Electrostatics and solvation for biomolecular systems

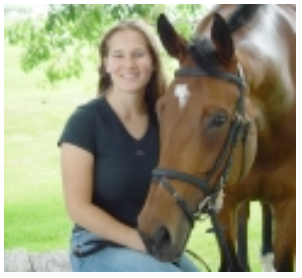
BIO 5325
Sept. 26, 2006

Nathan Baker
http://agave.wustl.edu/
nathan.baker@wustl.edu

Baker group

Mike Bradley
Molecular biophysics grad student

Sun Joo Lee
Molecular biophysics grad student



Rachel Rice
Molecular biophysics grad student

Todd Dolinsky
Programmer

Peter Jones
Programmer

Yuhua Song
Asst. Prof., Univ. Alabama

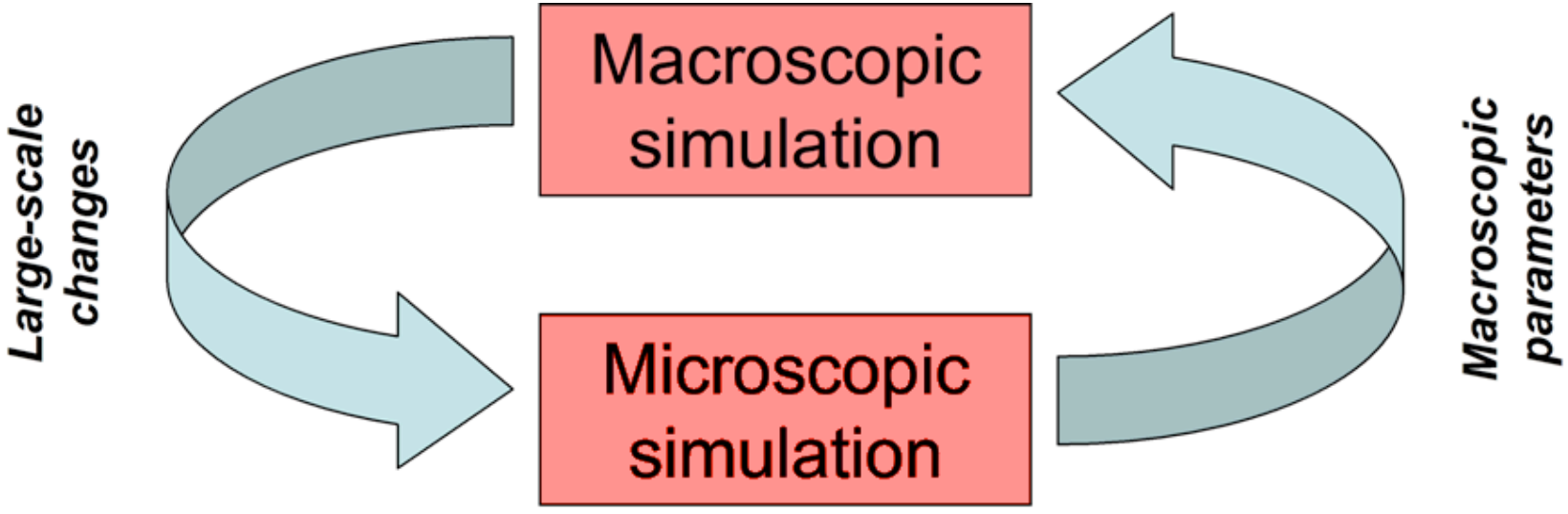
Feng Dong
Postdoc

Brett Olsen
Molecular & cell biology grad student

Jason Wagoner
Grad. student, Stanford

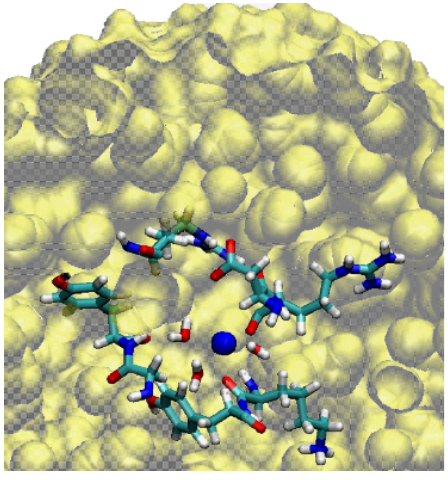
Baker group research

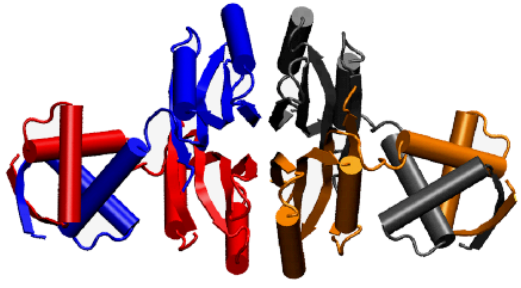
Multiscale modeling

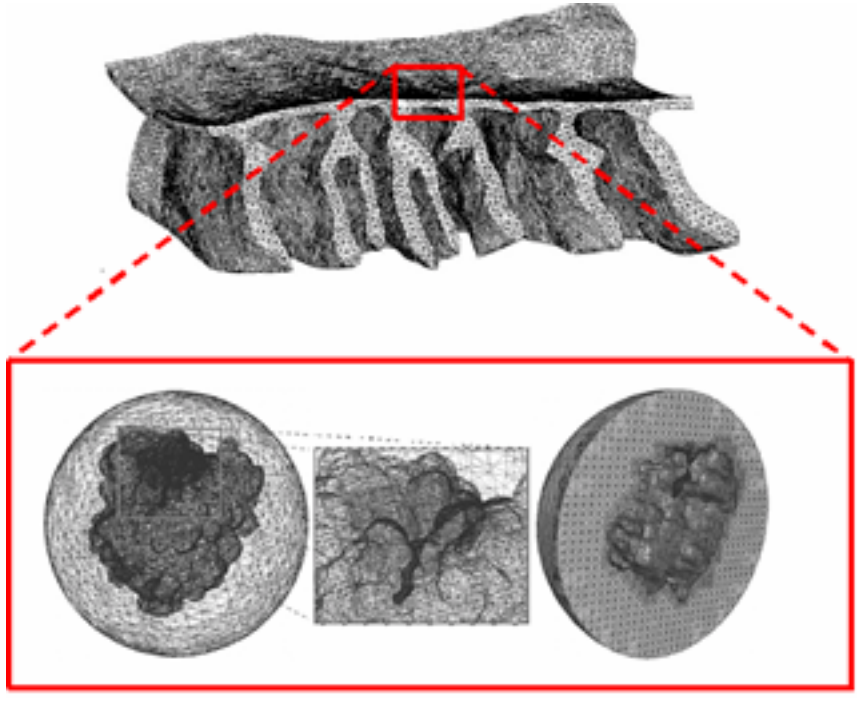


- How can continuum methods be systematically integrated with atomic-detail simulations in regions of interest?
- What are the limits of validity for continuum methods?
- What do we learn about the physics and chemistry of these systems through such integration?
- The solutions depend on the questions!

Biological "signaling"



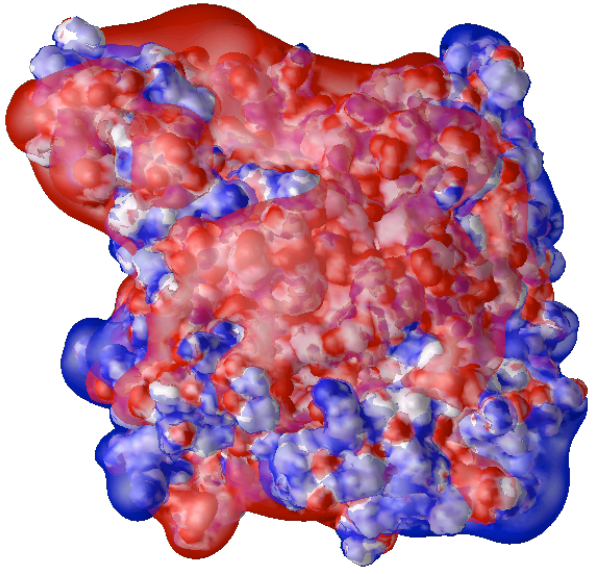




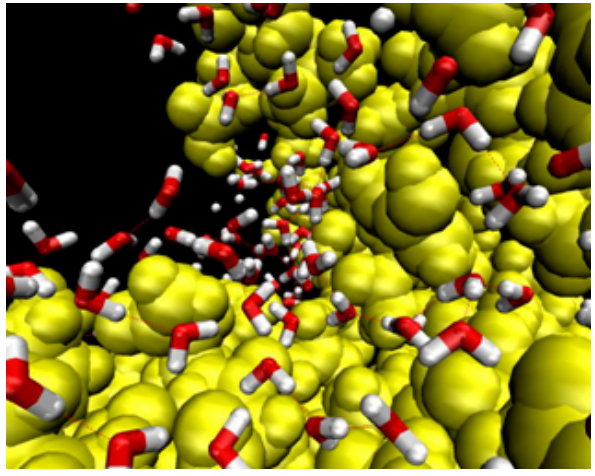
Allosteric proteins

Diffusion

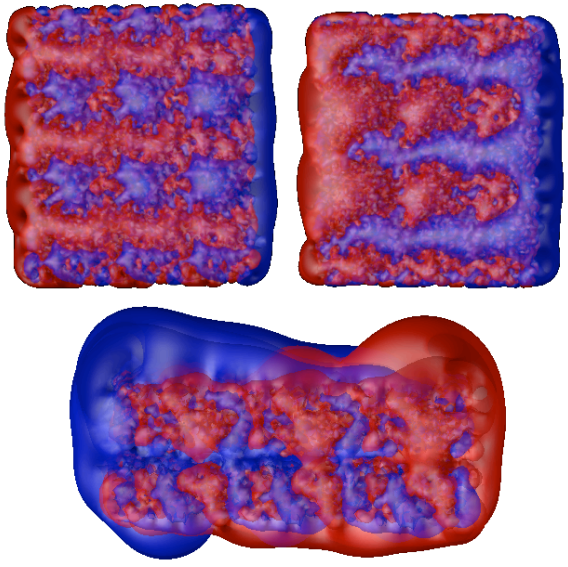
Water and ions



Electrostatics



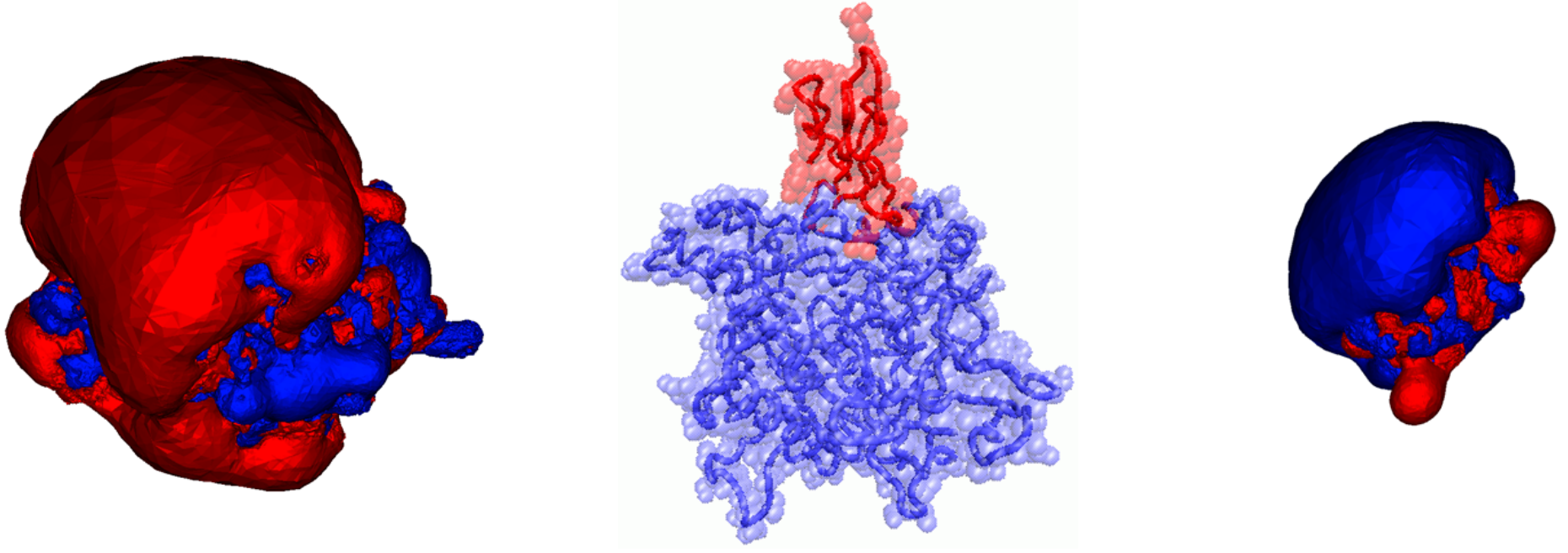
Solvation



Electromechanical coupling

Electrostatics in biomolecular interactions

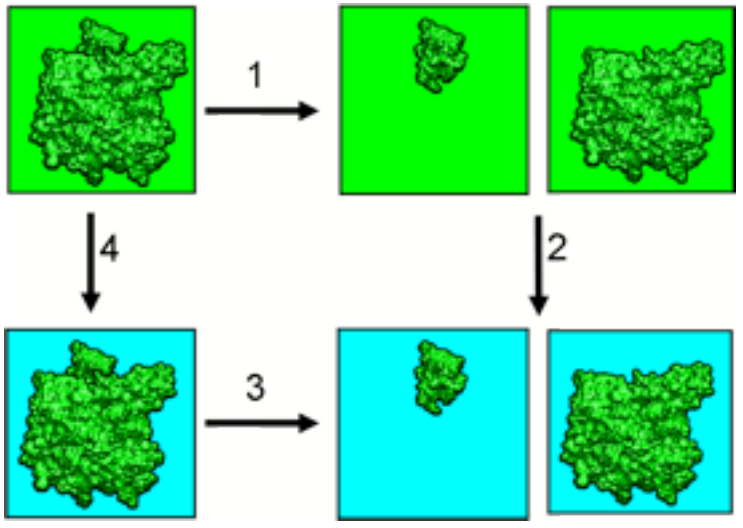
- Highly relevant to biological function
- Important tools in interpretation of structure and function
- Challenging computational problem: long-ranged and divergent

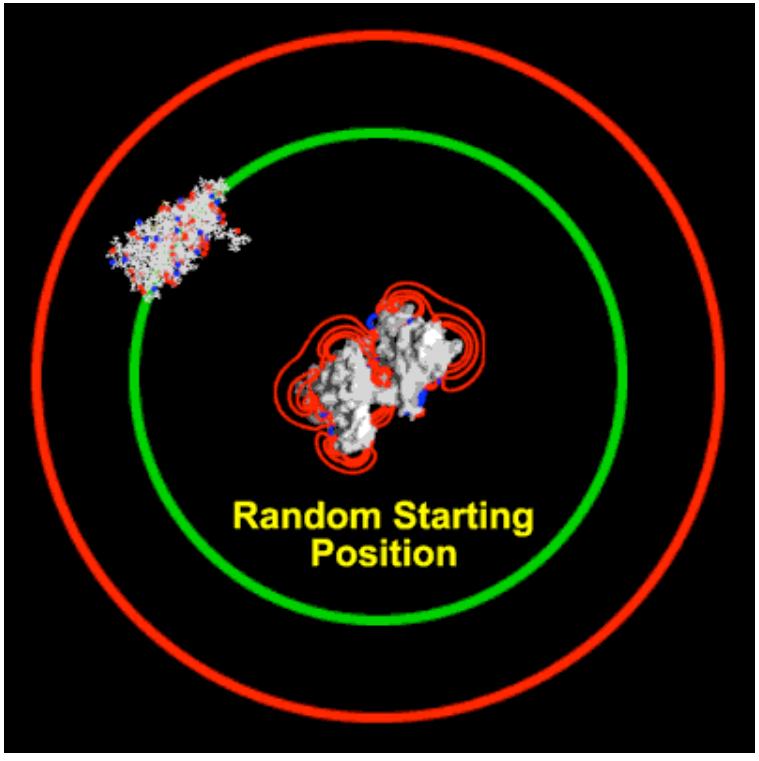


Acetylcholinesterase and fasciculin-2

Where do electrostatics matter?

- Computational applications
- Thermodynamics:
  - Solvation
  - Binding
  - Acid/base equilibria
  - Conformational changes
- Kinetics
  - Binding
  - Conformational changes

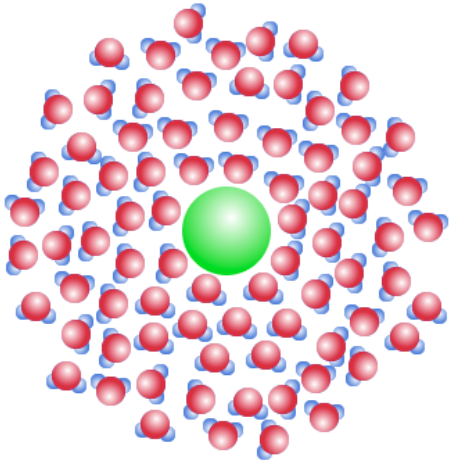


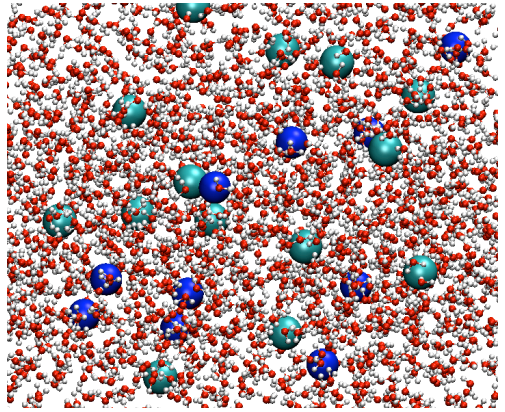


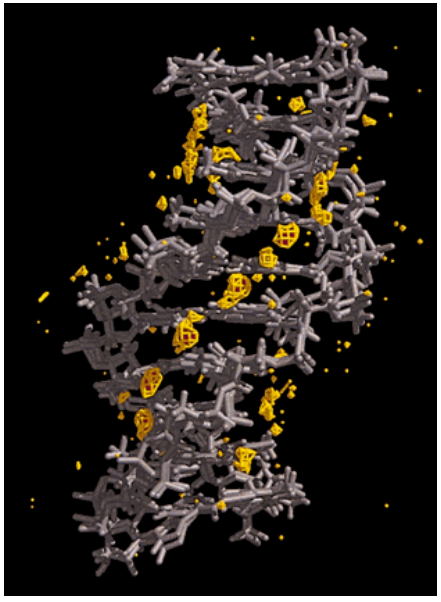
Movie courtesy of Dave Sept

How solvent interacts with biomolecules

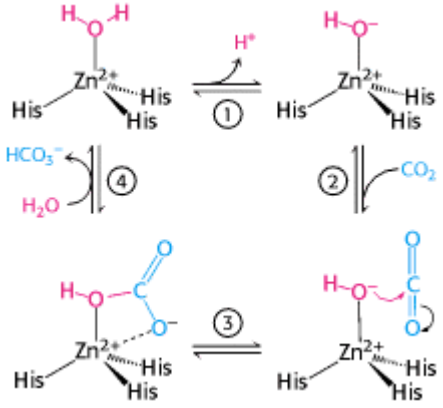
- Water properties
  - Dipolar solvent (1.8 D)
  - Hydrogen bond donor and acceptor
  - Polarizable
- Functional behavior:
  - Bulk polarization
  - Site binding or specific solvation
  - Preferential hydration
  - Acid/base chemistry
  - …







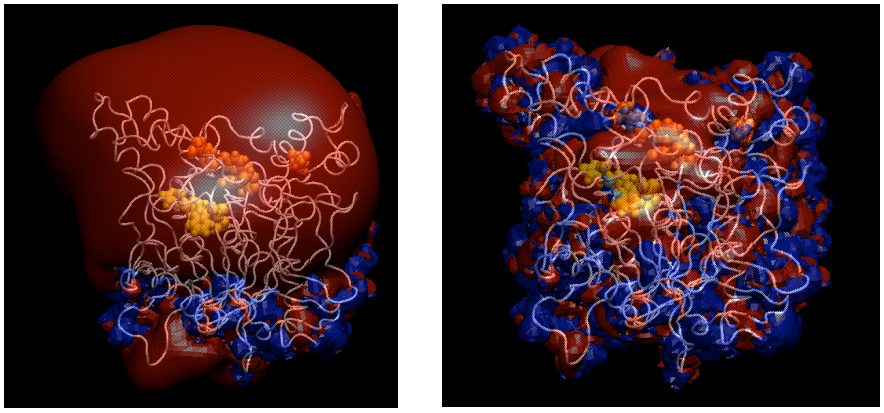
Spine of hydration in DNA minor groove (Kollman, et al.)

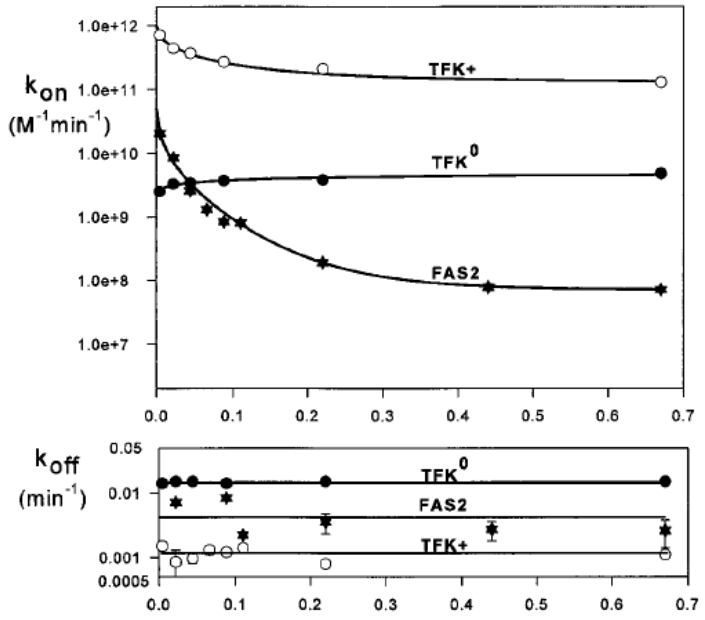


Carbonic anhydrase reaction mechanism (Stryer, et al.)

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

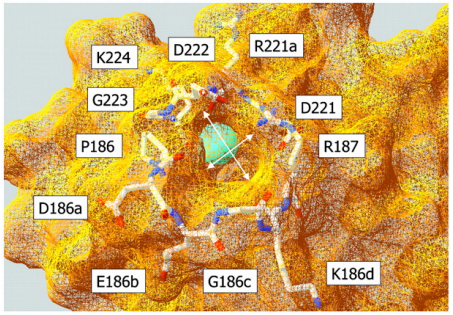




Electrostatic potential of AChE at 0 mM and 150 mM NaCl. Rate and binding affinity decrease with [NaCl] has been attributed to screening effects… although species-dependent influences have been observed. Radic Z, et al. 1997. J Biol Chem 272 (37): 23265-77.

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.

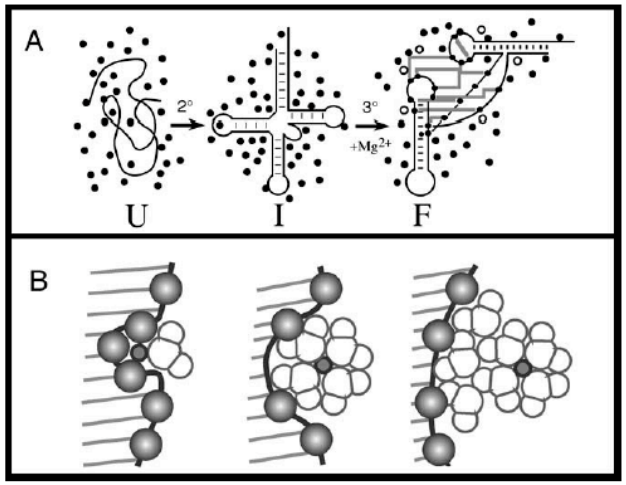
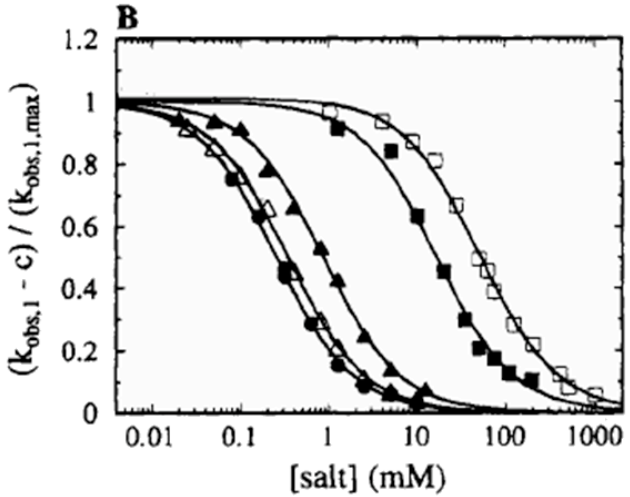


Figure 1 Ions and the RNA folding problem. Top panel: tRNA folding can be visualized as successive formation of secondary structure (intermediate, I) and tertiary structure (folded, F) from unfolded (U) RNA. Counterions associated with the RNAs are illustrated as black dots (monovalent) and open circles (divalent); Mg^{2+} is particularly effective in promoting tertiary structure folding. Bottom panel: Environments that might be seen by an RNA-associated ion differ in the extent of direct ion-RNA contacts, water-mediated ion-RNA contacts, and hydration. The ion is represented by a small circle; water molecules are outlined, and shaded spheres represent RNA backbone phosphates.

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



panel B. (B) Same as for panel A, except that various salts were added to mantATP. The value of $k_{obs,1}$ is plotted on a relative scale, as described in the text. The solid lines are the best fits of the data to hyperbolae: (●) $(NH_4)_2SO_4$, $K_{app} = (4.0 \pm 0.2) \times 10^3\ M^{-1}$; (△) Na_2SO_4, $K_{app} = (2.9 \pm 0.1) \times 10^3\ M^{-1}$; (▲) NaP_i, $K_{app} = (1.2 \pm 0.1) \times 10^3\ M^{-1}$; (■) NH_4Cl, $K_{app} = 59\ (\pm 4)\ M^{-1}$; (□) NaCl, $K_{app} = 20\ (\pm 1)\ M^{-1}$.

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

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 (1782-1864)

most stabilizing ***most destabilizing***

strongly solvated anions weakly solvated anions

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

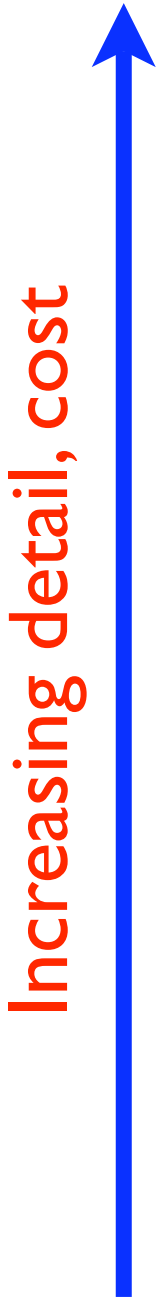
$$N(Me)_4^+ > NH_4^+ > Cs^+ > Rb^+ > K^+ > Na^+ > H^+ > Ca^{2+} > Mg^{2+} > Al^{3+}$$

weakly solvated cations strongly solvated cations

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

Implicit solvent models: basic overview

Modeling biomolecule-solvent interactions

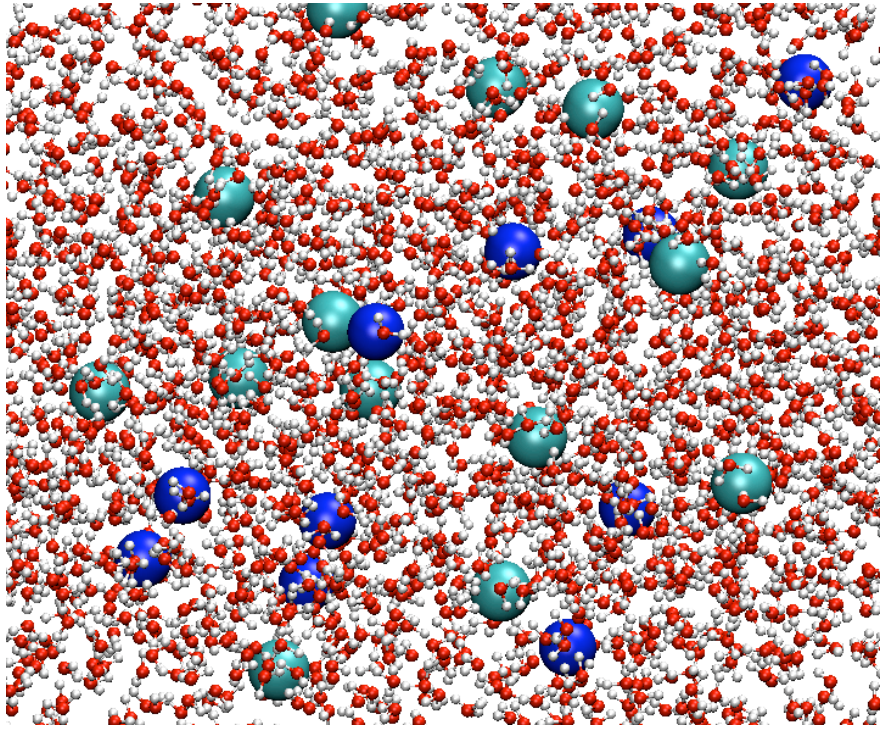


- **Solvent models**
- Quantum
- Explicit
  - Polarizable
  - Fixed charge
- Integral equation
  - RISM
  - 3D methods
  - DFT
- Primitive
  - Poisson equation
- Phenomenological
  - Generalized Born, et al
  - Modified Coulomb's law

- **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

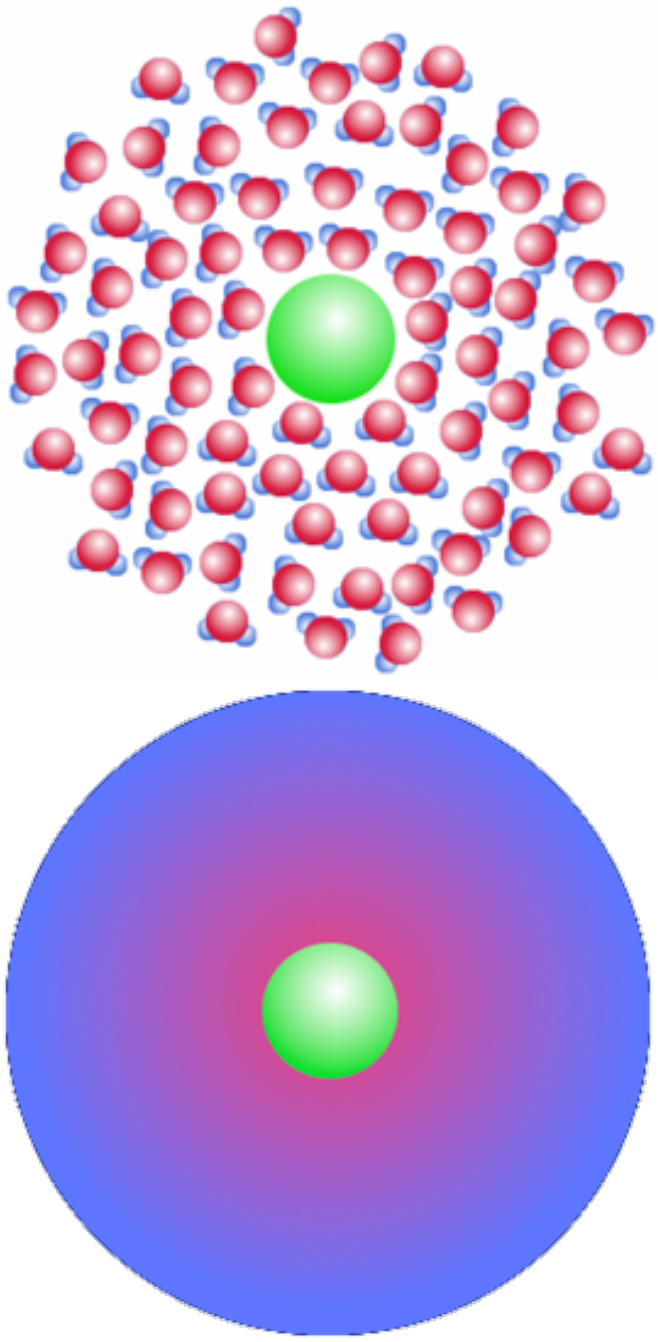
Explicit solvent simulations

- 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...



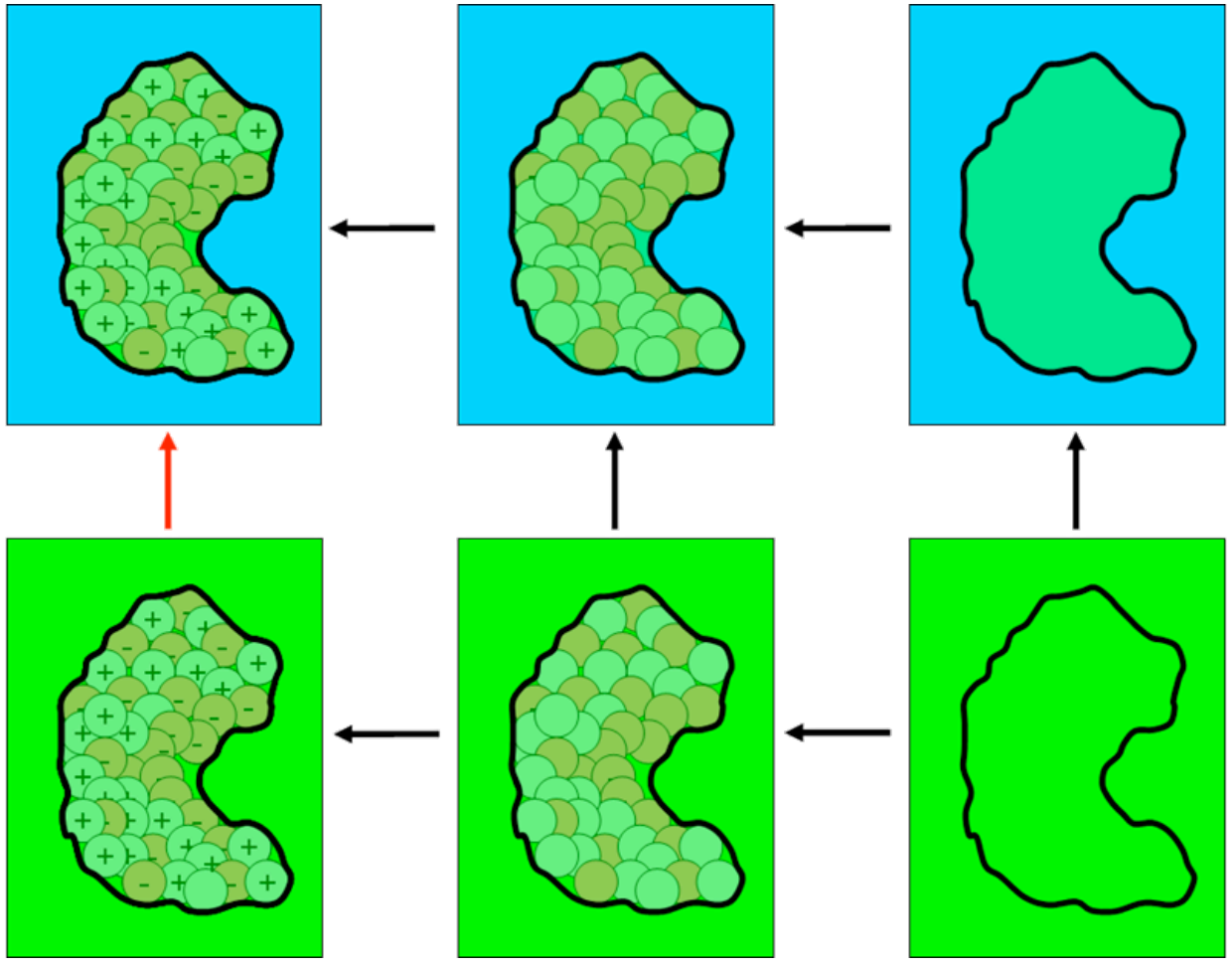
Electrostatics in a dielectric continuum

- Continuum dielectric...
- Has no atomic detail
- Is related to the *polarization* of the medium: redistribution of charges
- Responds *linearly* and *locally* to dampen an applied field
- Is characterized by a dielectric tensor
- Reduces the strength of electrostatic interactions relative to a vacuum



Solvation free energies

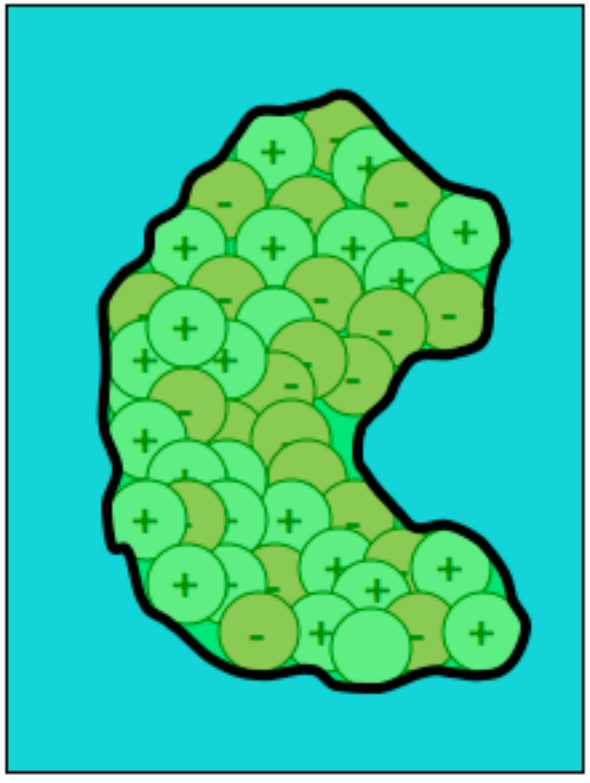
- Implicit solvent models
  - "Potentials of mean force" and solvation free energies
  - Mean forces
- Free energy cycle (as function of conformation)
  - Cavity creation
  - Dispersive interactions
  - Charging process



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

Poisson equation

- Classic equation of continuum electrostatics
- Can be derived from more microscopic models
- Assumptions:
  - No dielectric saturation (linear response)
  - No solvent-solvent correlation (local response)
  - No atomic detail, hydrogen bonding, chemistry, etc.

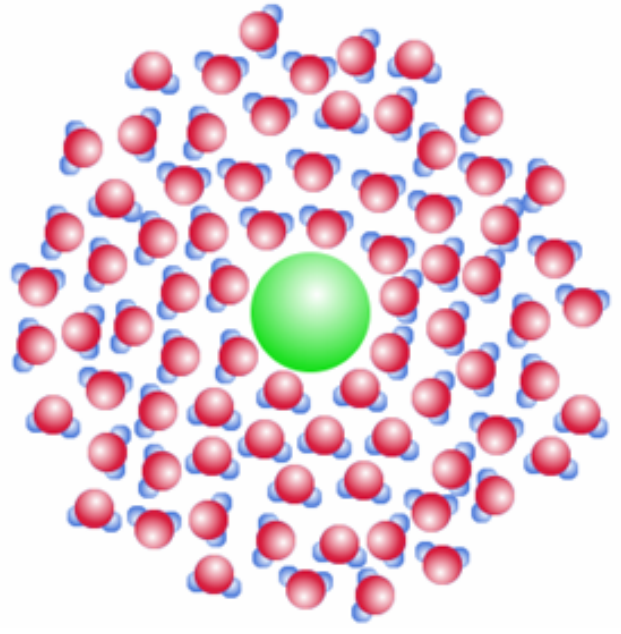


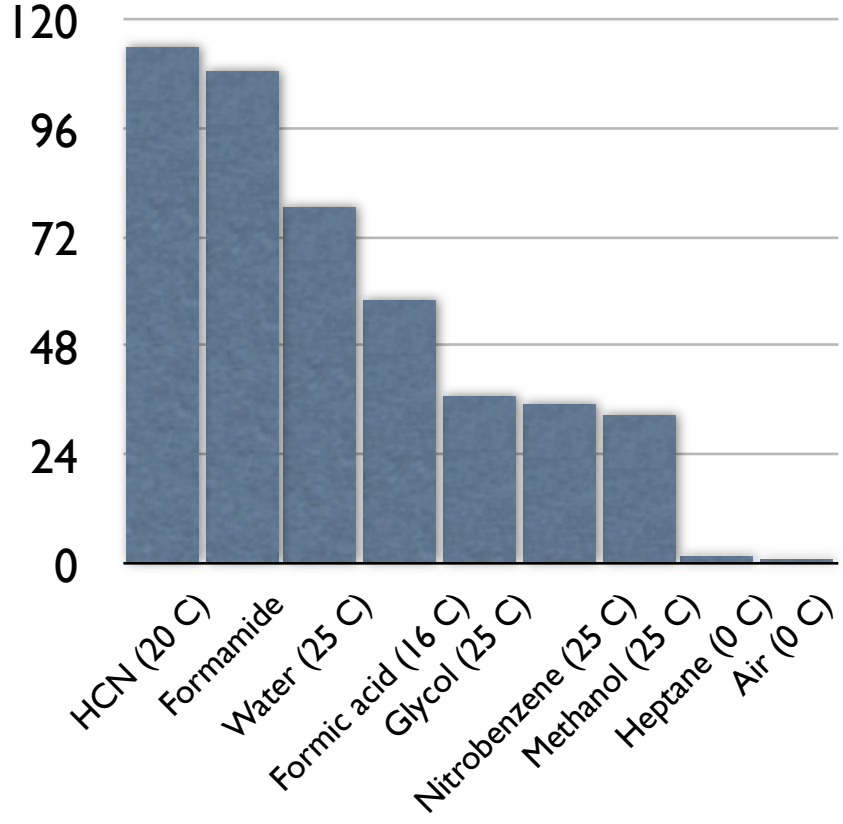
$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) = \rho(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

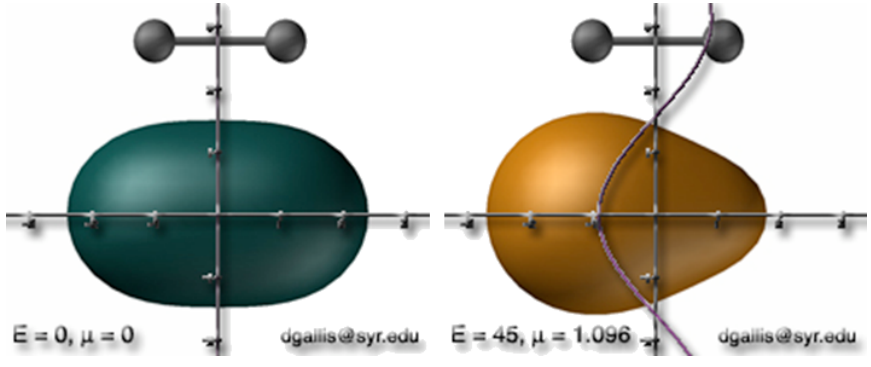
$$\phi(\mathbf{x}) = \phi_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial\Omega$$

Dielectric constants

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

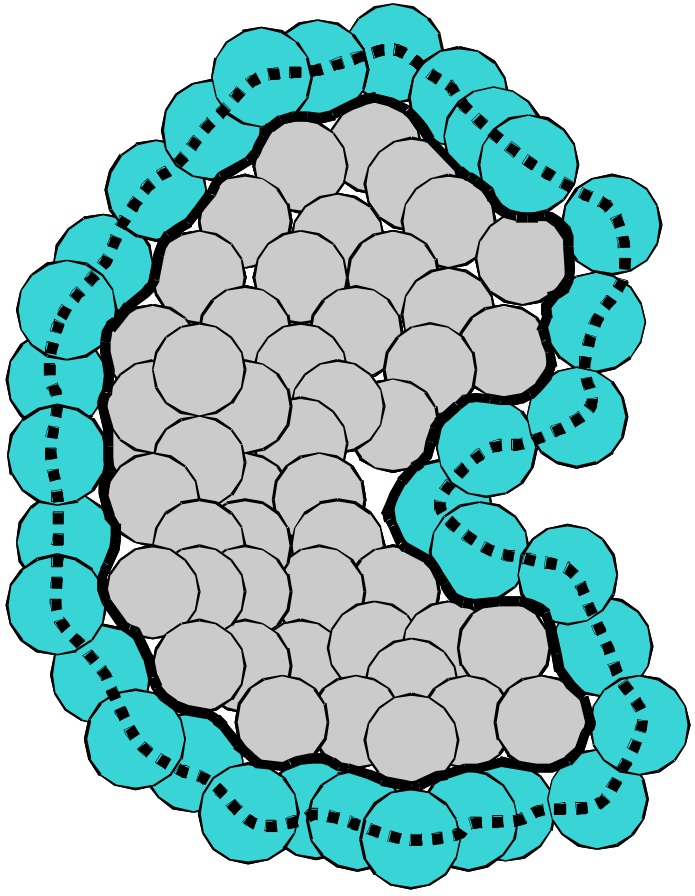






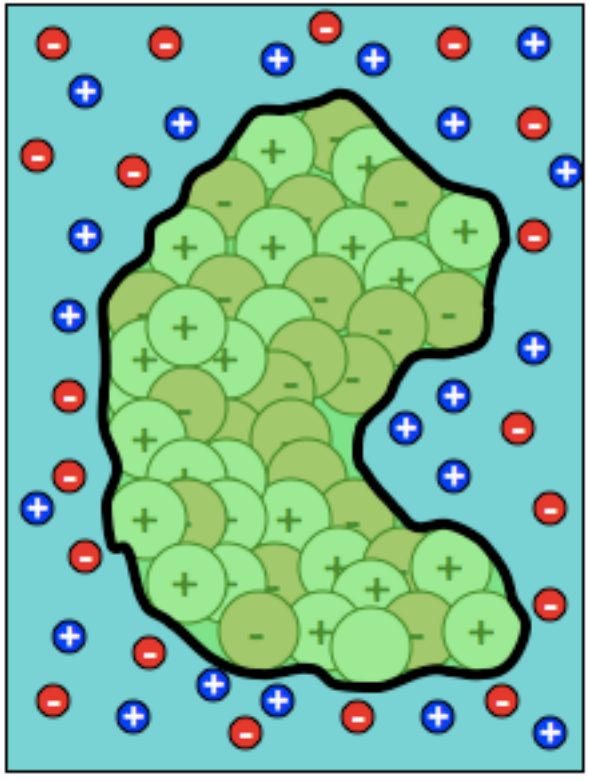
Molecular dielectric coefficients

- 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



Poisson-Boltzmann equation

- Mean field model
  - Include "mobile" charges
  - Boltzmann distribution
  - Ignore charge-charge correlations
    - Infinitesimal size
    - Weak ion-ion electrostatic interactions
    - Low densities
- Missing ion "chemistry"
  - No detailed ion-solute or ion-solvent interactions
  - No ion coordination, etc.

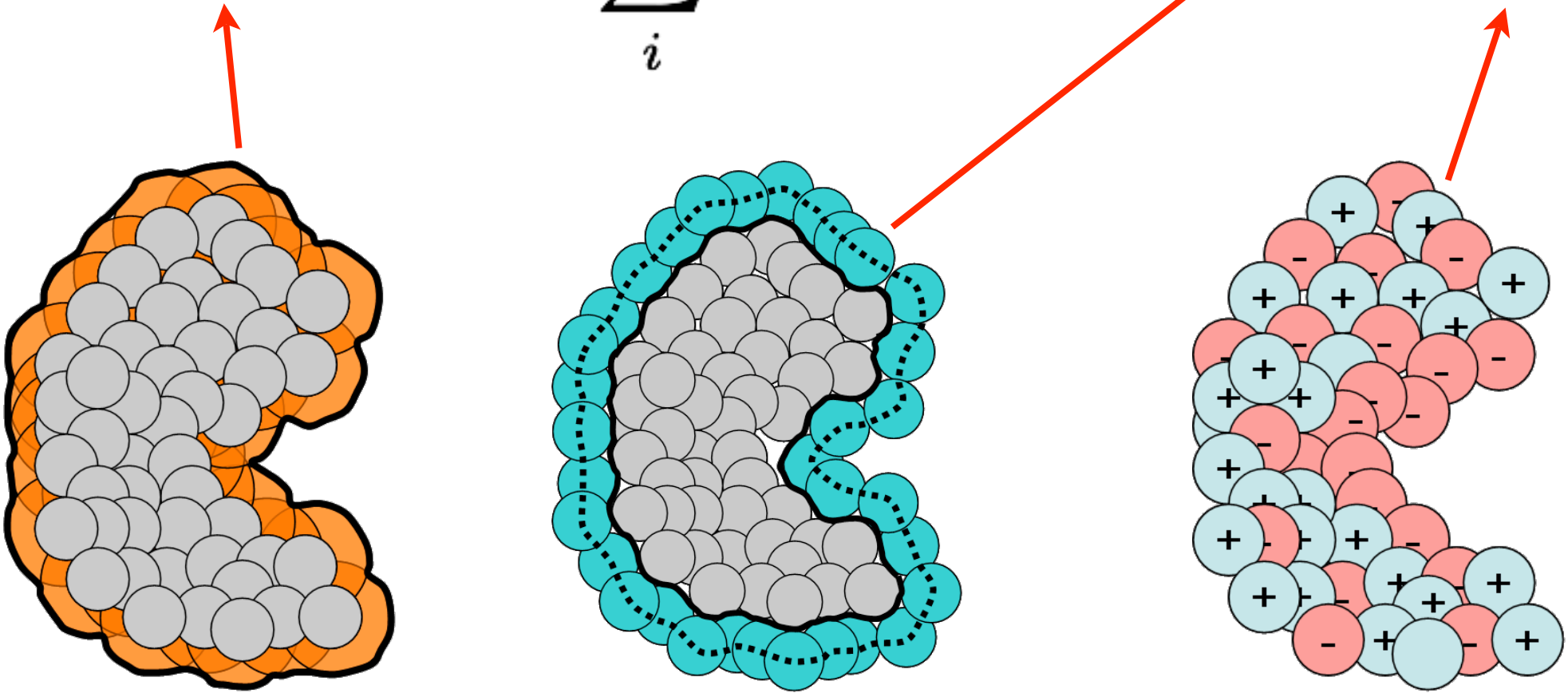


$$-\nabla \cdot \epsilon(\mathbf{x}) \nabla \phi(\mathbf{x}) = \rho_f(x) + \sum_i c_i q_i \exp\left(-\beta\left(q_i \phi(\mathbf{x}) + V_i(\mathbf{x})\right)\right)$$

The Poisson-Boltzmann equation: methods

Poisson-Boltzmann coefficients

$$-\nabla \cdot \epsilon(x) \nabla \phi(x) - \sum_i q_i c_i e^{-\beta(q_i \phi(x) + V_i(x))} = \rho(x)$$

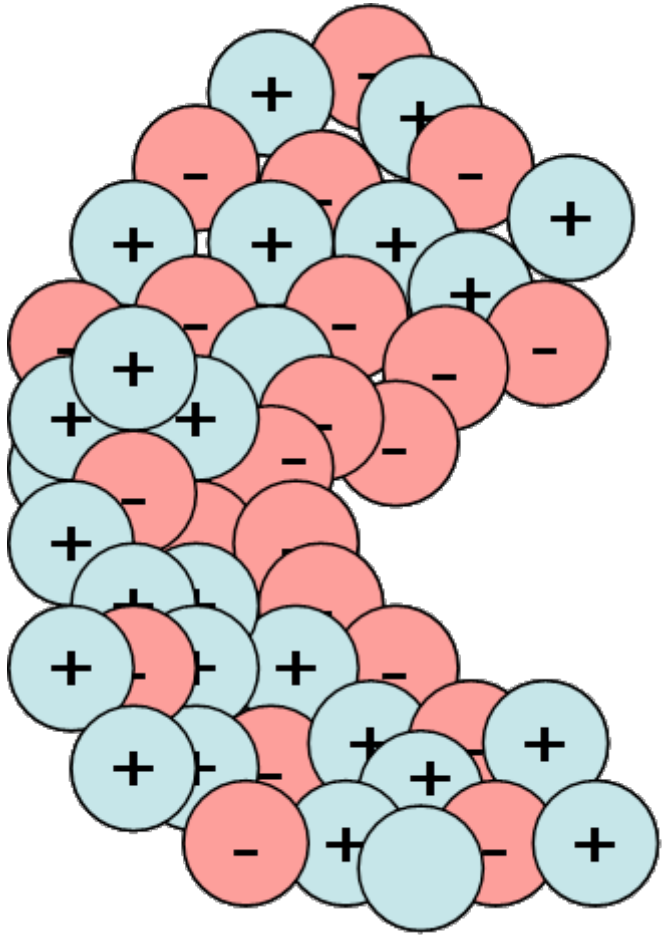


Baker NA, *Meth Enzymology*, **383**, 94-118, 2004; Baker NA, *Curr Opin Struct Biol*, **15**, 137-43, 2005.

Equation coefficients: "fixed" charge distribution

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

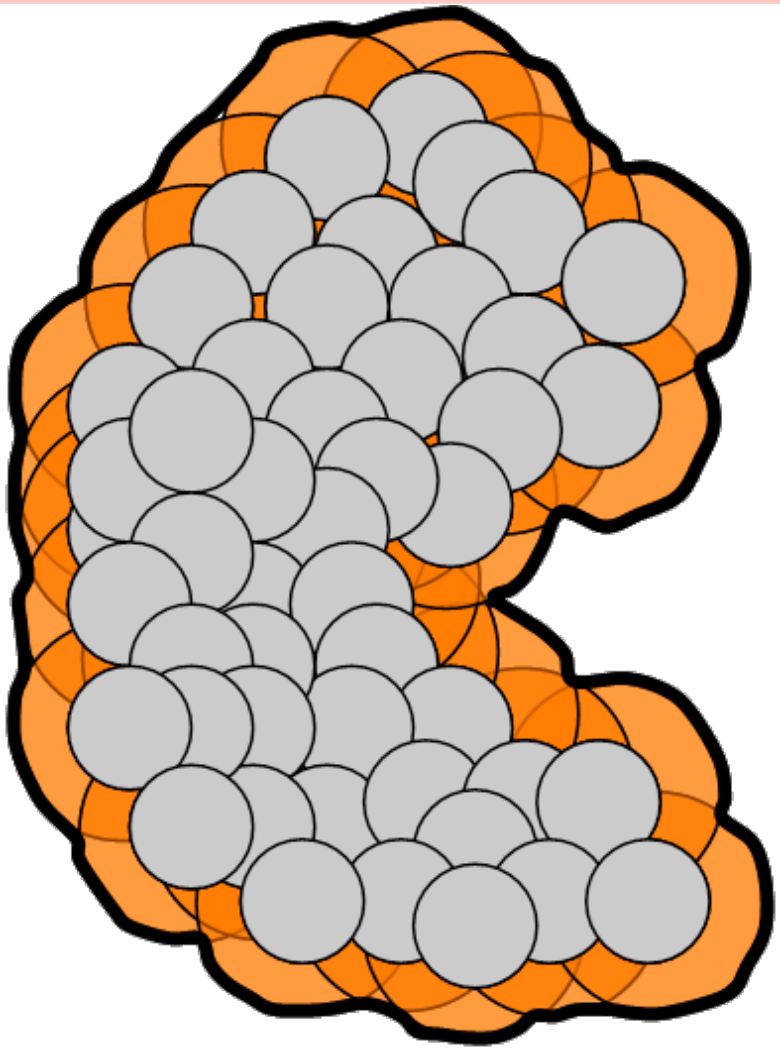
- Charges are *modeled as* delta functions: hard to represent
- Often discretized as splines to "smooth" the problem
- Higher-order charge distributions also possible



Equation coefficients: mobile ion distribution

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

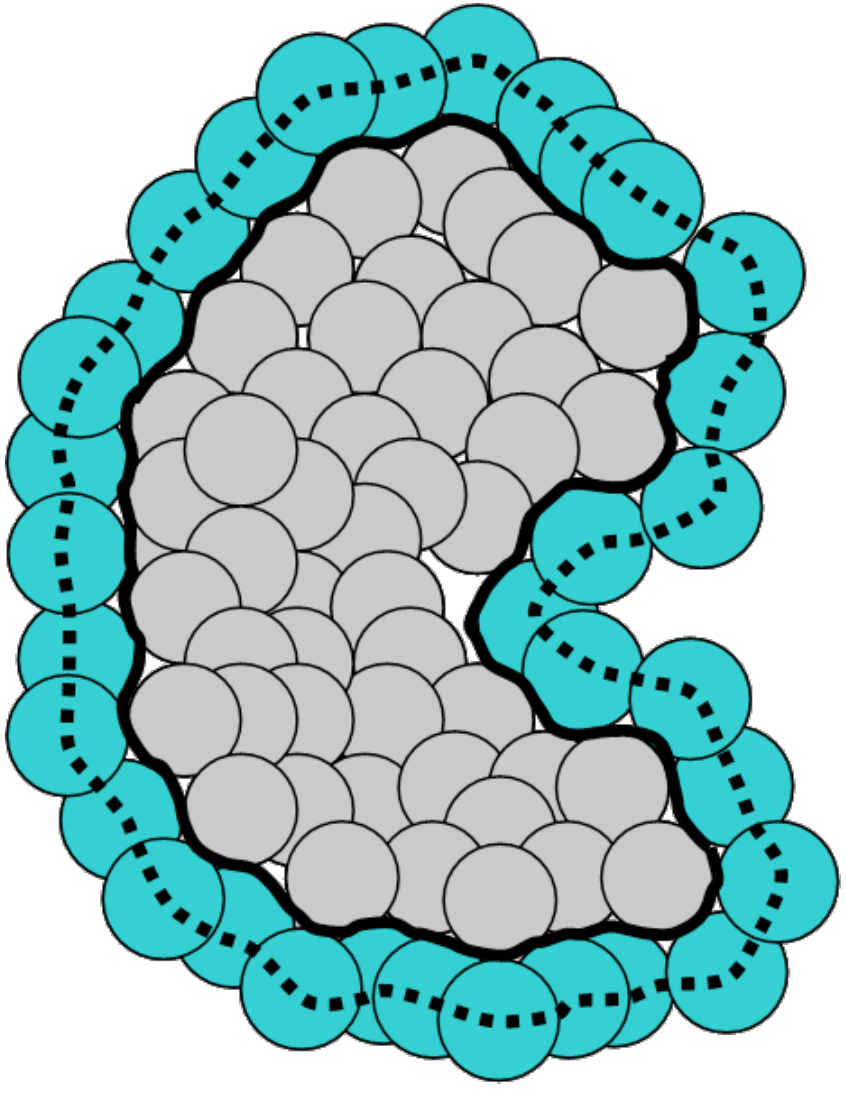
- Usually assume a single exclusion function for all ions
- Generally based on inflated van der Waals radii



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!

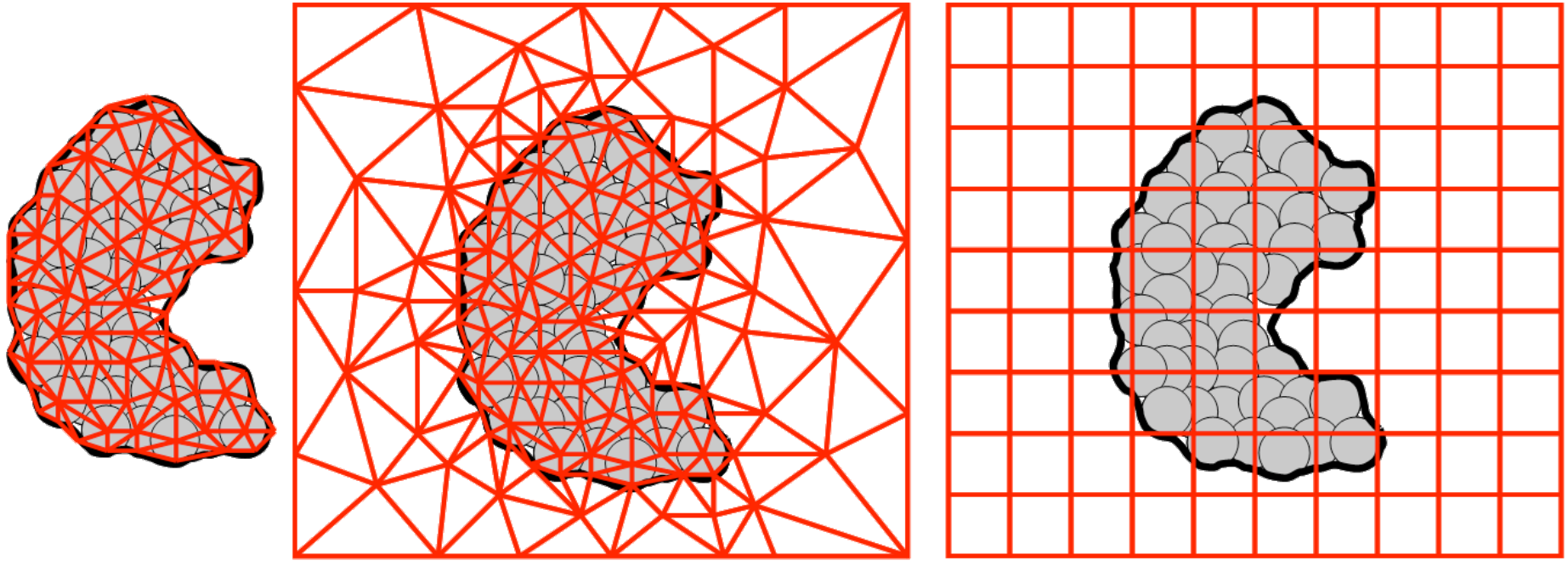


Basic steps

1. Determine the coefficients based on the biomolecular structure
2. Discretize the problem
3. Solve the resulting linear or nonlinear algebraic equations

Discretization methods

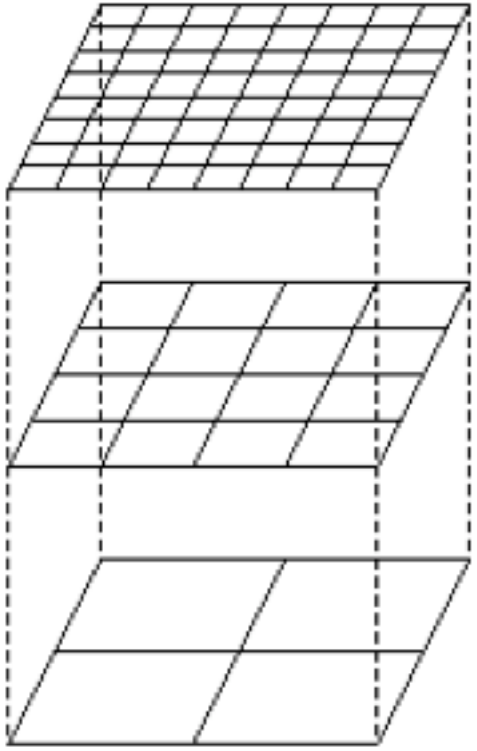
- Locally supported polynomial basis functions
- Sparse matrices
- Boundary element methods
- Finite element methods (Holst group FEtk)
- Multigrid (Holst group PMG)



Holst M, et al. *J Comput Chem* **21**, 1319-42, 2000; Baker N, et al. *J Comput Chem* **21**, 1343-52, 2000.

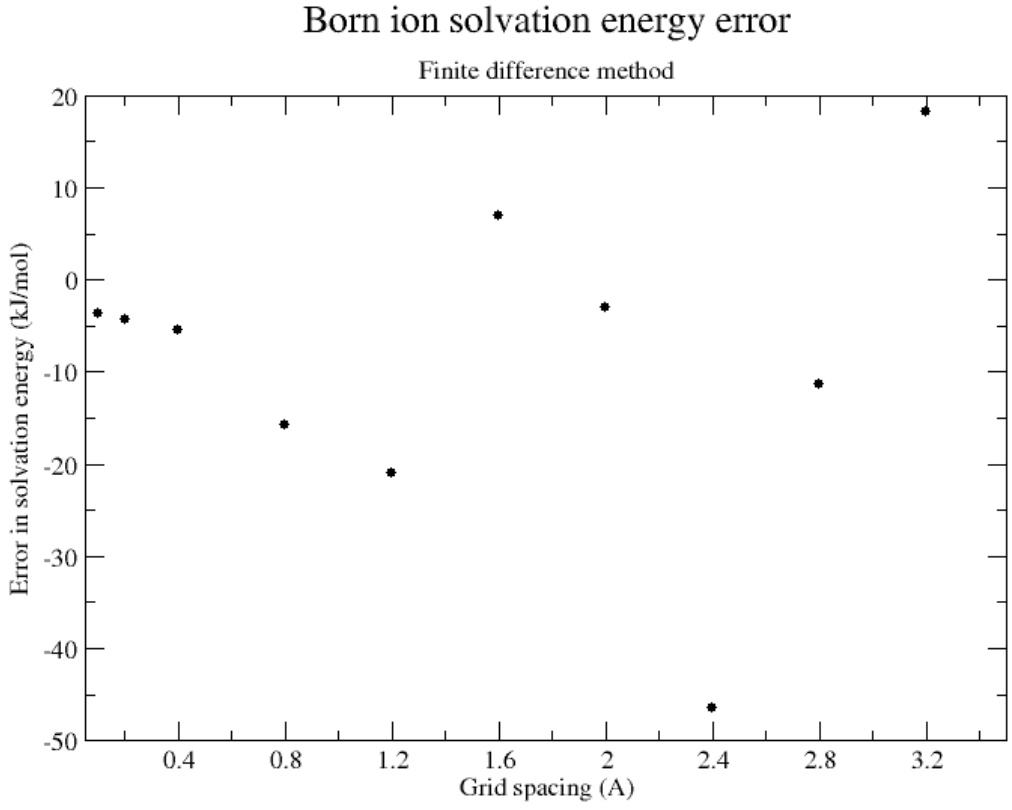
Numerical solvers

- Holst group software (FEtk and PMG)
- Iteratively solve algebraic systems:
  - Linear: multigrid
  - Nonlinear: Newton's method and multigrid
- Multigrid solvers offer optimal solution
- Big systems are still difficult
  - Significant memory
  - Long run-times
  - Need parallel solvers



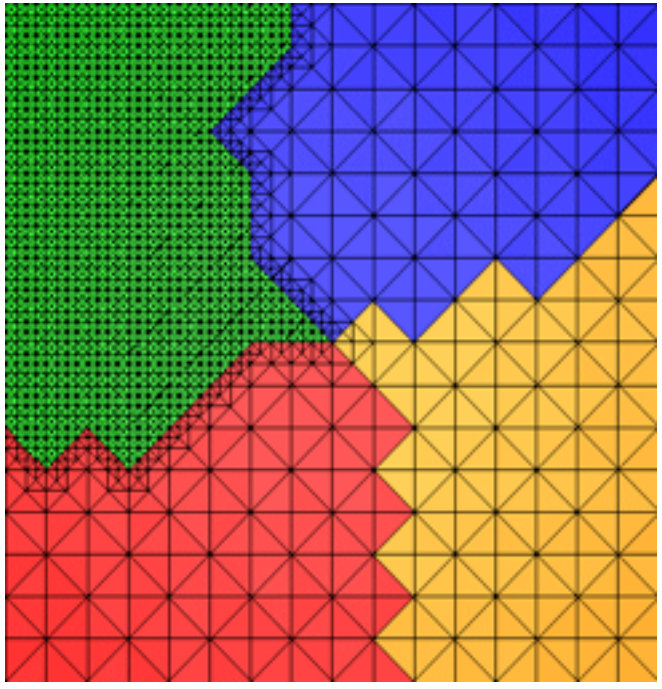
Errors in numerical solutions

- Electrostatic potentials are very sensitive to discretization
  - Grid spacings < 0.5 Å
  - Smooth surface discretizations
- Errors most pronounced next to a biomolecule
  - Large potential and gradients
  - High multipole order
- Errors decay with distance

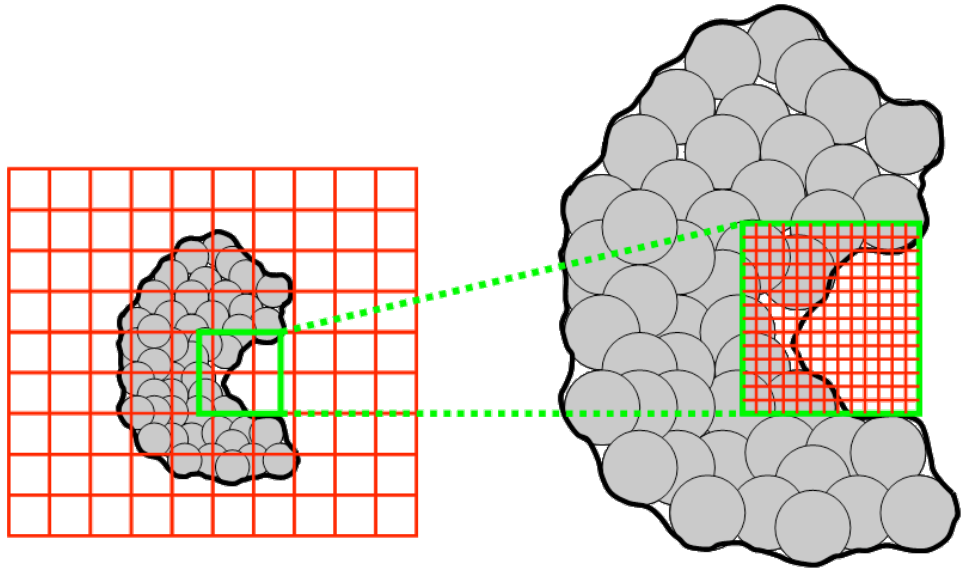


Focusing

- New parallel solution methods
- "Focusing" concept: low-resolution solutions acceptable for boundary conditions

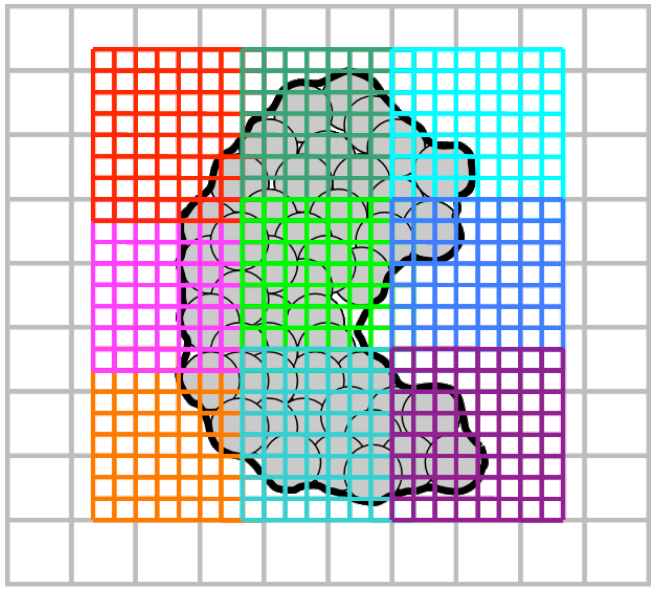


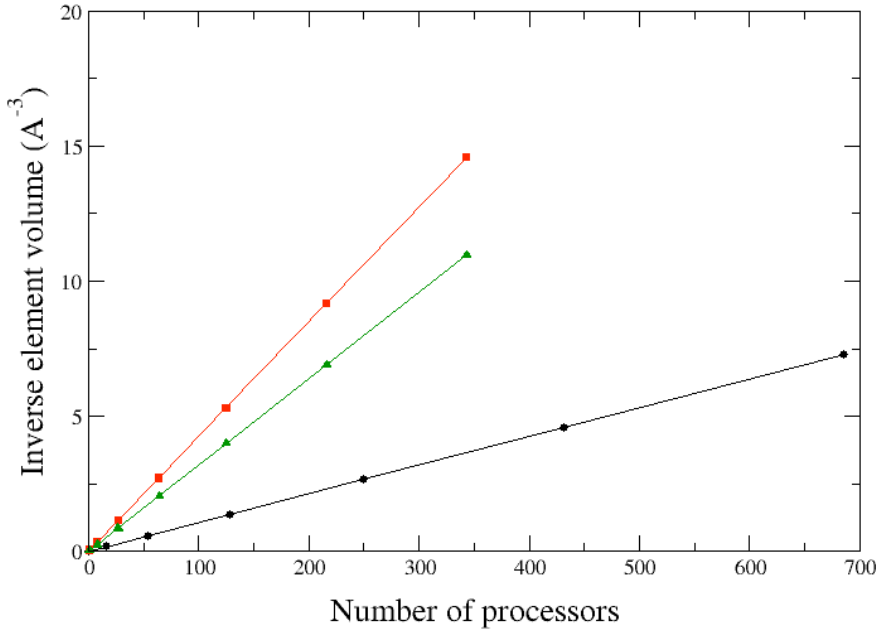
Similar to Bank-Holst parallel method



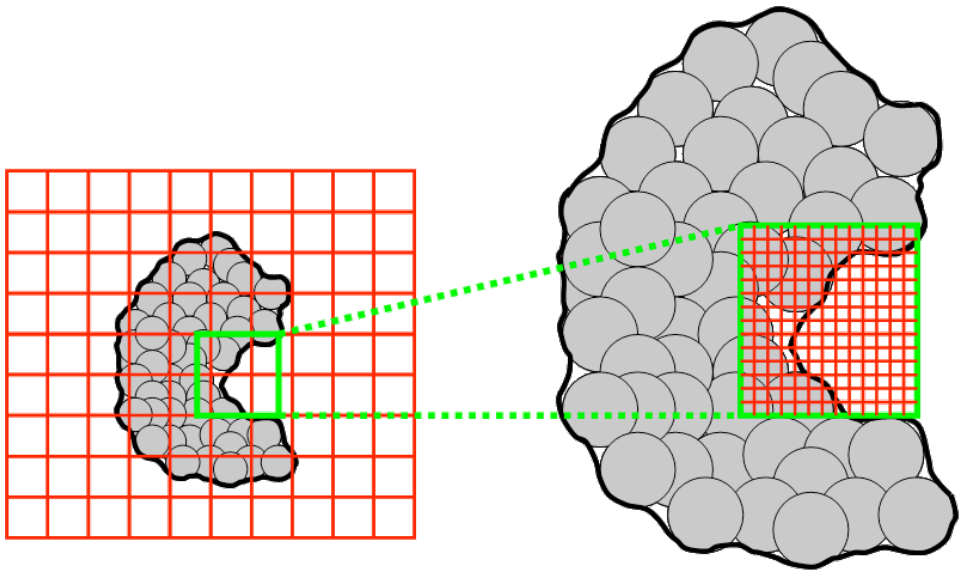
Focusing

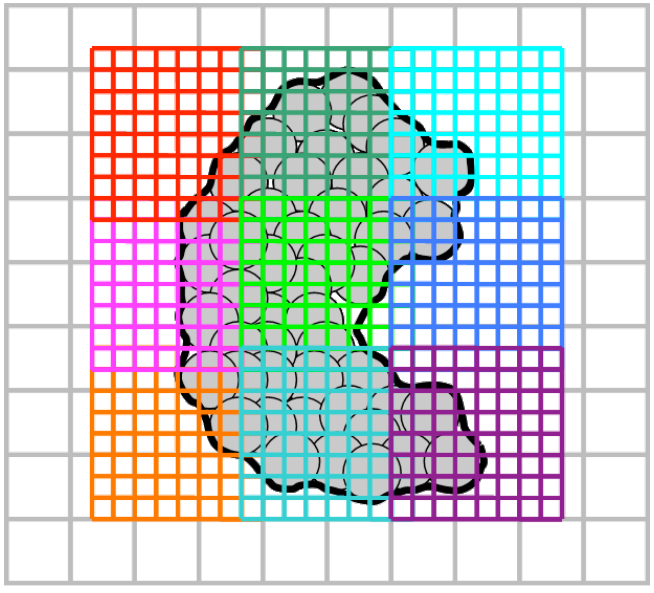
- Given the problem data and P processors of a parallel machine:
  - Each processor i = 1, …, P:
    - Obtains a coarse solution over the global domain
    - Subdivides the global domain into P subdomains, each of which is assigned a processor
    - Assigns boundary conditions to a fine discretization of its subdomain using the coarse global solution
    - Solves the equation on its subdomain
- A master processor collects observable data from other processors and controls I/O
- Very good linear scaling

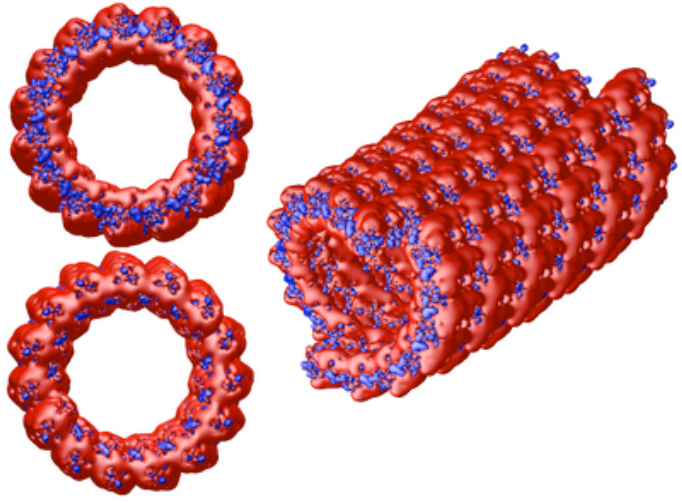


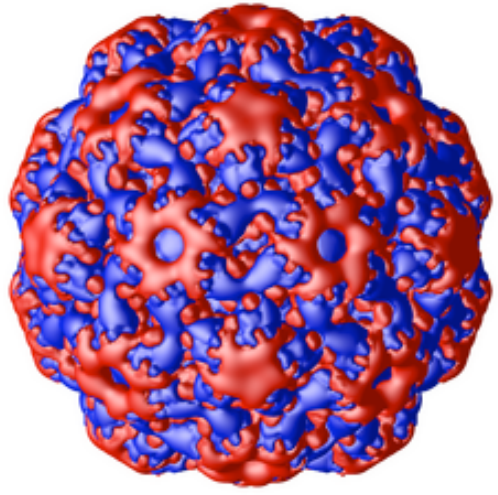


Parallel focusing









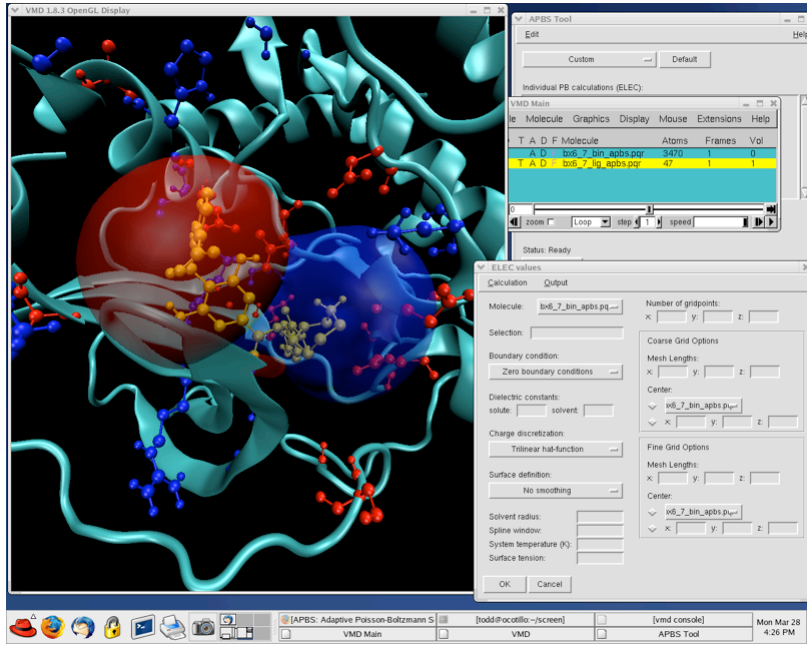
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Sept D, et al, *Protein Sci*, **12**, 2257-61, 2003; Konecny R, et al, *Biopolymers*, **82**, 106-20, 2006.

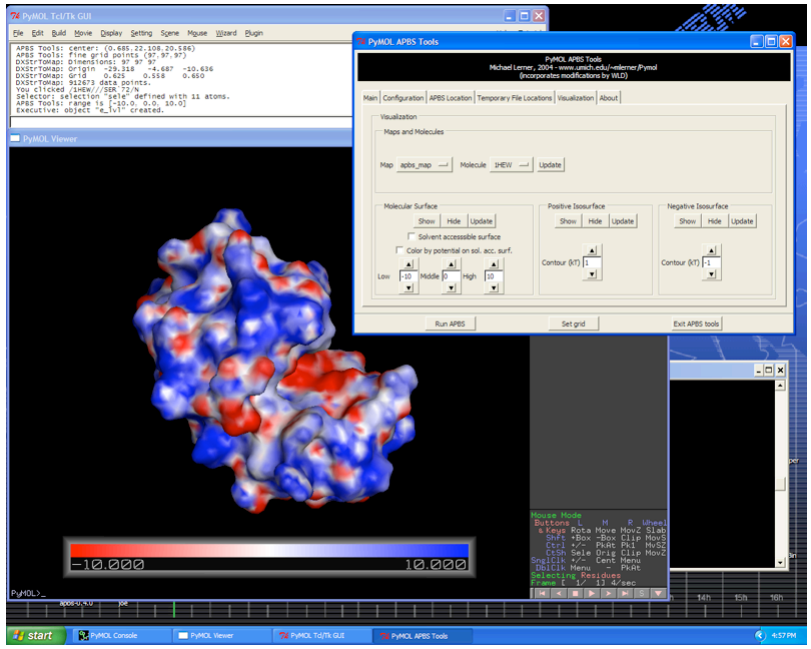
Electrostatics software

Software package	Description	URL	Availability
APBS	FD & FE MG	http://apbs.sf.net/	Free, open source
DelPhi	FD GS	http://trantor.bioc.columbia.edu/delphi	$250 academic
MEAD	FD SOR	http://www.scipps	Free, open source
UHBD	FD SOR	http://mccammon.ucsd.edu/uhbd.html	$300 academic
Jaguar	FE MG, SOR, CG	http://schrodinger.com	Commercial
CHARMM	FD MG	http://yuri.harvard.edu	$600 academic
AMBER	FD	http://amber.scripps.edu	$400 academic

Electrostatics software

- APBS (http://apbs.sf.net/)
  - PB electrostatics calculations
  - Fast finite element and multigrid solvers from Holst group
  - Web-based interface (Gemstone, Baldridge group & NBCR)
  - Works with most popular visualization software (VMD, PMV, PyMOL)
  - Links with CHARMM, AMBER, TINKER
- PDB2PQR (http://pdb2pqr.sf.net/)
  - Prepares PDB files for other calculations
  - Assigns titration states (PROPKA) and optimizes hydrogen positions
  - "Repairs" missing heavy atoms
  - Web-based and command-line



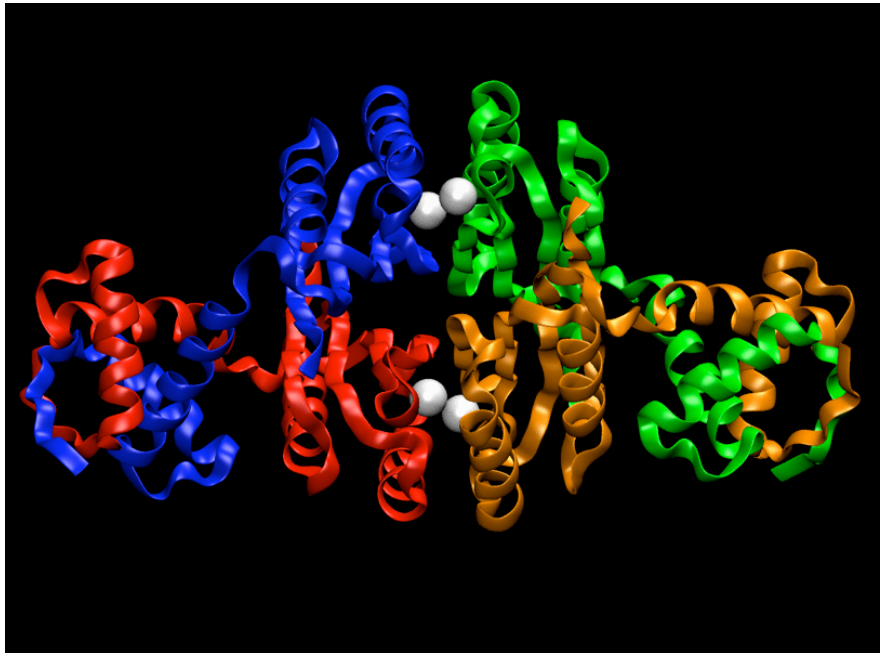


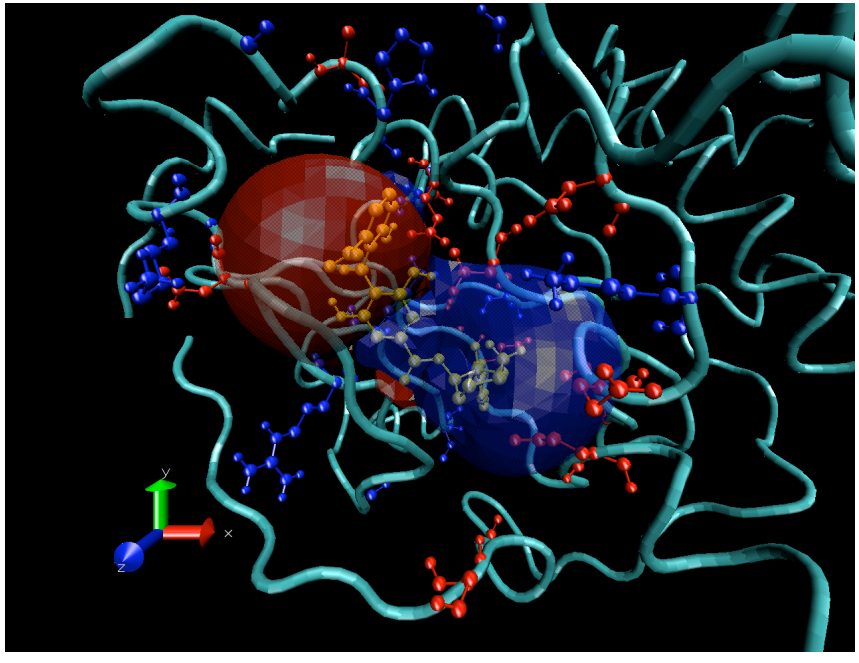
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001;
Dolinsky TJ, et al, *Nucl Acids Res*, **32**, W665-7, 2004.

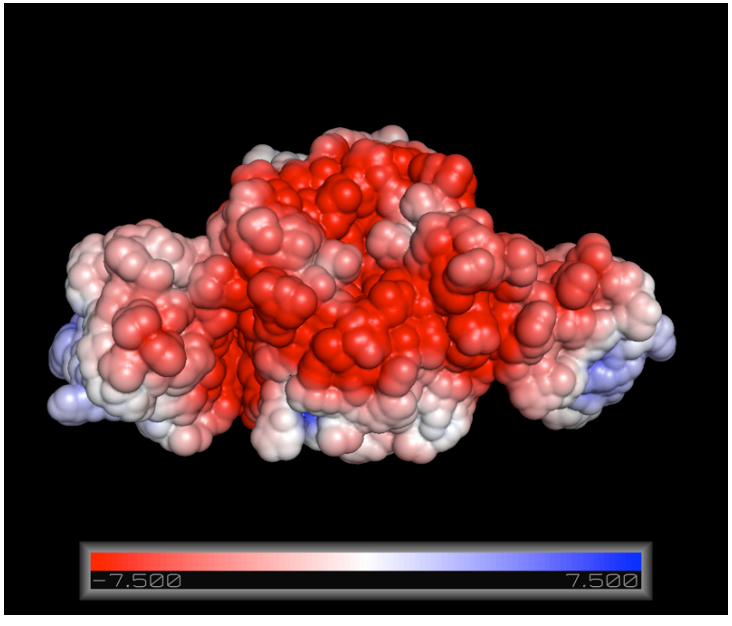
The Poisson-Boltzmann equation: applications

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

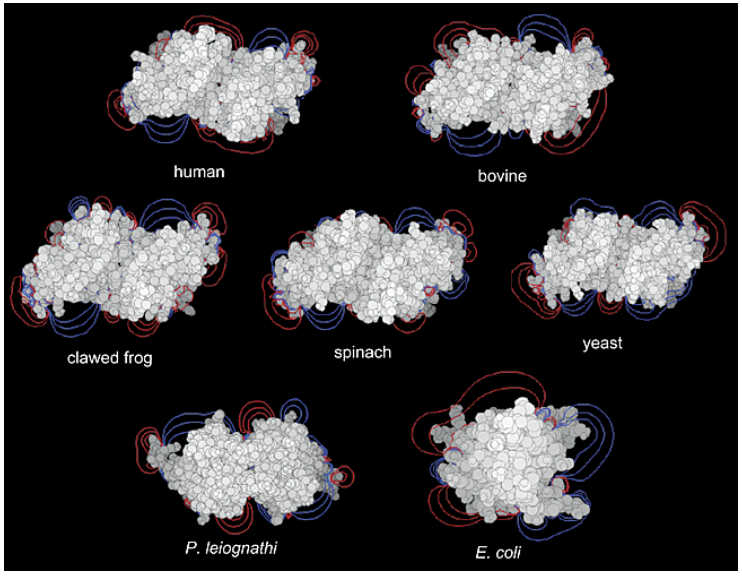




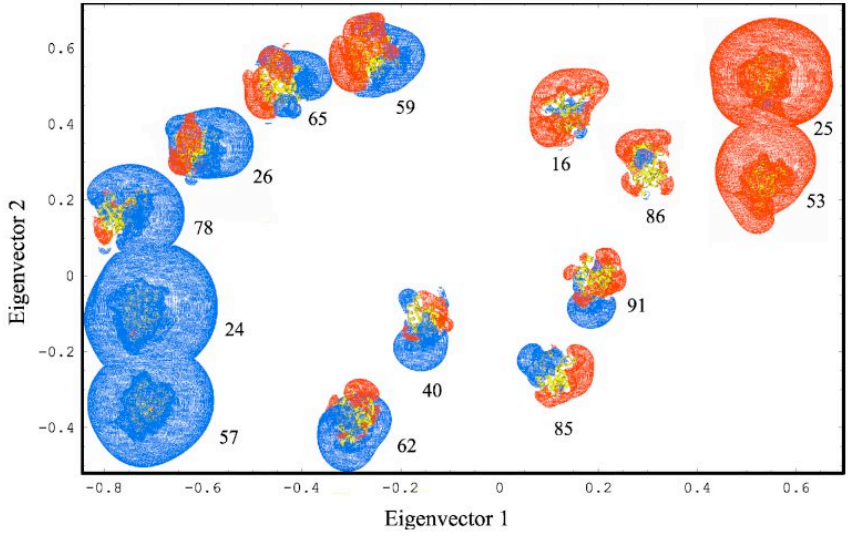


Quantitative comparison of electrostatic potentials

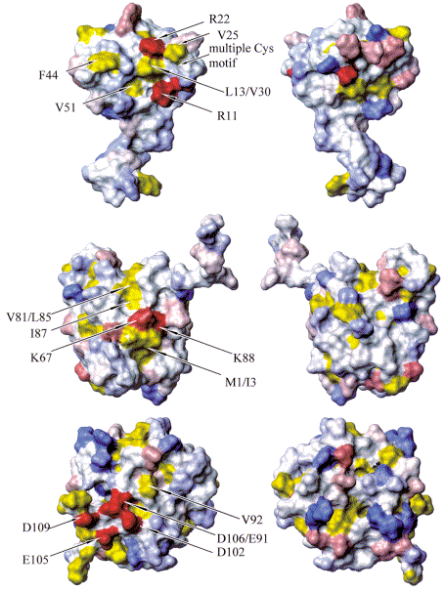
- Do electrostatic potentials tell us anything about biomolecular function?
  - Ligand binding
  - Active sites or shifted pK_as?
  - Structural (de)stabilization?



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

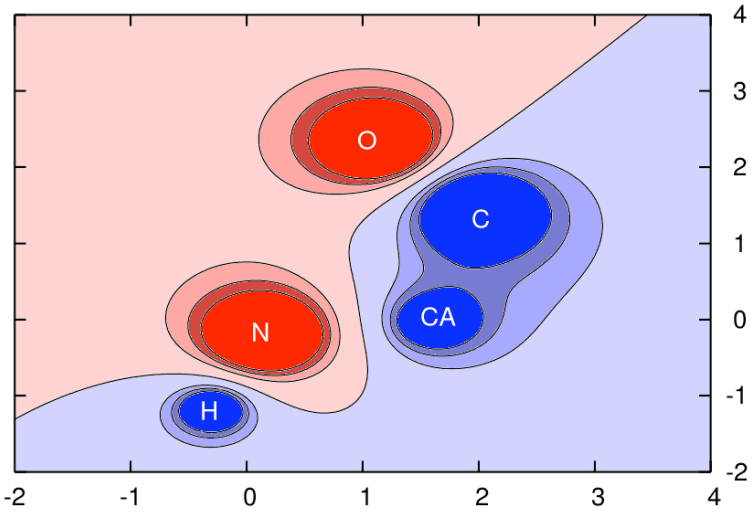


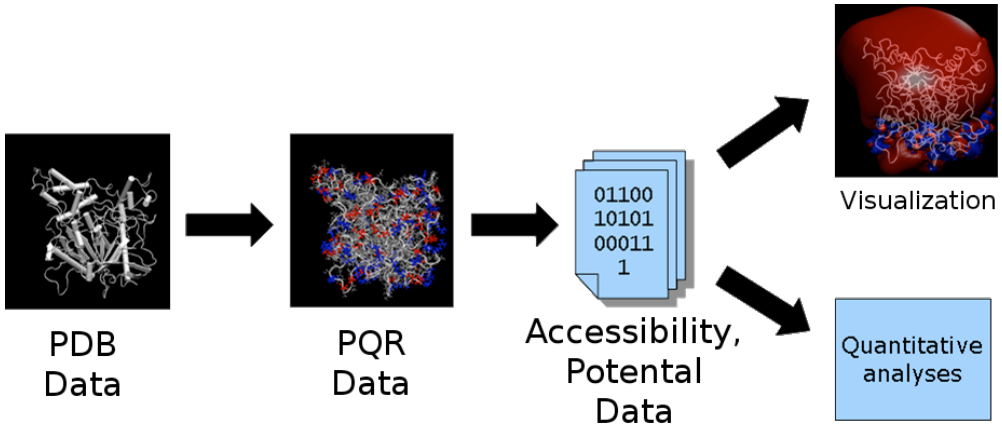
PH domain comparison: similar fold, similar electrostatics, different sequence. Blomberg N, et al. *Proteins* 37, 379-877 (1999).

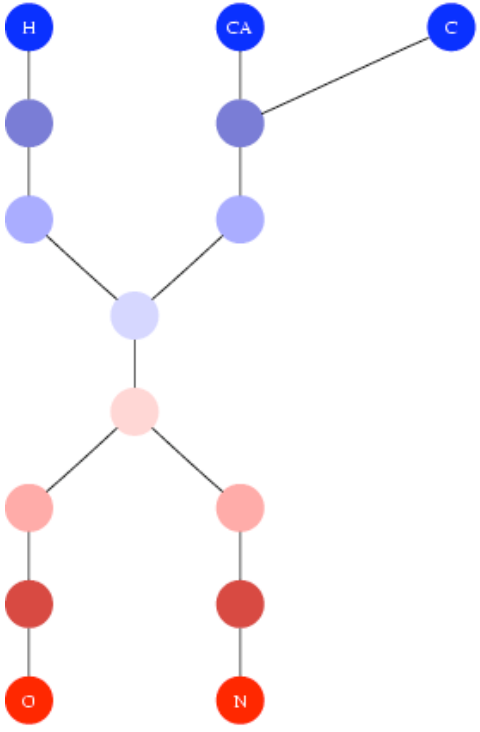


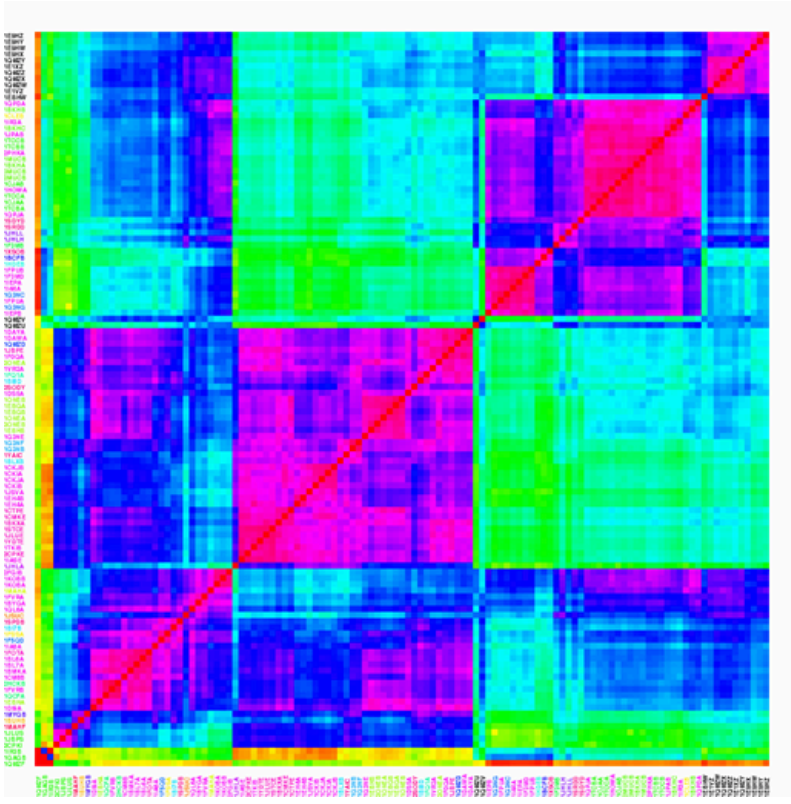
Putative Mg^{++} and RNA binding domains in an unclassified protein. Elcock AH, *J Mol Biol* 312, 885-96 (2001).

Multiresolution contour trees



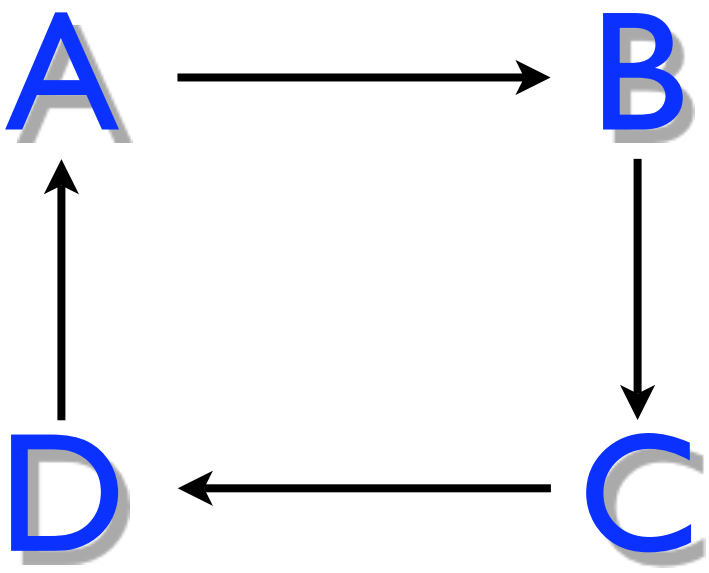






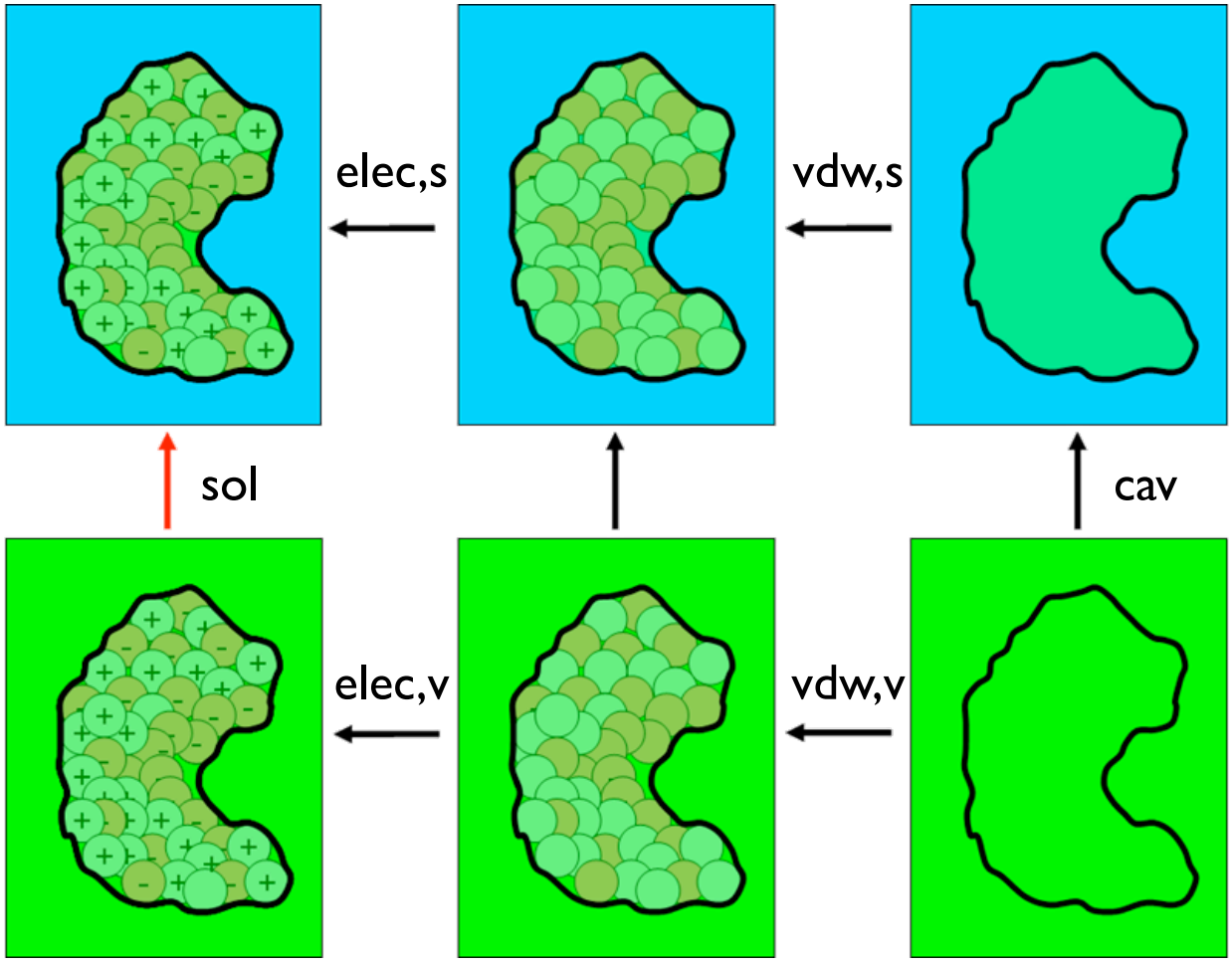
Free energy cycles

- 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



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

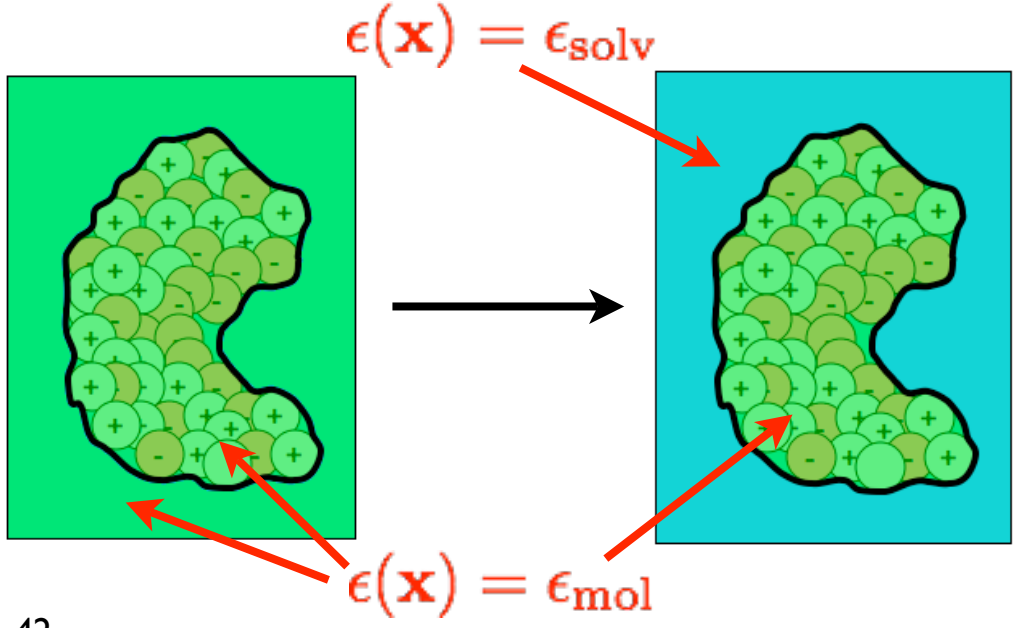
$$\Delta G_{\mathrm{sol}} = \Delta G_{\mathrm{elec,s}} - \Delta G_{\mathrm{elec,v}} + \Delta G_{\mathrm{vdw,s}} - \Delta G_{\mathrm{vdw,v}} + \Delta G_{\mathrm{cav}}$$

Solvation energies

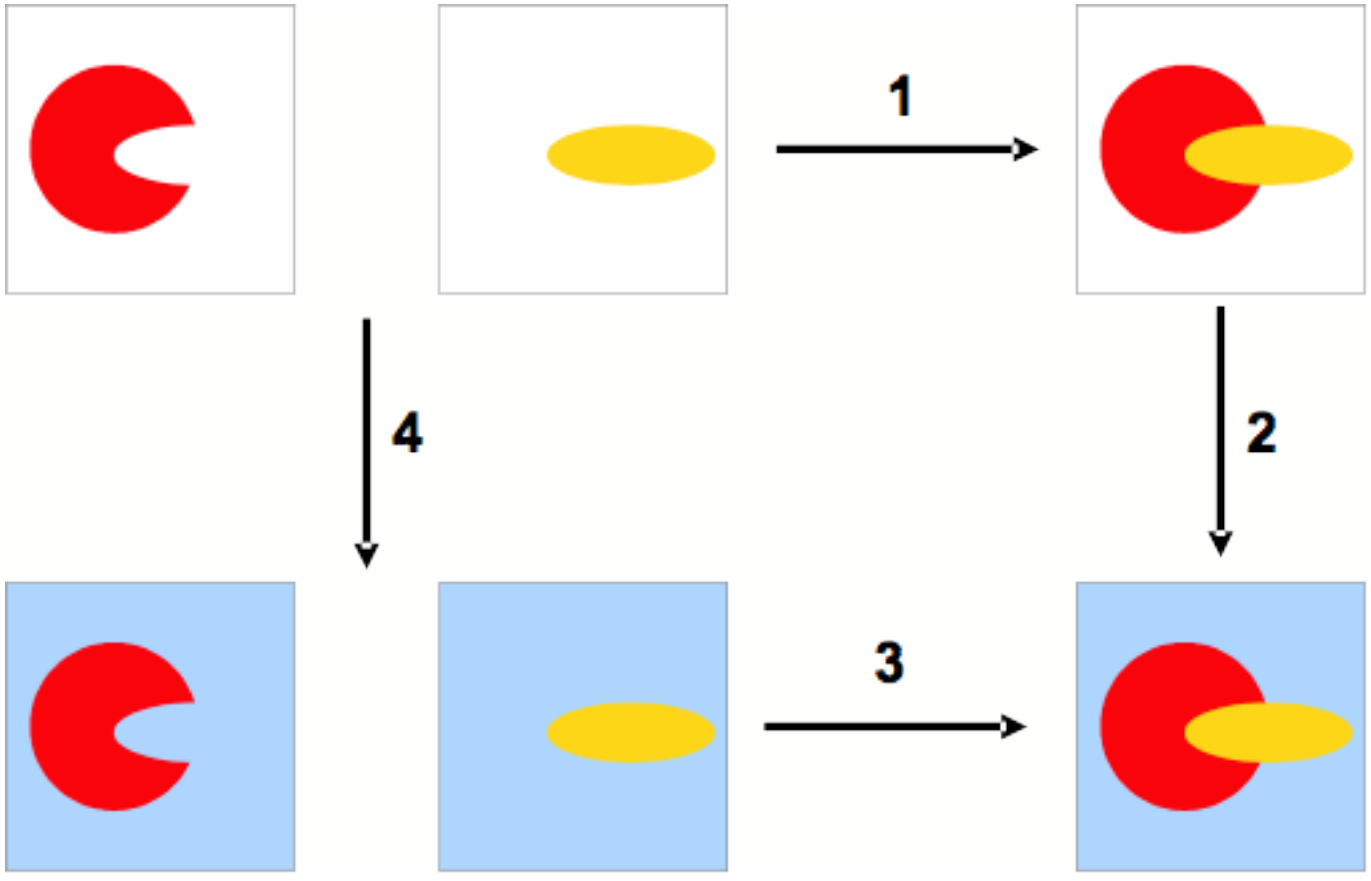
- Absolute energies are generally not useful: inaccurate
- Solvation: an excellent way to remove "self energies"
- Building block for most electrostatics calculations
- "Total energies" can be recovered by adding in vacuum polar and nonpolar contributions

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

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



Absolute binding free energy cycle

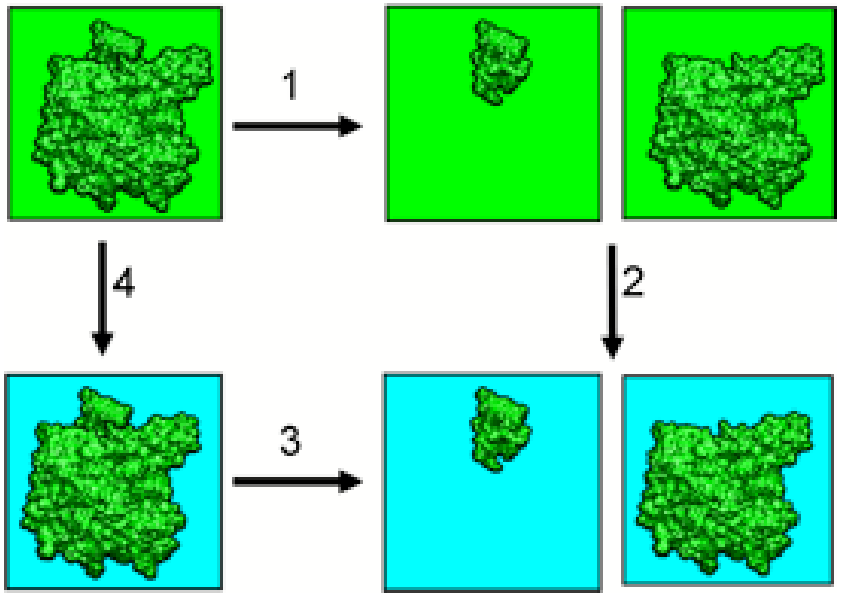


$$\Delta G_1 + \Delta G_2 - \Delta G_3 - \Delta G_4 = 0$$

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

Binding energies

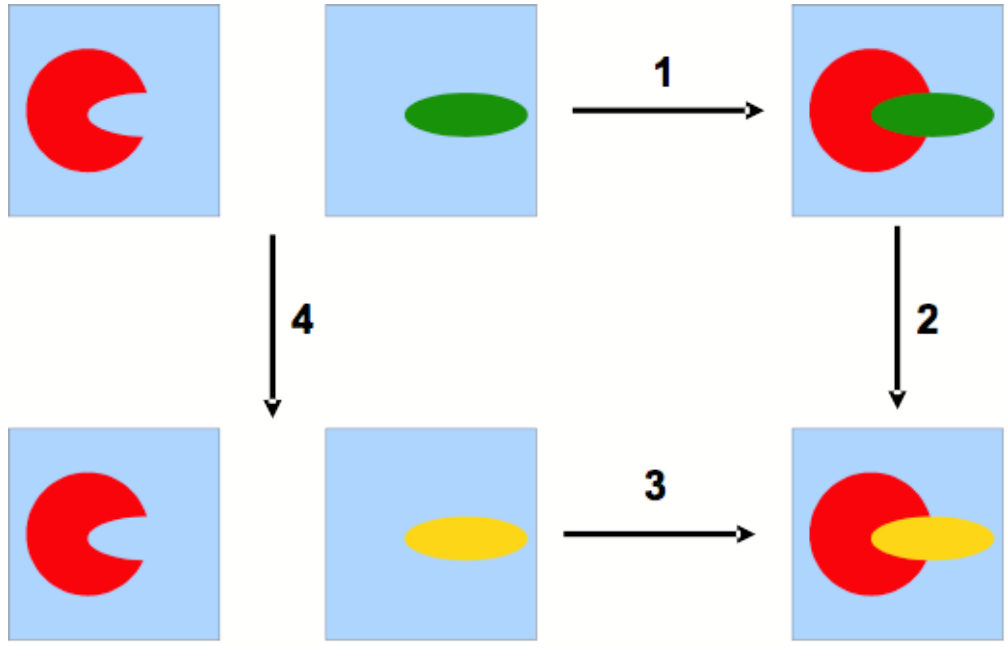
- Separate calculation into two steps:
  - Calculate electrostatic interaction for homogeneous dielectric (Coulomb's law)
  - Calculate solvation energy change upon binding
- Self-interactions are removed in solvation energy calculation
- Absolute binding energies are tricky...



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

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.

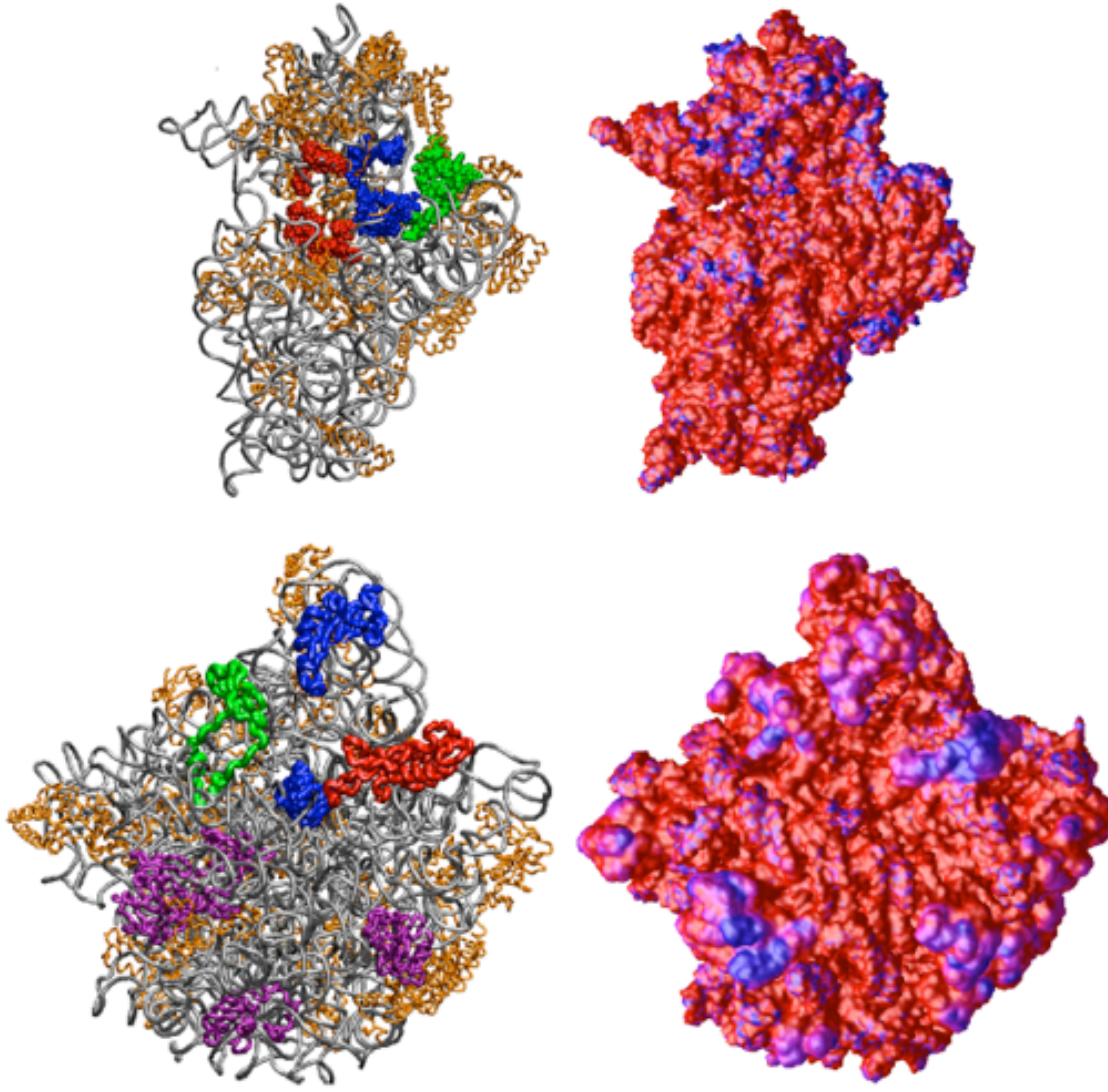


$$\Delta G_1 + \Delta G_2 - \Delta G_3 - \Delta G_4 = 0$$

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

Application to ribosomes

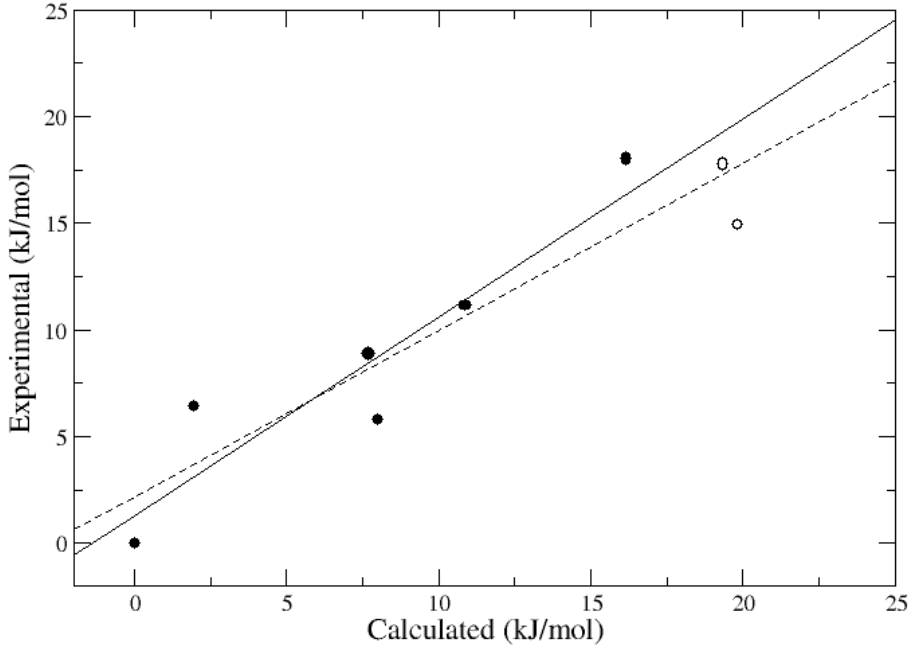
- Ribosome central to protein synthesis machinery
- Target for several pharmaceuticals
- Nucleoprotein composition make it computationally challenging
- Composed of two subunits (large and small):
  - 30S consists of 88,000 atoms and roughly 200 Å cube
  - 50S consists of more than 95,000 atoms and roughly 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

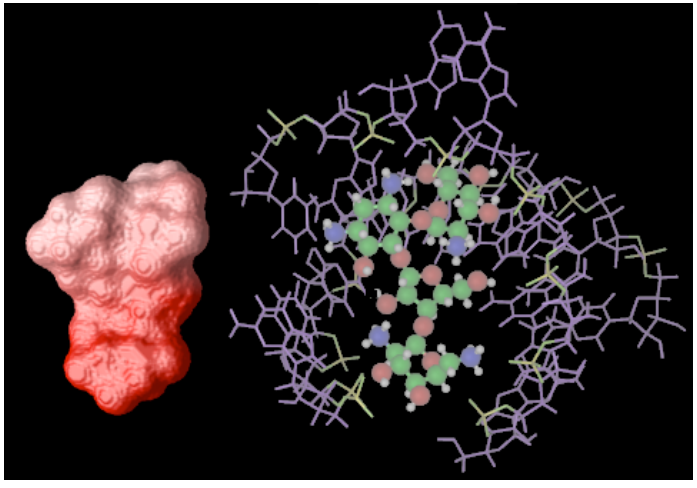


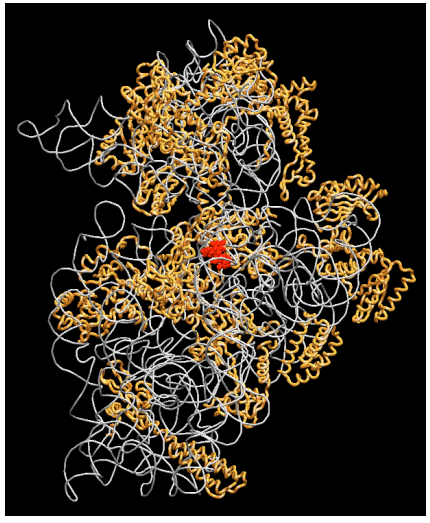
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Ma C, et al, *J Am Chem Soc*, **124**, 1438-42, 2002.

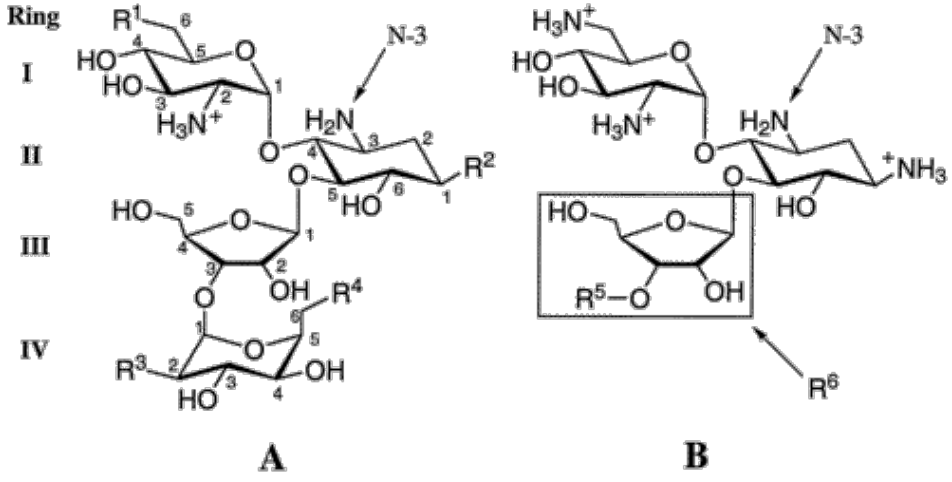
Ribosome-antibiotic binding

- 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





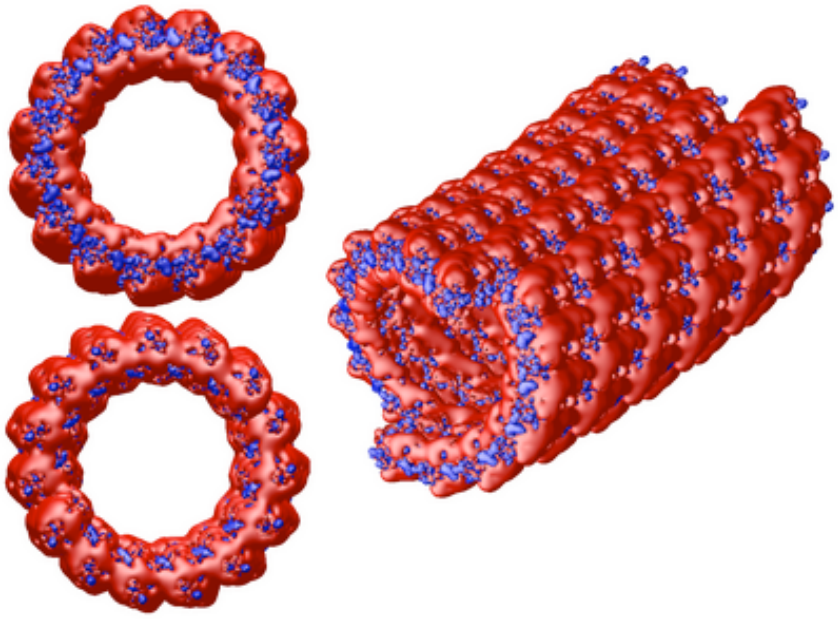


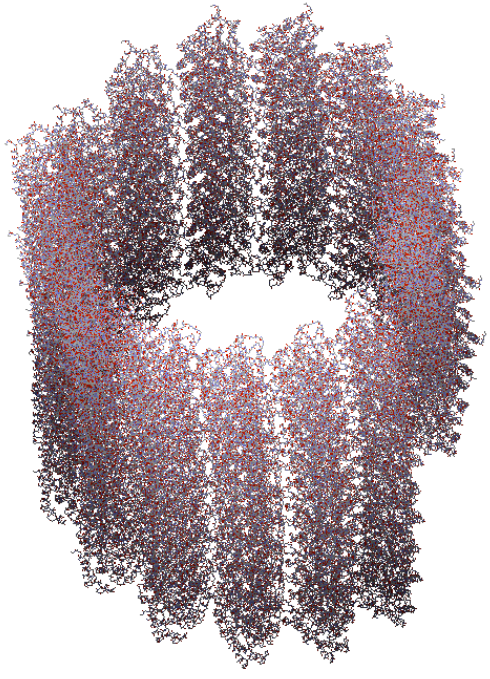


Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Ma C, et al, *J Am Chem Soc*, **124**, 1438-42, 2002.

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

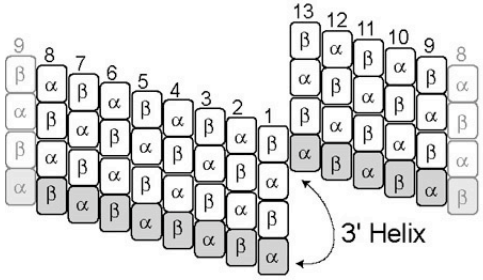




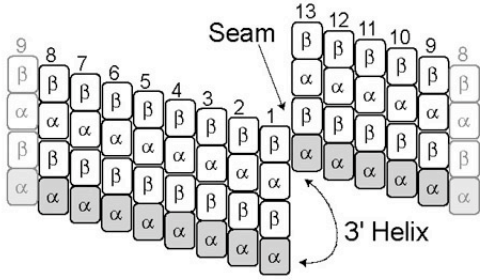
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001; Sept D, et al, *Protein Sci*, **12**, 2257-61, 2003.

Microtubule stability and assembly

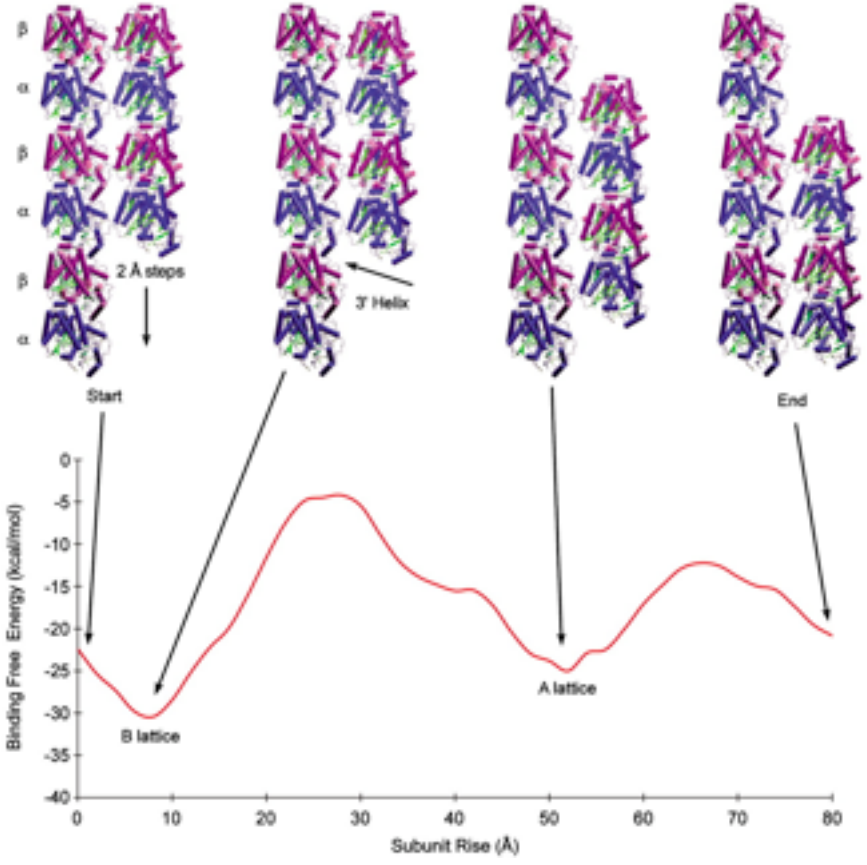
- Collaboration with Andy McCammon (UCSD) and Dave Sept (Wash U BME)
- Performed series of calculations on tubulin dimers and protofilament pairs
- Poisson-Boltzmann electrostatics and SASA apolar energies
- Observed 7 kcal/mol stronger interactions between protofilaments than within
- Determined energetics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



A lattice



B lattice



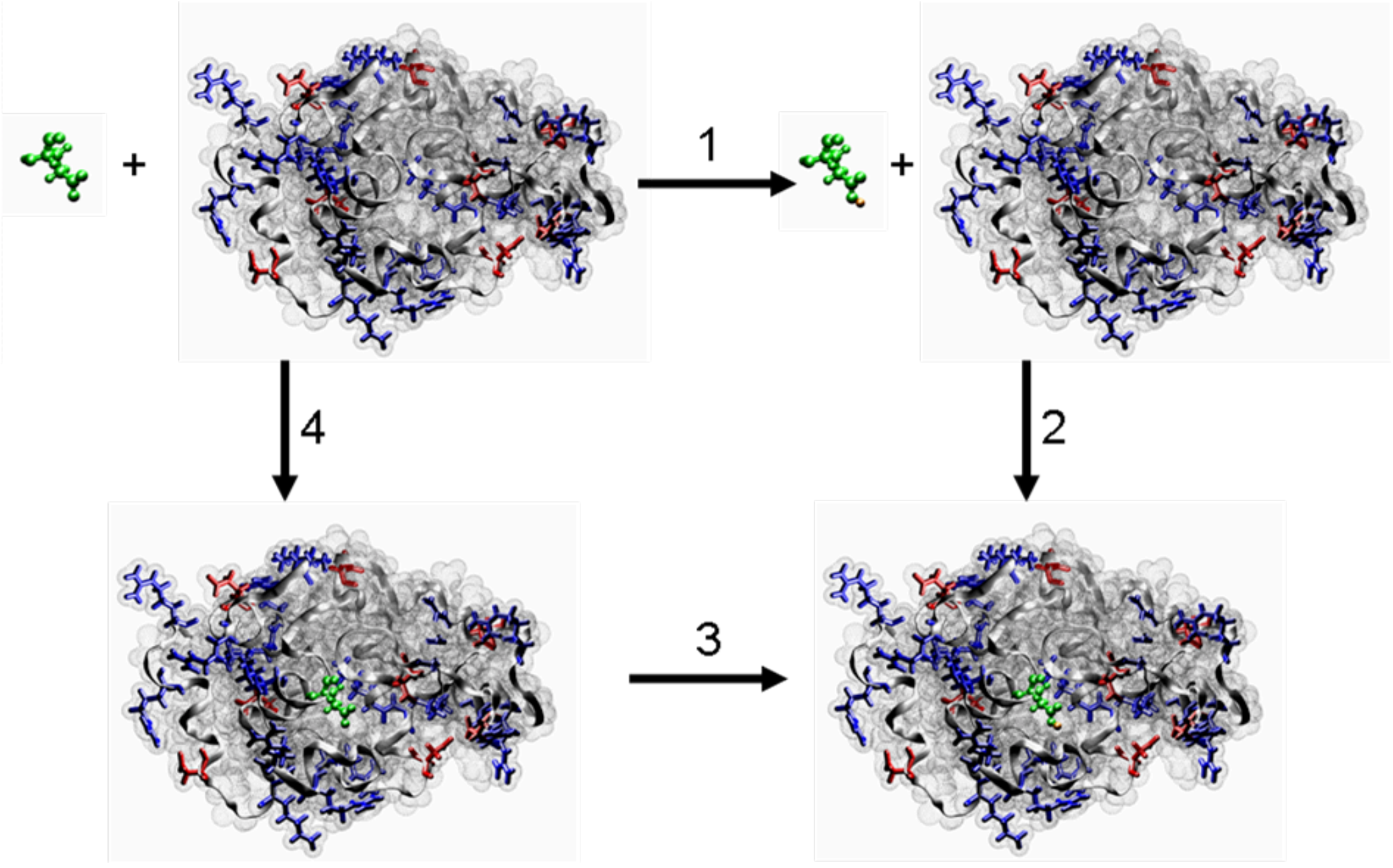
Baker NA, et al, *Proc Natl Acad Sci USA*, **98**, 10037-41, 2001;
Dolinsky TJ, et al, *Nucl Acids Res*, **32**, W665-7, 2004.

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



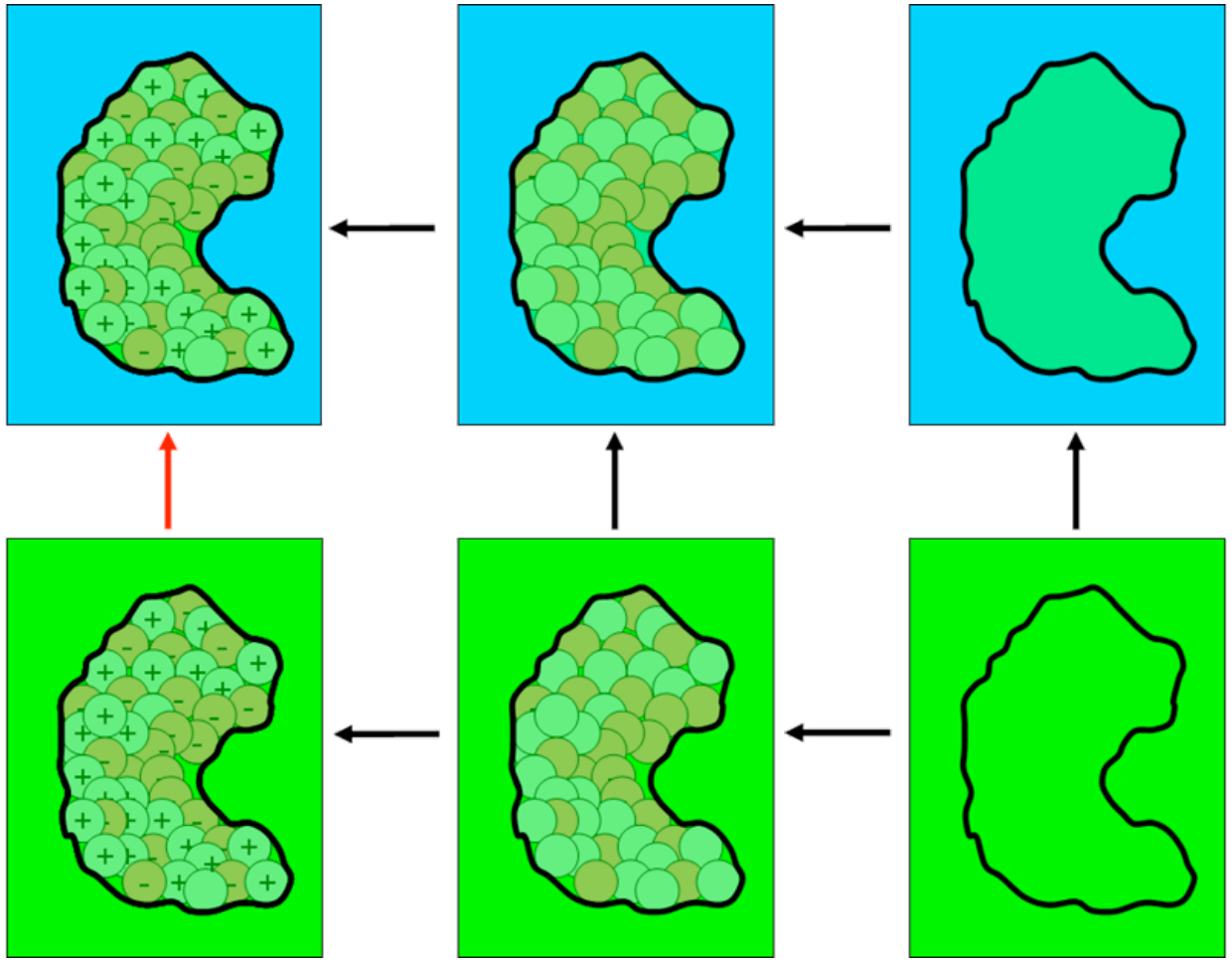
Implicit solvent models: how well do these really work?

Implicit solvent questions

- Where is the molecular detail of solvent and ions important?
- Where does nonlinear solvent/ion response matter?
- What is the correct description of nonpolar solvation?
- Ultimate goal: what is the correct interface between implicit/explicit solvent methods?

Solvation free energies

- Implicit solvent models
  - "Potentials of mean force" and solvation free energies
  - Mean forces
- Free energy cycle (as function of conformation)
  - Cavity creation
  - Dispersive interactions
  - Charging process



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

Nonpolar solvation: basic concepts

- "Oil and water"
- Solubility has temperature minimum
- Large positive heat capacity
- Temperature-sensitive structuring of water

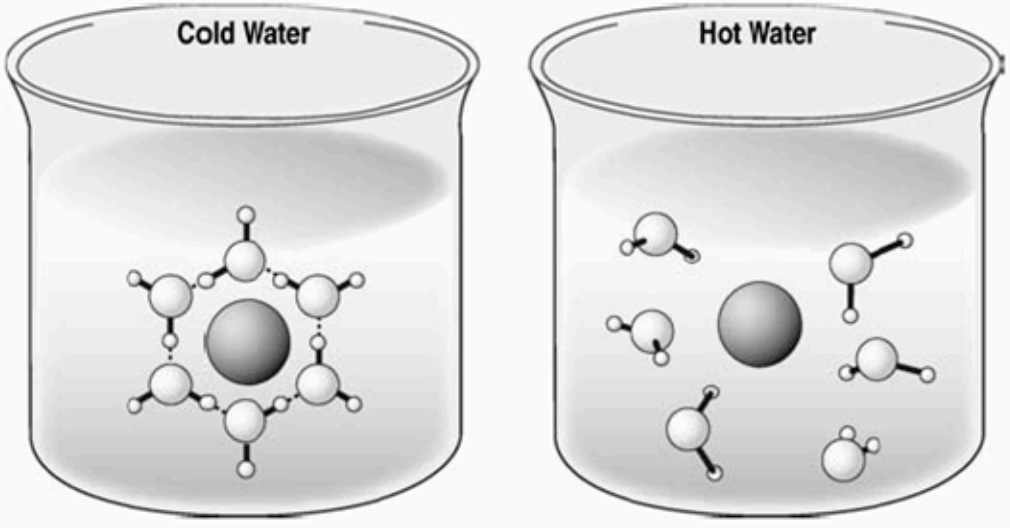
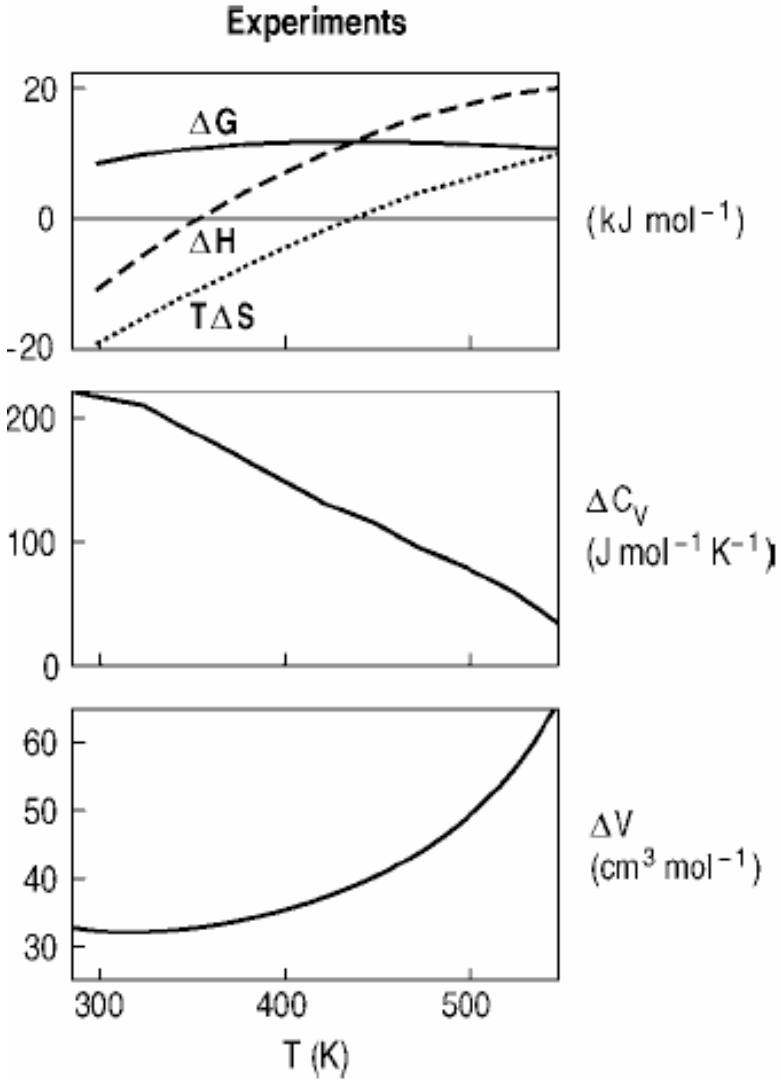


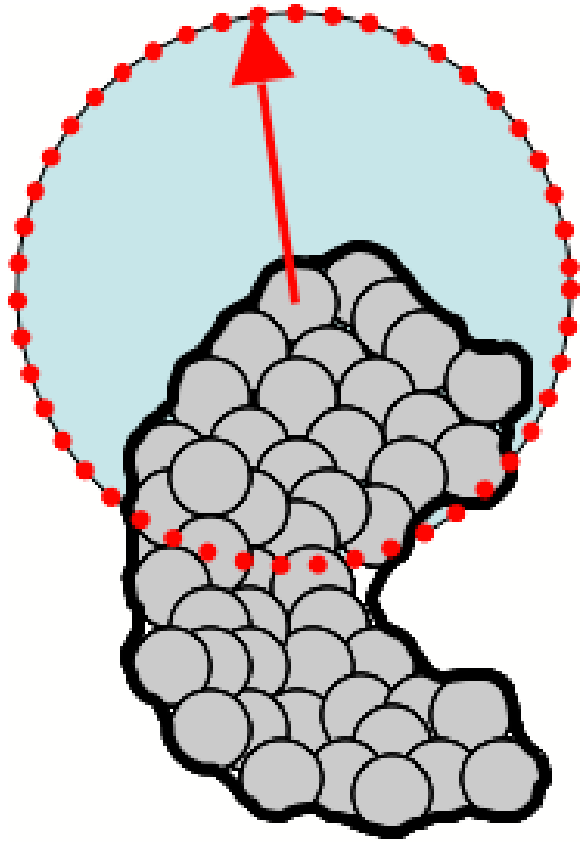
Figure 5 Iceberg model for the large heat capacity of transfer of nonpolar solutes into water.



Ar solvation and "iceberg model" figures from: Dill KA, et al. 2005. Modeling water, the hydrophobic effect, and ion solvation. Annu Rev Biophys Biomol Struct.

Nonpolar solvation: implementation

- Cavity terms
  - Area times surface tension; usually between 25 to 50 cal mol-1 Å-2 (model-dependent)
  - Volume and pressure; not commonly used
- Dispersion terms
  - WCA-like treatment: attractive potential integrated over solute-solvent distribution function for "hard" solute
  - Approximation: integral of attractive potential over solvent-accessible space



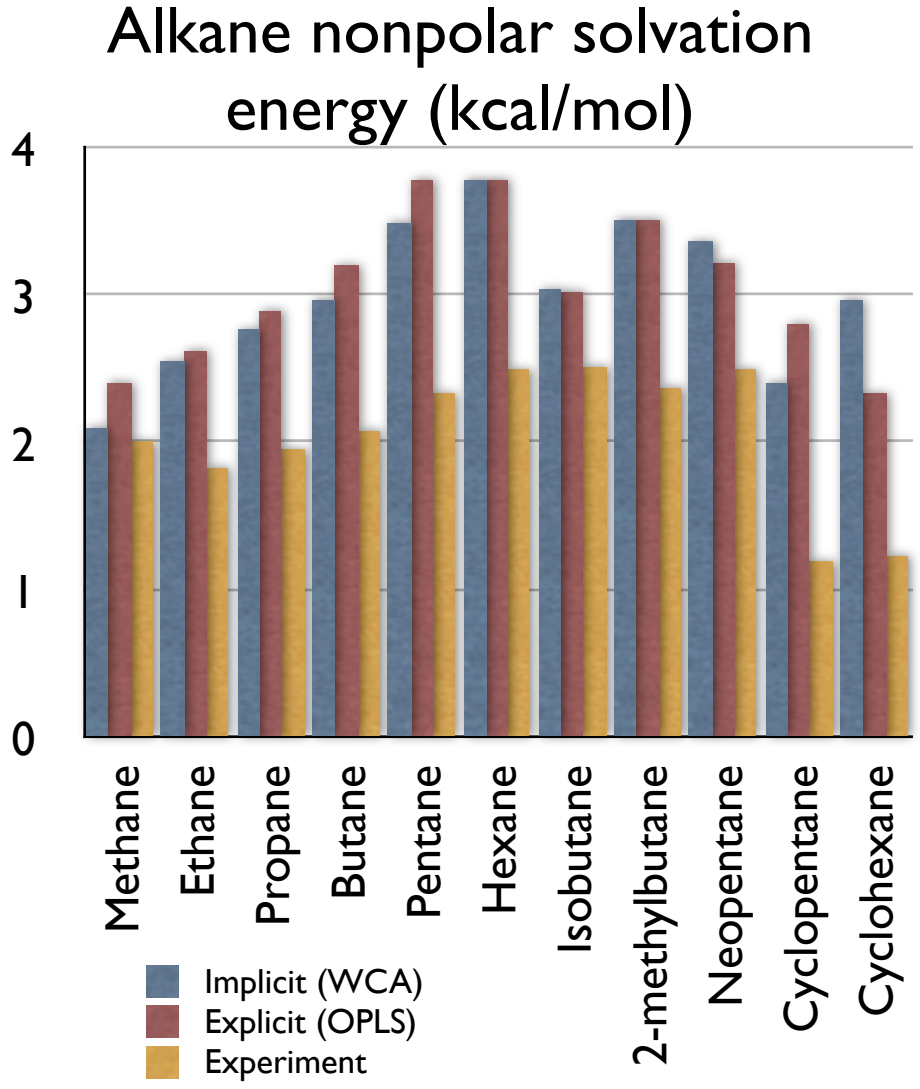
$$W^{(\mathrm{np})}(\mathbf{x}) = \gamma A(\mathbf{x};\sigma) + pV(\mathbf{x};\sigma) + \overline{\rho}\int_{\Omega} g_0(\mathbf{x},\mathbf{y};\sigma)U^{(\mathrm{np})}_{\mathrm{att}}(\mathbf{x},\mathbf{y};\sigma)d\mathbf{y}$$

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

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

Nonpolar solvation: alkanes

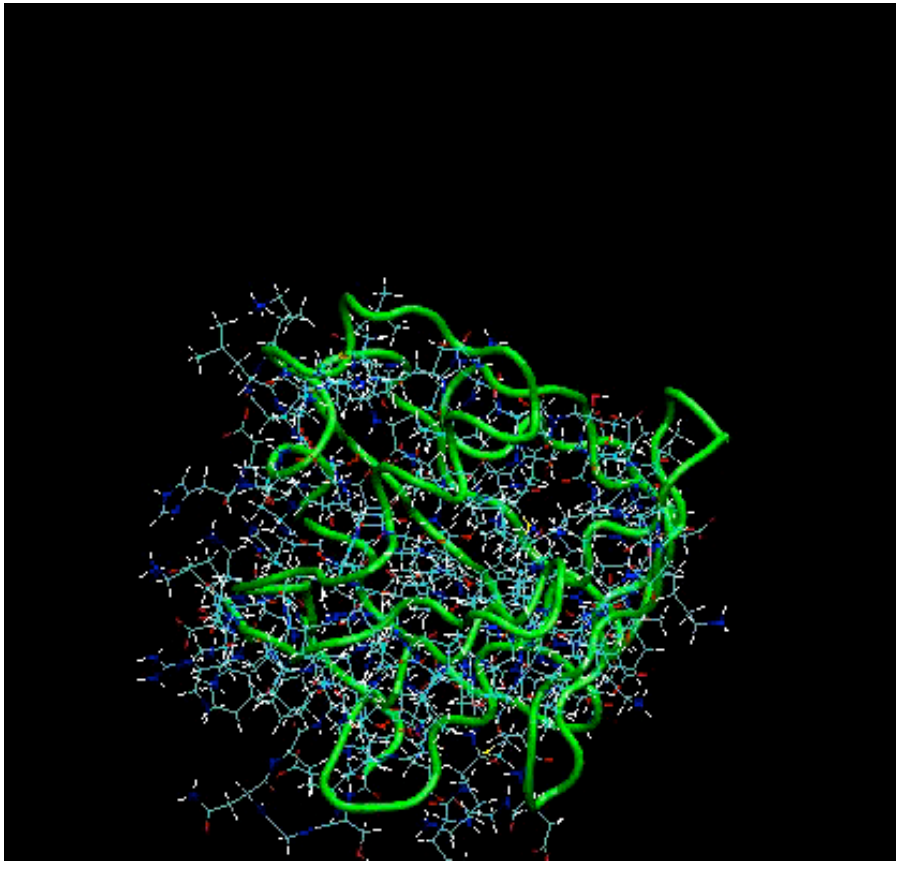
- Not very glamorous...
- ...but (relatively) easy to study.
- Good agreement with explicit solvent methods
- OK agreement with experiment



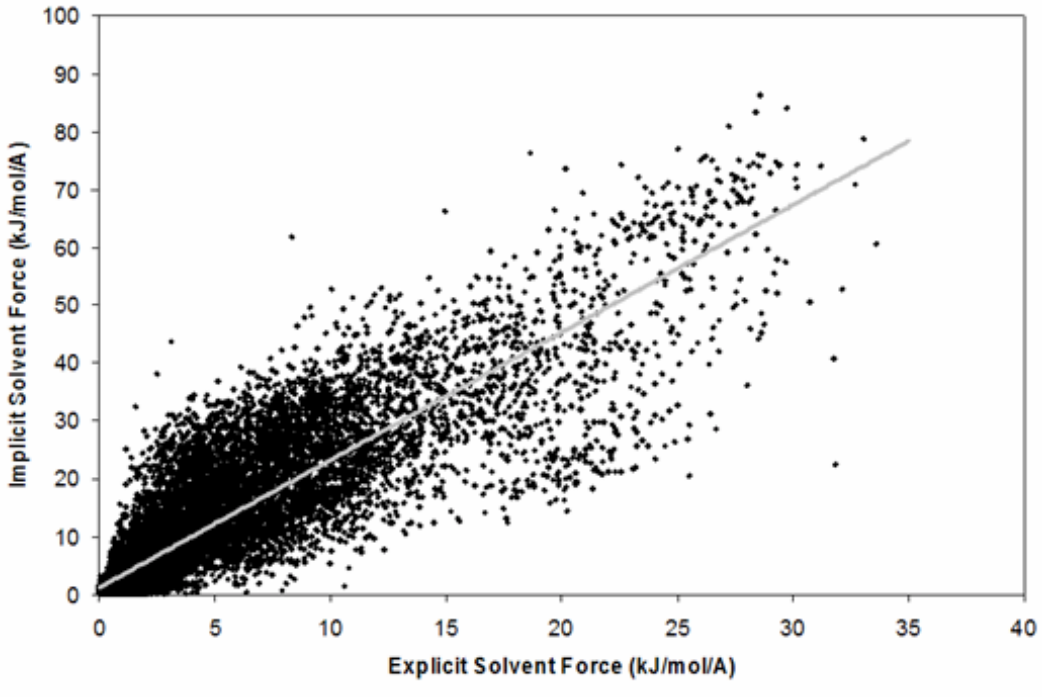
Wagoner JA, Baker NA. *Proc Natl Acad Sci USA*, **103**, 8331-6, 2006.

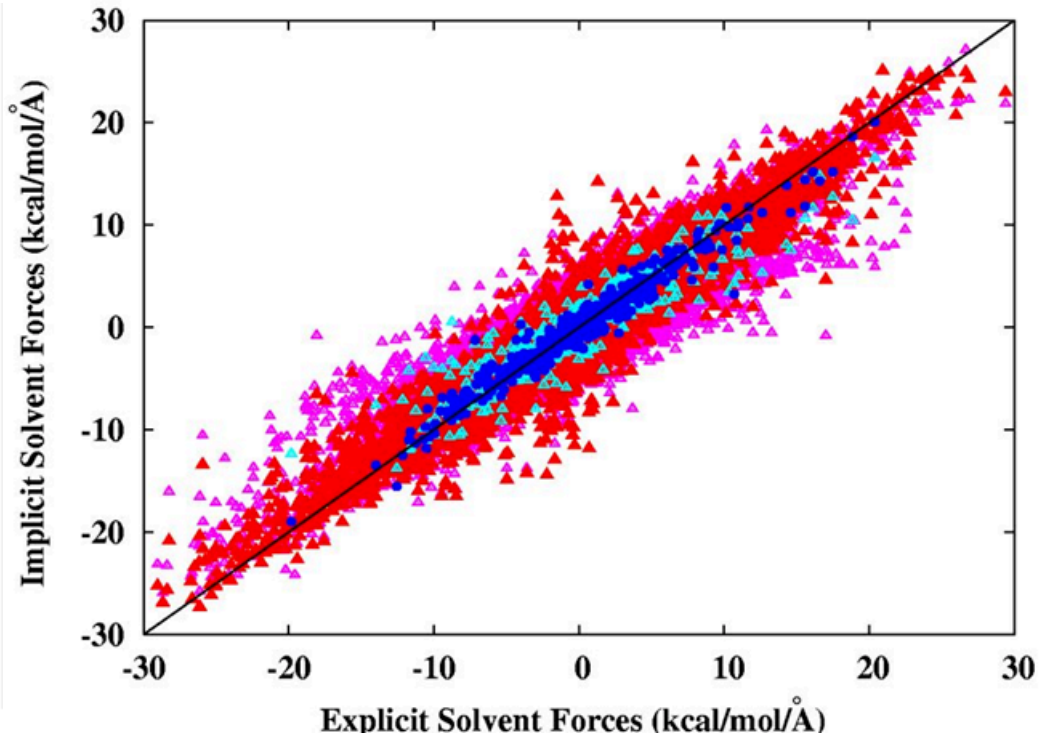
IFABP solvation

- NMR structures (PDB ID 1AEL)
- Highly solvated biomolecule
- Molecular dynamics simulations
  - Conformational clustering
  - Charged protein
  - Uncharged protein
- Solvation forces
  - Polar
  - Nonpolar



IFABP polar solvation

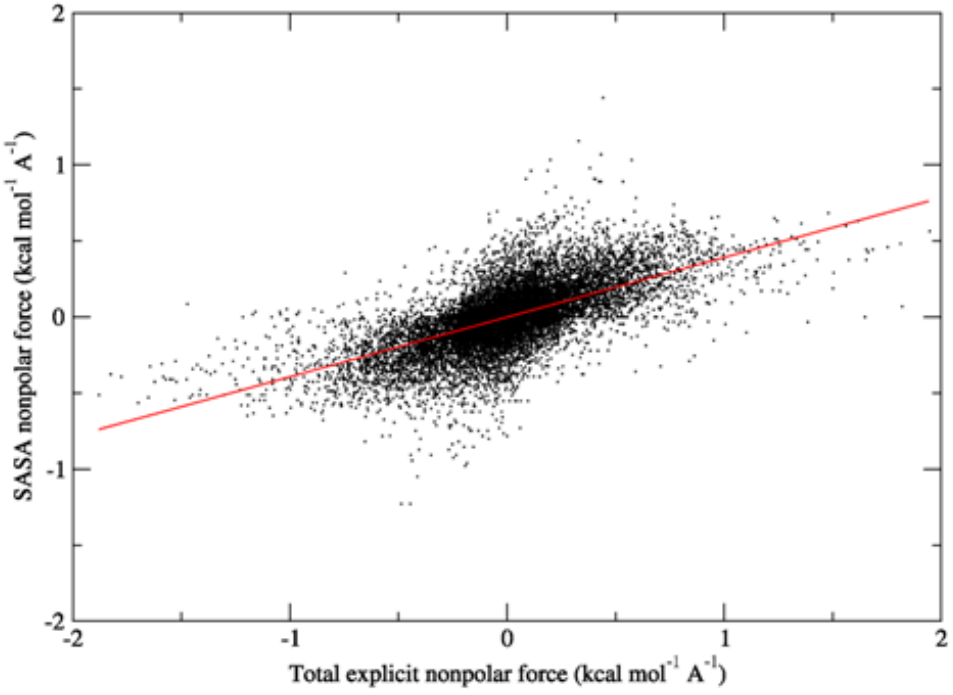


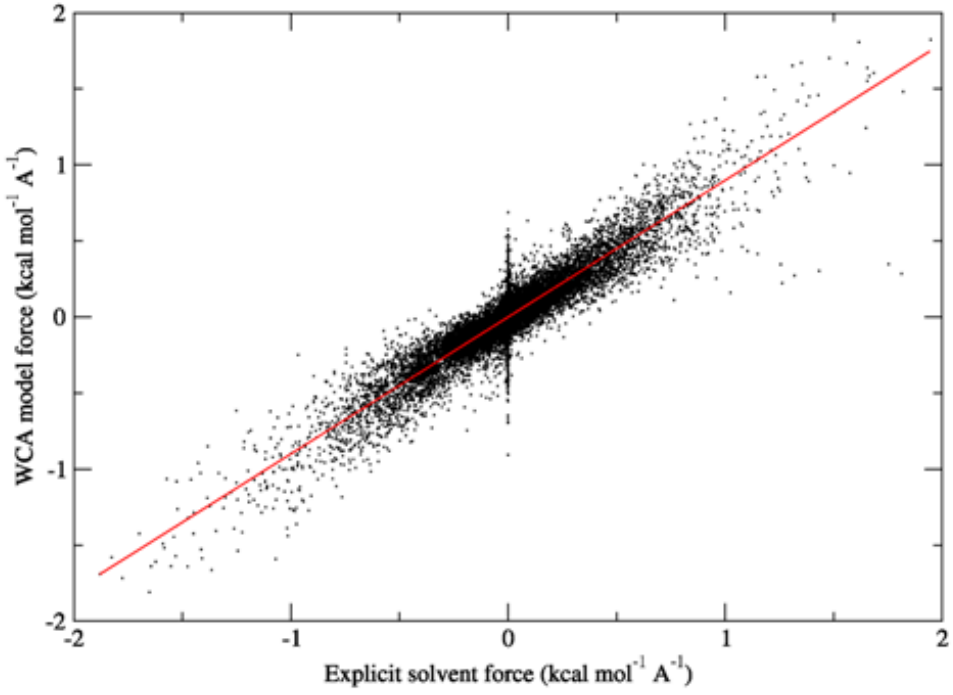


$$\mathbf{F}_i^{(p)}(\mathbf{x}) = -\frac{\partial W^{(p)}(\mathbf{x})}{\partial \mathbf{x}_i} = -\left\langle \frac{\partial U_{uv}^{(p)}}{\partial \mathbf{x}_i} + \frac{\partial U_{uv}^{(np)}}{\partial \mathbf{x}_i} \right\rangle_p + \left\langle \frac{\partial U_{uv}^{(np)}}{\partial \mathbf{x}_i} \right\rangle_{np}$$

Wagoner JA, Baker NA, *J Comput Chem*, **25**, 1623-9, 2004; Swanson JMJ, et al, submitted.

IFABP nonpolar solvation





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

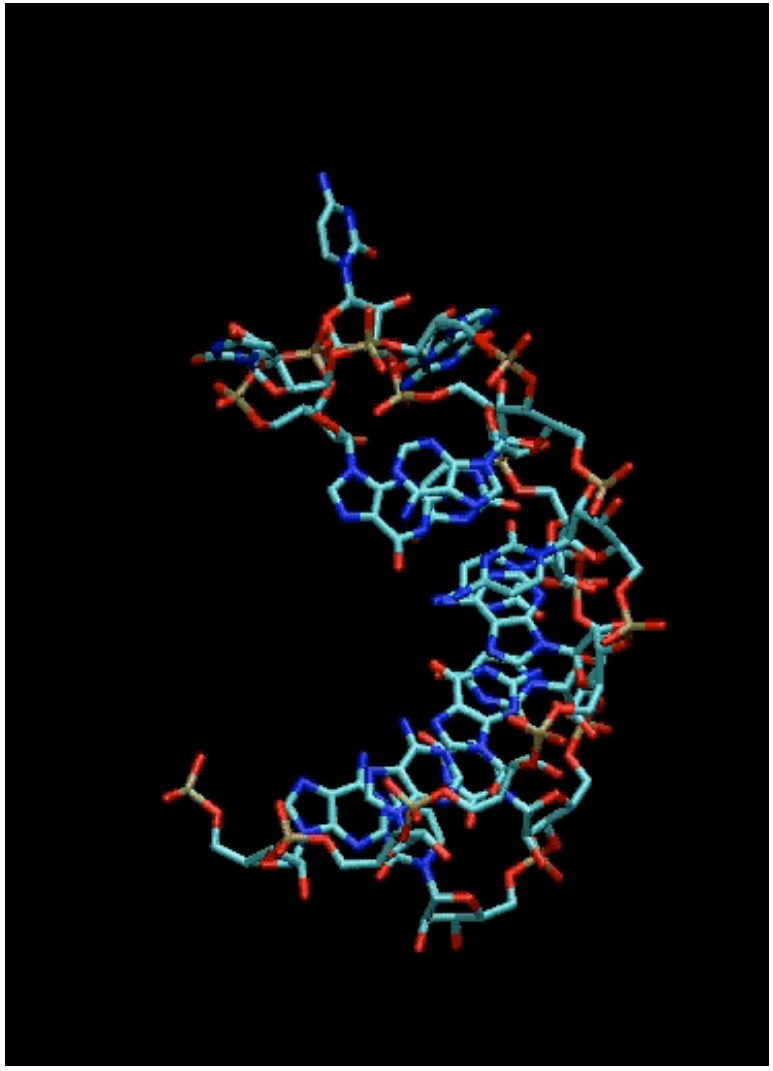
$$\mathbf{F}_i^{(\mathrm{np})}(\mathbf{x}) = -\gamma\frac{\partial A(\mathbf{x};\sigma)}{\partial \mathbf{x}_i} - p\frac{\partial V(\mathbf{x};\sigma)}{\partial \mathbf{x}_i} - \overline{\rho}\int_{\Omega} g_0(\mathbf{x},\mathbf{y};\sigma)\frac{\partial U_{\mathrm{att}}^{(\mathrm{np})}(\mathbf{x},\mathbf{y};\sigma)}{\partial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

WCA parameters: γ = 3 cal mol^{-1} Å^{-2}, p = 36 cal mol^{-1} Å^{-3}, σ = 1.13 Å

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

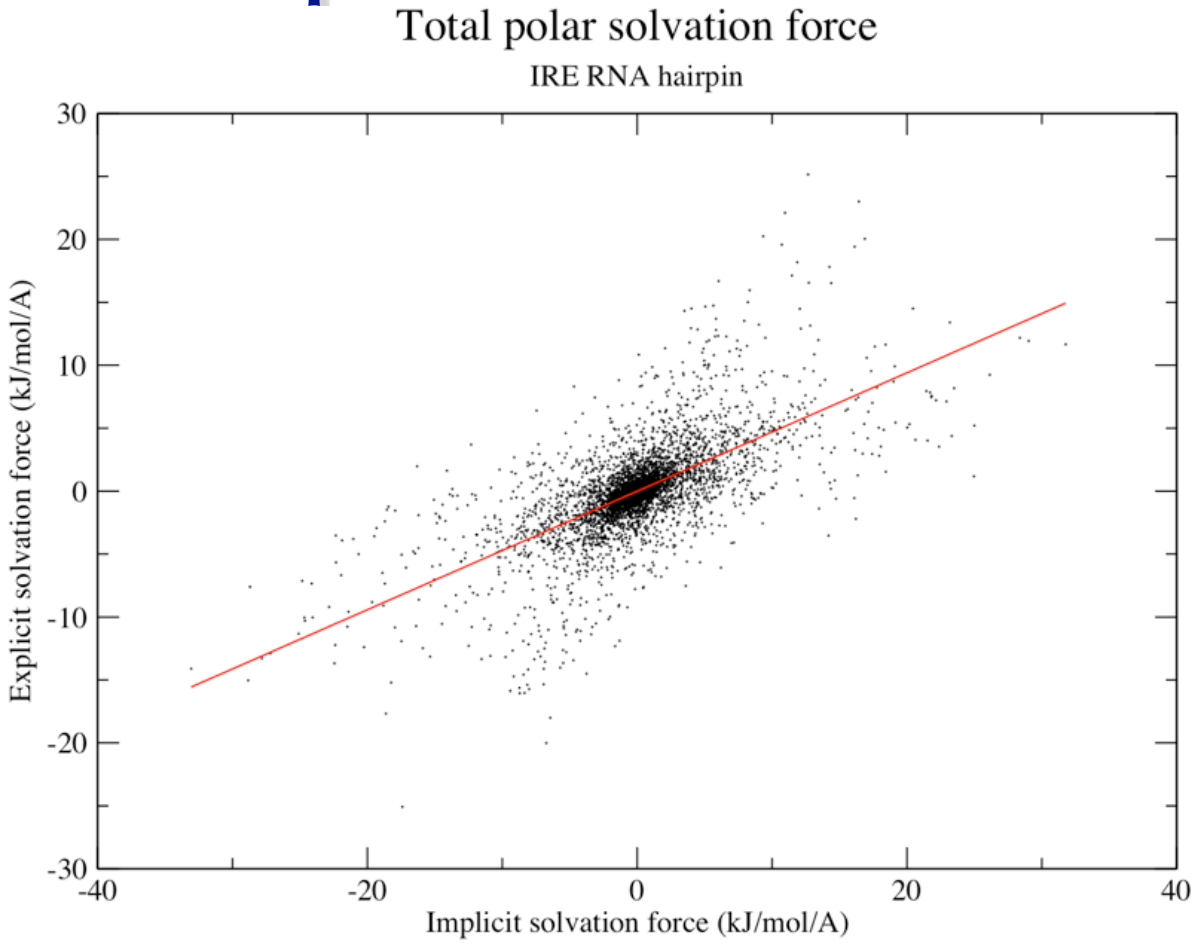
IRE solvation

- IRE hairpin
- Molecular dynamics simulation (AMBER94) agrees well with NMR data
- No specific divalent metal requirement



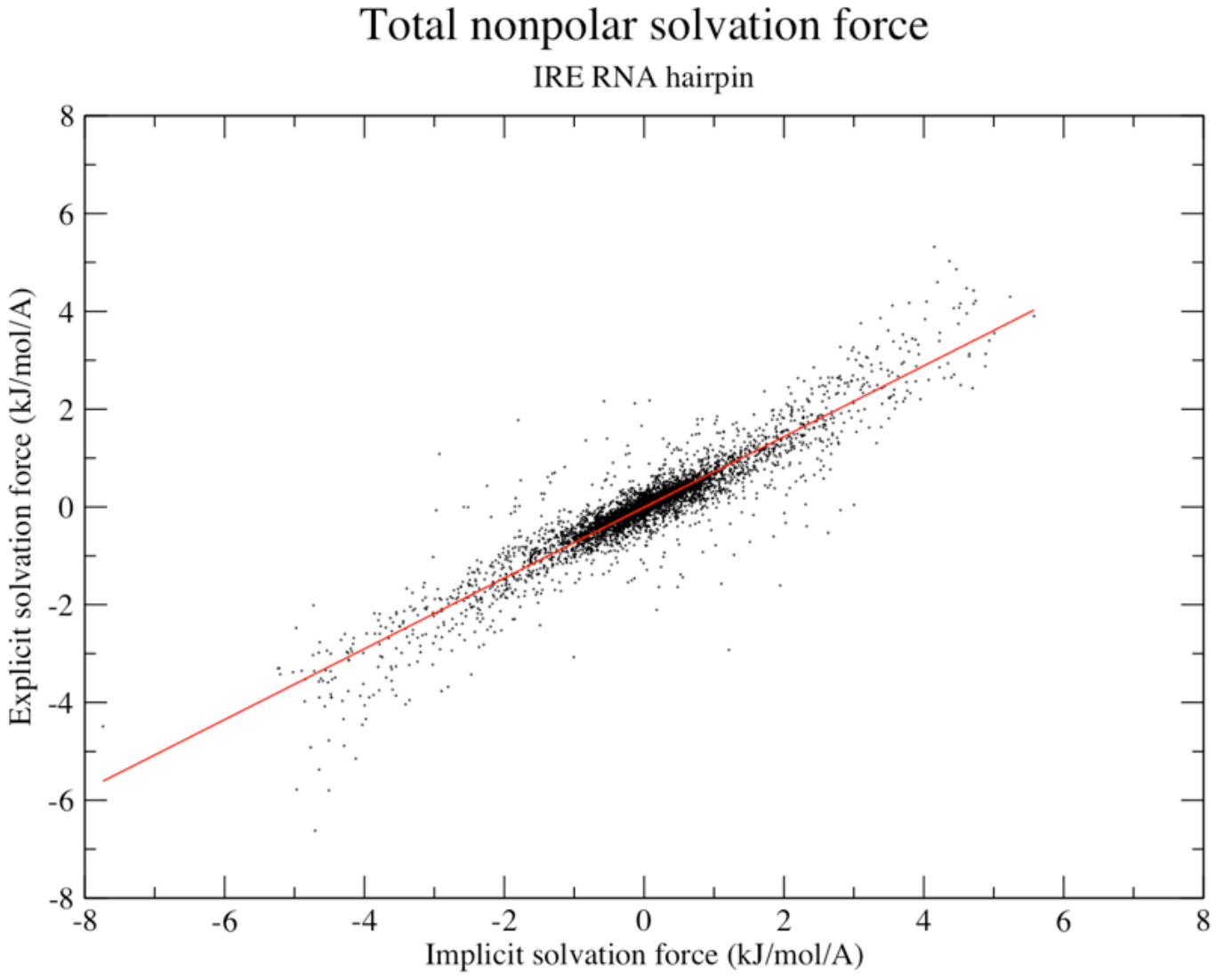
Showalter et al, *J Biomol NMR*, **32**, 179-93, 2004.

IRE polar solvation



$$\mathbf{F}_i^{(p)}(\mathbf{x}) = -\frac{\partial W^{(p)}(\mathbf{x})}{\partial \mathbf{x}_i} = -\left\langle \frac{\partial U_{uv}^{(p)}}{\partial \mathbf{x}_i} + \frac{\partial U_{uv}^{(np)}}{\partial \mathbf{x}_i} \right\rangle_p + \left\langle \frac{\partial U_{uv}^{(np)}}{\partial \mathbf{x}_i} \right\rangle_{np}$$

IRE nonpolar solvation



Same WCA parameters as IFAPB: γ = 3 cal mol^{-1} Å^{-2}, p = 36 cal mol^{-1} Å^{-3}, σ = 1.13 Å

Summary

- Do implicit solvent models “work”?
- Successes
  - Previous applications
  - Protein polar solvation forces
  - Protein and RNA nonpolar solvation forces (new model)
  - Alkane solvation energies
- Failures
  - RNA polar solvation forces
  - RNA and DNA ion densities
  - Specific ion binding

Acknowledgments

- People
  - Todd Dolinsky
  - Feng Dong
  - Dave Gohara
  - Jason Wagoner
- Funding
  - NIGMS
  - NBCR
  - Sloan Foundation
- Collaborators
  - Chandrajit Bajaj
  - Mike Holst
  - Jan Jensen
  - Andy McCammon
  - Jens Nielsen
  - Rohit Pappu
  - Jay Ponder
  - Dave Sept
  - Xiaoyu Zhang