisson-Boltzmann equation



Electrostatics in a homogeneous dielectric

- An *isotropic* dielectric continuum exhibits the same response in all directions
- The dielectric tensor can be reduced to a scalar
- For a homogeneous isotropic dielectric, electrostatic energies are still governed by Coulomb's law (with a dielectric coefficient)

$$U = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon r}$$

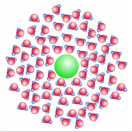
$$\mathbf{F} = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon r^2}\frac{\mathbf{r}}{r}$$

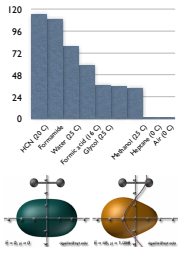
Dielectric constant

17

Dielectric constants

- Several contributions to polarizability
  - Electronic polarizability
  - Intramolecular rearrangement
  - Reorientation of permanent dipole moment
  - Hydrogen bonding networks

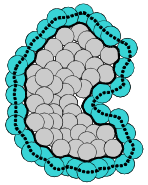




18

Molecular dielectric coefficients

- A heterogeneous molecule like a biomolecule shouldn't really be represented by a continuum dielectric…
- …however, that doesn't keep people from trying
- Multiple dielectric values:
  - 1 = vacuum
  - 2-4 = atomic polarizability (solid)
  - 4-10 = some libration, minor sidechain rearrangement
  - 10-20 = significant internal rearrangement
- Multiple surface definitions:
  - van der Waals
  - Splines
  - Molecular surface



19

Gauss' law, Gauss' theorem, and Poisson equation

- Gauss' law: the integral of the displacement over a surface equals the enclosed charge (general conservation relation)

$$\int_{\partial\Omega} \epsilon(\mathbf{s})\mathbf{E}(\mathbf{s}) \cdot d\mathbf{s} = \int_{\Omega} \frac{\rho(\mathbf{x})}{\epsilon_0} d\mathbf{x}$$

- Gauss' theorem: the integral of a flux over a closed surface equals the enclosed divergence

$$\int_{\partial\Omega} \mathbf{v}(\mathbf{s}) \cdot d\mathbf{s} = \int_{\Omega} \nabla \cdot \mathbf{v}(\mathbf{x}) d\mathbf{x}$$

$$\int_{\partial\Omega} \epsilon(\mathbf{s})\mathbf{E}(\mathbf{s}) \cdot d\mathbf{s} = \int_{\Omega} \nabla \cdot (\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x})) \, d\mathbf{x}$$

- Poisson's equation: divergence of the displacement equals the charge density

$$\int_{\Omega} \nabla \cdot (\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x})) \, d\mathbf{x} = \int_{\Omega} \frac{\rho(\mathbf{x})}{\epsilon_0} d\mathbf{x}$$

$$\nabla \cdot (\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x})) = \frac{\rho(\mathbf{x})}{\epsilon_0}$$

20

Poisson equation: structural elements

- Charge distribution & boundary conditions: solute atom positions and charges
- Dielectric function: solute atom radii, positions; solvent radius; polarizabilities
- Assumptions: linear and local response; no mobile ions

$$
\begin{aligned}
-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) &= \rho(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \\
\phi(\mathbf{x}) &= \phi_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial\Omega
\end{aligned}
$$

Poisson equation: energies

- Total energies obtained from:
  - Integral of polarization energy
  - Sum of charge-potential interactions

$$
\begin{aligned}
G[\phi] &= \frac{1}{4\pi} \int \left\{ \rho(\mathbf{x})\phi(\mathbf{x}) - \frac{\epsilon(\mathbf{x})}{2} \left[\nabla\phi(\mathbf{x})\right]^2 \right\} d\mathbf{x} \\
&= -\frac{1}{8\pi} \int \epsilon(\mathbf{x}) \left[\nabla\phi(\mathbf{x})\right]^2 d\mathbf{x} \\
&= -\frac{1}{8\pi} \int \rho(\mathbf{x})\phi(\mathbf{x}) d\mathbf{x} = -\frac{1}{8\pi} \sum_i q_i \phi(\mathbf{x}_i)
\end{aligned}
$$

Electrostatic energy example: Born ion

The Born ion

- What is the energy of transferring a non-polarizable ion from between two dielectrics?
  - Free energy for charging a sphere in solvent and vacuum
  - No *polar* energy for transferring the uncharged sphere to solvent

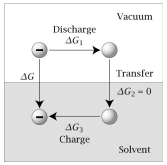


Image from Dill textbook.

Born ion: solvation energies

- Integrate polarization for dielectric media
- Assume ion is non-polarizable
- Subtract energies between media

$$\begin{aligned} G_i &= \frac{\epsilon_0}{2}\int_{\text{solvent}} \epsilon_i \left[\nabla\phi_i(\mathbf{x})\right]^2 d\mathbf{x} \\ &= \frac{\epsilon_0}{2}\int_a^\infty \epsilon_i \left(-\frac{q}{4\pi\epsilon_0\epsilon_i r^2}\right)^2 4\pi r^2 dr \\ &= \frac{q^2}{8\pi\epsilon_0\epsilon_i a} \\ \Delta G &= G_2 - G_1 \\ &= \frac{q^2}{8\pi\epsilon_0 a}\left(\frac{1}{\epsilon_2} - \frac{1}{\epsilon_1}\right) \end{aligned}$$

Poisson-Boltzmann theory

- Simplifies to Debye-Hückel theory
- Continuum dielectric (Poisson equation)
- Non-correlated implicit ions (mean field theory)
- Limitations:
  - Low ion concentration
  - Low ion valency
  - No specific interactions: ion-solute, ion-ion, ion-solvent, solute-solvent, ...
- Going to provide a very simple derivation (***other approaches are more entertaining!***)

Poisson-Boltzmann derivation: Step 1

- Start with Poisson equation to describe solvation and electrostatics
- Supplement biomolecular charge distribution with mobile ion term

$$-\nabla\cdot\epsilon(\mathbf{x})\nabla\phi(\mathbf{x}) = \rho_f(\mathbf{x}) + \rho_m(\mathbf{x})$$

Poisson-Boltzmann equation: Step 2

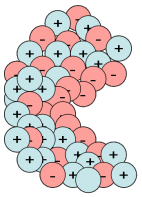
- Choose mobile ion distribution form
  - Boltzmann distribution implies no ion-ion correlation
  - Apparent lack of normalization implies grand canonical ensemble
  - No detailed structure for ion desolvation
- Result: nonlinear partial differential equation
- Don't forget boundary conditions!

$$-\nabla\cdot\epsilon(\mathbf{x})\nabla\phi(\mathbf{x}) = \rho_f(\mathbf{x}) + \sum_m q_m c_m e^{-\beta[q_m\phi(\mathbf{x})+V_m(\mathbf{x})]}$$

Equation coefficients: "fixed" charge distribution

$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) = \rho_f(\mathbf{x}) + \sum_m q_m c_m e^{-\beta[q_m \phi(\mathbf{x}) + V_m(\mathbf{x})]}$$

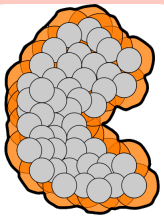
- Charges are *modeled as* delta functions: hard to represent
- Often discretized as splines to "smooth" the problem
- Higher-order charge distributions also possible



Equation coefficients: mobile ion distribution

$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) = \rho_f(\mathbf{x}) + \sum_m q_m c_m e^{-\beta[q_m \phi(\mathbf{x}) + V_m(\mathbf{x})]}$$

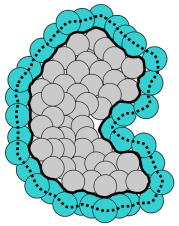
- Usually assume a single exclusion function for all ions
- Generally based on inflated van der Waals radii



Equation coefficients: dielectric function

$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) = \rho_f(\mathbf{x}) + \sum_m q_m c_m e^{-\beta[q_m \phi(\mathbf{x}) + V_m(\mathbf{x})]}$$

- Describes change in local polarizability
  - Low dielectric interior (2-20)
  - High dielectric exterior (80)
- Many definitions
  - Molecular
  - Solvent-accessible
  - van der Waals
  - Smoothed (Gaussian, spline)
- Results can be *very sensitive* to surface definition!



PB special cases: symmetric electrolyte

- Assume similar steric interactions for each species with solute
- Simplify two-term exponential series to hyperbolic sine

$$\begin{aligned}
\rho_m(\mathbf{x}) &= qce^{-\beta[q\phi(\mathbf{x})+V(\mathbf{x})]} - qce^{-\beta[-q\phi(\mathbf{x})+V(\mathbf{x})]} \\
&= qce^{-\beta V(\mathbf{x})}\left[e^{-\beta q\phi(\mathbf{x})} - e^{\beta q\phi(\mathbf{x})}\right] \\
&= -2qce^{-\beta V(\mathbf{x})}\sinh\left[\beta q\phi(\mathbf{x})\right] \\
&= -\bar{\kappa}^2(\mathbf{x})\sinh\left[\beta q\phi(\mathbf{x})\right]
\end{aligned}$$

$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi + \bar{\kappa}^2(\mathbf{x}) \sinh\left[\beta q \phi(\mathbf{x})\right] = \rho_f(\mathbf{x})$$

PB special cases: linearization

- Assume similar steric interactions for each species with solute
- Assume very small local electrostatic energies
- Taylor series expansion of exponential
- Bulk solution electroneutrality

$$\begin{aligned}\rho_m(\mathbf{x}) &= \sum_m q_m c_m e^{-\beta[q_m\phi(\mathbf{x})+V_m(\mathbf{x})]} \\ &\approx e^{-\beta V(\mathbf{x})}\sum_m q_m c_m\left[1-\beta q_m\phi(\mathbf{x})\right] \\ &= -\left[\beta e^{-\beta V(\mathbf{x})}\sum_m q_m^2 c_m\right]\phi(\mathbf{x}) \\ &= -\bar{\kappa}^2(\mathbf{x})\phi(\mathbf{x})\end{aligned}$$

$$-\nabla\cdot\epsilon(\mathbf{x})\nabla\phi+\bar{\kappa}^2(\mathbf{x})\phi(\mathbf{x})=\rho_f(\mathbf{x})$$

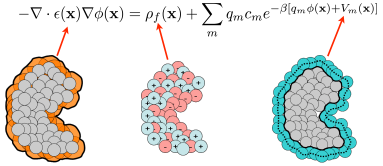
Poisson-Boltzmann energies

- Similar to Poisson equation
- Functional: integral of solution over domain
- Solution extremizes energy
- Basis for calculating forces: charge-field, dielectric boundary, osmotic pressure

$$\begin{aligned}G[\phi] &= \frac{1}{4\pi}\int_\Omega\left\{\rho_f(\mathbf{x})\phi(\mathbf{x})-\frac{\epsilon(\mathbf{x})}{2}\left[\nabla\phi(\mathbf{x})\right]^2+\sum_m c_m\left[e^{-\beta[q_m\phi(\mathbf{x})+V_m(\mathbf{x})]}-1\right]\right\}d\mathbf{x} \\ &\approx \frac{1}{4\pi}\int_\Omega\left\{\rho_f(\mathbf{x})\phi(\mathbf{x})-\frac{\epsilon(\mathbf{x})}{2}\left[\nabla\phi(\mathbf{x})\right]^2+\frac{\bar{\kappa}^2(\mathbf{x})}{2}\left[\phi(\mathbf{x})\right]^2\right\}d\mathbf{x} \\ &= -\frac{1}{8\pi}\int_\Omega\rho_f(\mathbf{x})\phi(\mathbf{x})d\mathbf{x}\end{aligned}$$

Poisson-Boltzmann equation

$$-\nabla\cdot\epsilon(\mathbf{x})\nabla\phi(\mathbf{x})=\rho_f(\mathbf{x})+\sum_m q_m c_m e^{-\beta[q_m\phi(\mathbf{x})+V_m(\mathbf{x})]}$$



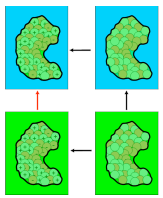
$$G[\phi] = \frac{1}{4\pi}\int_\Omega\left\{\rho_f(\mathbf{x})\phi(\mathbf{x})-\frac{\epsilon(\mathbf{x})}{2}\left[\nabla\phi(\mathbf{x})\right]^2+\sum_m c_m e^{-\beta V_m(\mathbf{x})}\left[e^{-\beta q_m\phi(\mathbf{x})}-1\right]\right\}d\mathbf{x}$$

$$\mathbf{F}_i[\phi]=-\frac{\partial G[\phi]}{\partial\mathbf{r}_i} = -\frac{1}{4\pi}\int_\Omega\left\{\frac{\partial\rho_f(\mathbf{x})}{\partial\mathbf{r}_i}\phi(\mathbf{x})+\frac{1}{2}\frac{\partial\epsilon(\mathbf{x})}{\partial\mathbf{r}_i}\left[\nabla\phi(\mathbf{x})\right]^2+\sum_m c_m\frac{\partial e^{-\beta V_m(\mathbf{x})}}{\partial\mathbf{r}_i}\left[e^{-\beta q_m\phi(\mathbf{x})}-1\right]\right\}d\mathbf{x}$$

Reaction field | Dielectric boundary | "Osmotic"

Reminder: polar solvation

- Charging free energies
  - Solvent: dielectric effects through Poisson equation
  - Ions: mean-field screening effects through Poisson-Boltzmann equation
- What about the uncharged steps?

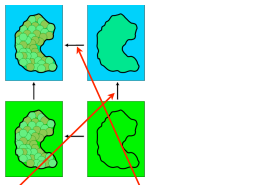


Nonpolar solvation (implicit)

- It's not just surface area!
- WCA formalism:
  - Cavity creation
    - Small length scales: proportional to volume (pressure) and area (surface tension)
    - Large length scales: proportional to area (surface tension)
  - Dispersive interactions
    - Modeled by WCA formalism
    - Integral of potential over solvent-accessible volume

Adapted from: Levy RM, Zhang LY, Gallicchio E, Felts AK. 2003. *J Am Chem Soc* **125** (31): 9523-9530.

Nonpolar solvation: implementation

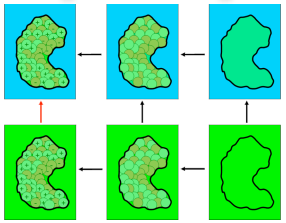


$$W^{(\mathrm{np})}(\mathbf{x}) = \gamma A(\mathbf{x};\sigma) + pV(\mathbf{x};\sigma) + \bar{\rho}\int_\Omega g_0(\mathbf{x},\mathbf{y};\sigma)U_{\mathrm{att}}^{(\mathrm{np})}(\mathbf{x},\mathbf{y};\sigma)d\mathbf{y}$$

$$\mathbf{F}_i^{(\mathrm{np})}(\mathbf{x}) = -\gamma\frac{\partial A(\mathbf{x};\sigma)}{\partial \mathbf{x}_i} - p\frac{\partial V(\mathbf{x};\sigma)}{\partial \mathbf{x}_i} + \bar{\rho}\int_\Omega g_0(\mathbf{x},\mathbf{y};\sigma)\frac{\partial U_{\mathrm{att}}^{(\mathrm{np})}(\mathbf{x},\mathbf{y};\sigma)}{\partial \mathbf{x}_i}d\mathbf{y}$$

Wagoner JA, Baker NA. *Proc Natl Acad Sci USA*, **103**, 8331-6, 2006.

Putting it all back together

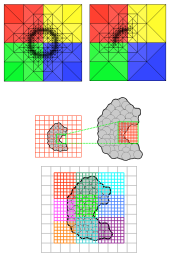


Adapted from: Levy RM, Zhang LY, Gallicchio E, Felts AK. 2003. *J Am Chem Soc* **125** (31): 9523-9530.

Software for continuum electrostatics and solvation

Solving the PB equation

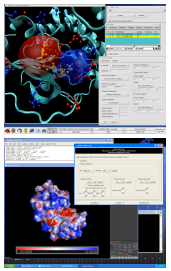
- Parallel adaptive finite element methods
  - Bank and Holst, *SIAM Review*, 2003
  - *A posteriori* residual-based error estimators
  - PB-specific customization
  - FEtk-based solution (http://www.fetk.org/)
- Parallel focusing methods
  - Baker et al, *Proc Natl Acad Sci*, 2001
  - Loosely related to Bank-Holst method
  - PMG-based solution (http://www.fetk.org/)



41

Implicit solvent tools

- APBS (http://apbs.sf.net/)
  - PB electrostatics calculations
  - Freely available
  - Fast finite element (FEtk) and multigrid (PMG) solvers from Holst group (http://fetk.org)
  - Works with most popular visualization software (VMD, PMV, PyMOL)
  - Links with CHARMM, AMBER, TINKER*
- PDB2PQR (http://pdb2pqr.sf.net/)

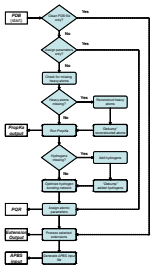


Baker NA, et al *Proc Natl Acad Sci USA*, **98**, 10037, 2001;
*Schnieders MJ, et al. *J Chem Phys*, **126**, 124114, 2007.

42

PDB2PQR

- PDB2PQR (http://pdb2pqr.sf.net/)
  - Collaborative project: Jens Nielsen, Jan Jensen, and Gerhard Klebe groups
  - Prepares PDB files for other calculations
  - Assigns titration states (PROPKA) and optimizes hydrogen positions
  - "Repairs" missing heavy atoms
  - Assigns parameters
  - Web-based and command-line
  - Freely available (GPL or BSD) and extensible



Dolinsky TJ, et al. *Nucleic Acids Res*, **35**, W522-5, 2007;
Dolinsky TJ, et al. *Nucleic Acids Res*, **32**, W665-7, 2007.

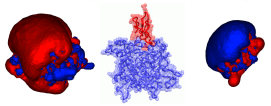
43

Advanced electrostatics and solvation

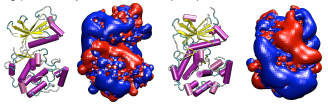
Applications of continuum electrostatics

Visualization and analysis of electrostatic potentials

Electrostatic potential comparisons



The interaction of AChE (structure: center, blue; electrostatic potential: left) with its inhibitor FAS2 (structure: center, red; electrostatic potential: right) is electrostatically driven. Blue surfaces denote positive potential isocontours; red surfaces denote negative.

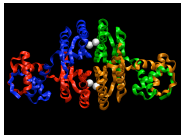


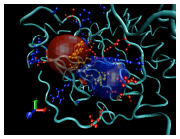
The electrostatic potential and structure of two cAMP-dependent kinases: 1FOT:A (left) and 1JBP:E (right). The two proteins share only 51% sequence identity and adopt different conformations but still exhibit the same electrostatic potential motifs, share the same fold, and perform the same basic biological function.

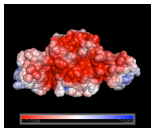
47

Inspection of ligand binding sites

- Balanol protein kinase A binding (Wong CF, et al. J Med Chem 44, 1530-9 (2001)
- NikR Ni(II) and DNA binding



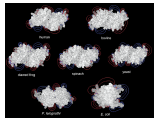




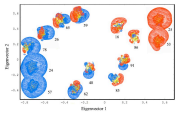
48

Quantitative comparison of electrostatic potentials

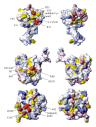
- Do electrostatic potentials tell us anything about biomolecular function?
  - Ligand binding
  - Active sites or shifted pK_as?
  - Structural (de)stabilization?



SOD comparison. Livesay DR, et al. *Biochemistry* 42, 3464-73 (2003)

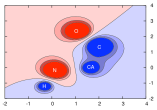


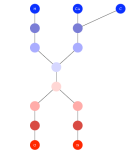
PH domain comparison: similar fold, similar electrostatics, different sequence. Blomberg N, et al. *Proteins* 37, 379-877 (1999).

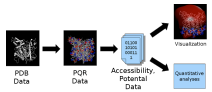


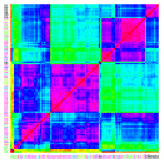
Putative Mg^{++} and RNA binding domains in an unclassified protein. Elcock AH, *J Mol Biol* 312, 885-96 (2001).

Multiresolution contour trees





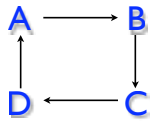




Thermodynamics

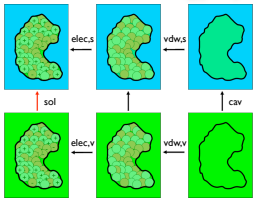
Free energy cycles

- At the heart of most calculations...
- ...because we can't usually directly calculate the quantity of interest
- Most important principle:
  - Energy is a state function
  - Integral of energy changes over a closed cycle is zero



$$\Delta G_{A\to B} + \Delta G_{B\to C} + \Delta G_{C\to D} + \Delta G_{D\to A} = 0$$

Solvation free energy cycle



$$\Delta G_{\mathrm{sol}} - \Delta G_{\mathrm{elec,s}} - \Delta G_{\mathrm{vdw,s}} - \Delta G_{\mathrm{cav}} + \Delta G_{\mathrm{vdw,v}} + \Delta G_{\mathrm{elec,v}} = 0$$

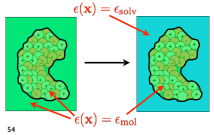
$$\Delta G_{\mathrm{sol}} = \Delta G_{\mathrm{elec,s}} - \Delta G_{\mathrm{elec,v}} + \Delta G_{\mathrm{vdw,s}} - \Delta G_{\mathrm{vdw,v}} + \Delta G_{\mathrm{cav}}$$

Solvation energies

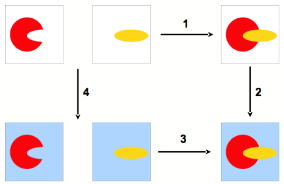
- Absolute energies are generally not useful: inaccurate
- Solvation: an excellent way to remove "self energies"
- Building block for most electrostatics calculations
- "Total energies" can be recovered by adding in vacuum polar and nonpolar contributions

$$\begin{aligned} -\epsilon_{\mathrm{h}} \nabla^2 \phi_{\mathrm{h}}(\mathbf{x}) &= \rho(\mathbf{x}) \\ -\nabla \cdot \epsilon_{\mathrm{ih}} \nabla \phi_{\mathrm{ih}}(\mathbf{x}) &= \rho(\mathbf{x}) \end{aligned}$$

$$\begin{aligned} \Delta_{\mathrm{solv}} G &= G\left[\phi_{\mathrm{ih}}\right] - G\left[\phi_{\mathrm{h}}\right] \\ &= -\frac{1}{8\pi} \sum_i \left[\phi_{\mathrm{ih}}(\mathbf{x}_i) - \phi_{\mathrm{h}}(\mathbf{x}_i)\right] \end{aligned}$$



Absolute binding free energy cycle

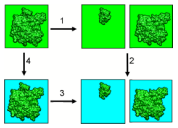


$$\Delta G_1 + \Delta G_2 - \Delta G_3 - \Delta G_4 = 0$$

$$\Delta G_3 = \Delta G_1 + \Delta G_2 - \Delta G_4$$

Binding energies

- Separate calculation into two steps:
  - Calculate electrostatic interaction for homogeneous dielectric (Coulomb's law)
  - Calculate solvation energy change upon binding
- Self-interactions are removed in solvation energy calculation
- Absolute binding energies are tricky...



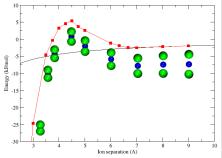
$$\begin{aligned} \Delta G_3 &= \Delta G_1 + (\Delta G_2 - \Delta G_4) \\ &= \Delta_{\mathrm{solv}} G + \Delta_{\mathrm{coul}} G \end{aligned}$$

Ion desolvation PMF

- Two nonpolarizable ions
  - Solve for polar energy as a function of separation
  - Poison equation
- Increase in energy as water is “squeezed” out
  - Desolvation effect
  - Smaller volume of polarized water
- Important points
  - Non-superposition of ion potentials
  - Reaction field causes repulsion at short distances
  - Dielectric medium “focuses” field

Polar binding energy (PMF): two ions

- Water dielectric
- Two ions: 3 Å radii, non-polarizable, opposite charges
- Basic calculation:
  - Calculate solvation energies of isolated ions
  - Calculate solvation energy of “complex”
  - Subtract solvation energies
  - Add vacuum Coulomb’s law

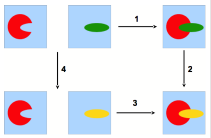


Polar binding energy: how-to

- Method #1 (allows for conformational change)
  - Calculate solvation energies for complex and isolated components. Use focusing as needed.
  - Subtract to calculate solvation energy change upon binding.
  - Calculate Coulombic energies for complex and isolated components ***using same internal dielectric constant!*** Subtract to calculate Coulombic energy change upon binding.
  - Add solvation and Coulombic energy changes.
- Method #2 (fast but dangerous!)
  - Calculate absolute energies for complex and isolated components. Using focusing as needed. ***Use the same grid, dielectric, etc. parameters for all calculations!!!***
  - Subtract.

Relative binding free energy cycle

- Usually better accuracy
- Cancellation of numerical errors
- Cancellation of hard-to-quantify terms
- Useful for predicting mutations, changes in functional groups, etc.

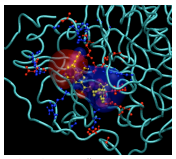


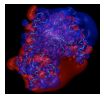
$$\Delta G_1 + \Delta G_2 - \Delta G_3 - \Delta G_4 = 0$$

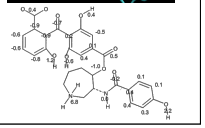
$$\Delta\Delta G = \Delta G_1 - \Delta G_3 = \Delta G_4 - \Delta G_2$$

Binding energy example

- Protein kinase A inhibition by balanol
- Wong CF, et al. *J Med Chem* 44, 1530-9 (2001)
- Continuum electrostatics analysis of protein mutations and functional group changes on binding affinity

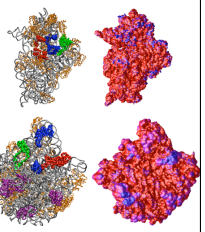






Application to ribosomes

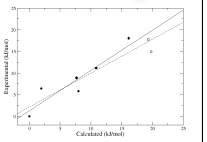
- Ribosome central to protein synthesis machinery
- Target for several pharmaceuticals
- Nucleoprotein composition make it computationally challenging
- Composed of two subunits (large and small):
  - 30S consists of 88,000 atoms and roughly 200 Å cube
  - 50S consists of more than 95,000 atoms and roughly 200 Å cube
- Function involves several interesting features:
  - Protein-nucleic acid association
  - Protein-protein association
  - Conformational changes
  - Salt dependence (type and quantity)
- Solved on 343 processors of Blue Horizon to 0.41 Å (30S) and 0.43 Å (50S) resolution

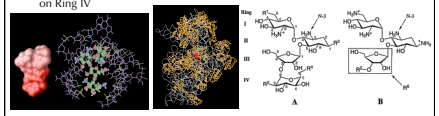


Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Ma C, et al, *J Am Chem Soc*, **124**, 1438-42, 2002.

Ribosome-antibiotic binding

- Determine binding energies between 30S ribosomal subunit and aminoglycoside antibiotics
- Good agreement for experimental and computational relative binding free energies: 0.78 ± 0.13 slope with small molecules, 0.95 ± 0.19 slope without
- Suggests importance of basic groups on Ring IV

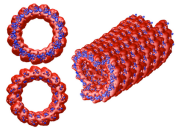




Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Ma C, et al, *J Am Chem Soc*, **124**, 1438-42, 2002.

Application to microtubules

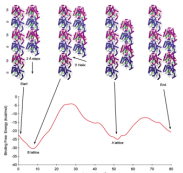
- Important cytoskeletal components: structure, transport, motility, division
- Typically 250-300 Å in diameter and up to millimeters in length
- Computationally difficult due to size (1,500 atoms Å^{-1}) and charge (-4.5 e Å^{-1})
- Solved LPBE at 150 mM ionic strength on 686 processors for 600 Å-long, 1.2-million-atom microtubule
- Resolution to 0.54 Å for largest calculation: quantitative accuracy



Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Sept D, et al, *Protein Sci*, **12**, 2257-61, 2003.

Microtubule stability and assembly

- Collaboration with Andy McCammon (UCSD) and Dave Sept (Wash U BME)
- Performed series of calculations on tubulin dimers and protofilament pairs
- Poisson-Boltzmann electrostatics and SASA apolar energies
- Observed 7 kcal/mol stronger interactions between protofilaments than within
- Determined energetics for helix properties; predict correct minimum for experimentally-observed A (52 Å) and B (8-9 Å) lattices



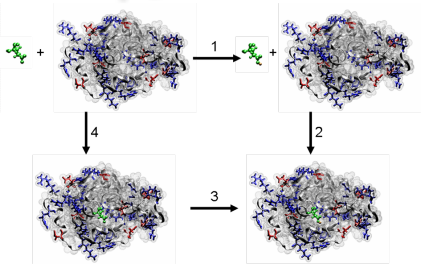
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Dolinsky TJ, et al, *Nucl Acids Res*, **32**, W665-7, 2004.

pK_a calculations

- Want acid dissociation constant for residues in a particular structural context
- Use "model" pK_as for amino acids
- Calculate "intrinsic" pK_a from two calculations:
  - Binding of unprotonated residue
  - Binding of protonated residue
- Calculate titration state and actual from sampling of coupled pK_as
- Conformational distributions can matter

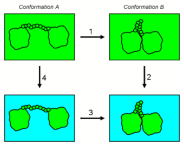
Amino acid	α-carboxylic acid	α-amino	Side chain
Alanine	2.35	9.87	
Arginine	2.01	9.04	12.48
Asparagine	2.02	8.80	
Aspartic acid	2.10	9.82	3.86
Cysteine	2.05	10.25	8.00
Glutamic acid	2.10	9.47	4.07
Glycine	2.35	9.78	
Histidine	1.77	9.18	6.10
Isoleucine	2.23	9.76	
Leucine	2.33	9.74	
Lysine	2.18	8.95	10.53
Methionine	2.28	9.21	
Phenylalanine	2.58	9.24	
Proline	2.00	10.60	
Serine	2.21	9.15	
Threonine	2.09	9.10	
Tryptophan	2.38	9.39	
Tyrosine	2.20	9.11	10.07
Valine	2.29	9.72	

pK_a calculations



Conformational changes: two conformations

- Same idea as binding free energies
  - Calculate polar energy change due to conformational change in homogeneous dielectric (Coulomb's law)
  - Calculate polar solvation energy change due to conformation change in inhomogeneous dielectric
  - Subtract.



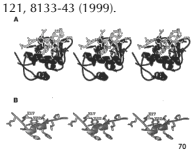
Conformational change: multiple conformations

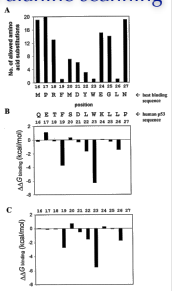
- MM/PBSA: include contribution from multiple conformations to energy
- Typically used for binding energy
- Accounts for conformational distribution effects on
  - Intra- and intermolecular energy (mechanics)
  - Solvation (Poisson-Boltzmann and apolar)
  - Entropy (quasi-harmonic)

$$
\begin{aligned}
Z_i &= \int e^{-\beta[U(\mathbf{x}_i)+W(\mathbf{x}_i)]} d\mathbf{x}_i \approx z_i^{\text{int}} e^{-\beta\langle U(\mathbf{x}_i)+W(\mathbf{x}_i)\rangle} \\
\Delta G &= -\frac{1}{\beta}\log\left(\frac{Z_{AB}}{Z_A Z_B}\right) \\
&\approx -\frac{1}{\beta}\log\left(\frac{z_{AB}^{\text{int}}}{z_A^{\text{int}} z_B^{\text{int}}}\right) + \langle U_{AB}(\mathbf{x}_i) + W_{AB}(\mathbf{x}_i)\rangle \\
&\quad - \langle U_A(\mathbf{x}_i) + W_A(\mathbf{x}_i)\rangle - \langle U_B(\mathbf{x}_i) + W_B(\mathbf{x}_i)\rangle
\end{aligned}
$$

MM-PBSA: computational alanine scanning

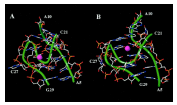
- Examine the interface of oncoprotein MDM2 with N-terminus of tumor suppressor p53
- Apply MM-PBSA methods with normal mode entropies
- Surprisingly good results!
- Massova I, Kollman PA. *J Am Chem Soc* 121, 8133-43 (1999).

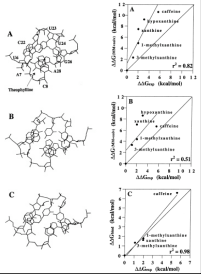




MM-PBSA: RNA-ligand interactions

- Calculate binding free energy of theophylline to RNA 33-mer
- Use normal mode entropy calculation
- Compare with thermodynamic integration
- Reasonable agreement between computational (-7.5 kcal/mol) and experimental (-9.0 kcal/mol) binding energies
- Pretty good relative binding free energies
- Gouda H, Kuntz ID, Case DA, Kollman PA. 2003. Biopolymers 68 (1): 16-34.





Summary

- Continuum electrostatics:
  - Linear and local response
  - Mean field ion behavior
- Numerical methods
- Applications
  - Structural bioinformatics and other analyses
  - Thermodynamics
  - Binding affinities
  - Solvation energies
- Kinetics
  - Forces
  - Rate constants
  - Dynamics