Fast Methods for Simulation of Biomolecule Electrostatics

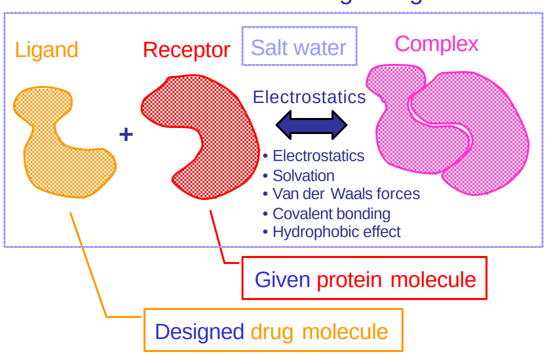
Shihhsien Kuo
Michael Altman
Jaydeep Bardhan
Bruce Tidor
Jacob White

November 12, 2002

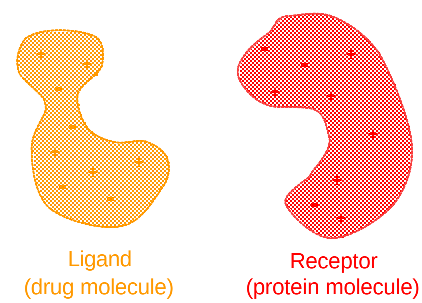
Outline

- Problem statement
- Finite-difference approach and problems
- Integral equation method and advantages
- Fast solver implementation
- Computational results
- Conclusion and Future work

The Problem of Drug Design

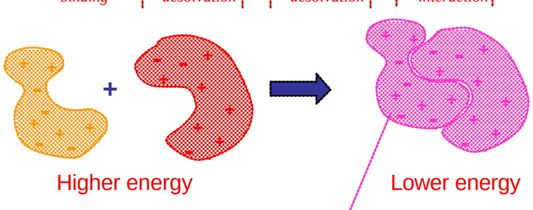


Electrostatics View



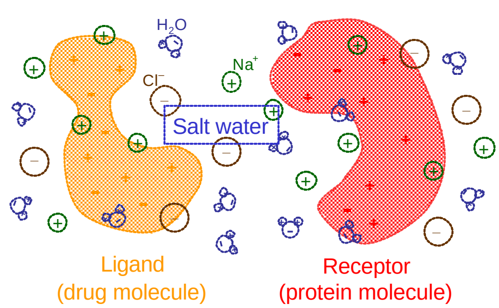
Minimize Electrostatic Binding Energy

$$E_{binding} = \left|E^{ligand}_{desolvation}\right| + \left|E^{receptor}_{desolvation}\right| - \left|E_{interaction}\right|$$

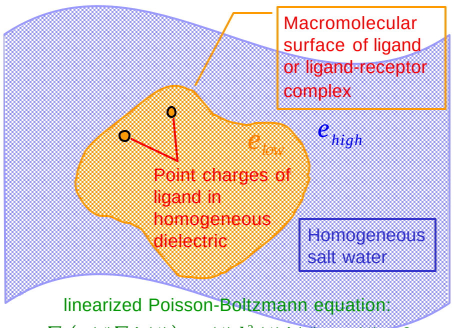


Determine the charge distribution in the ligand so that it is "Energetically Optimized" to bind

An Electrostatic Analysis Problem

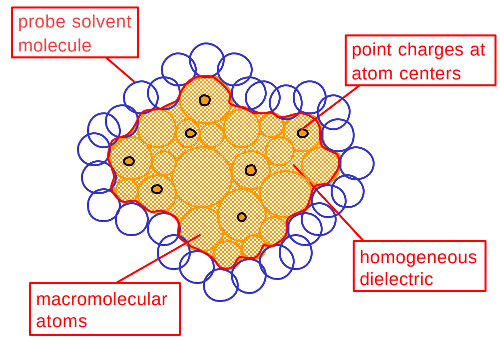


A Simplified Physical Description

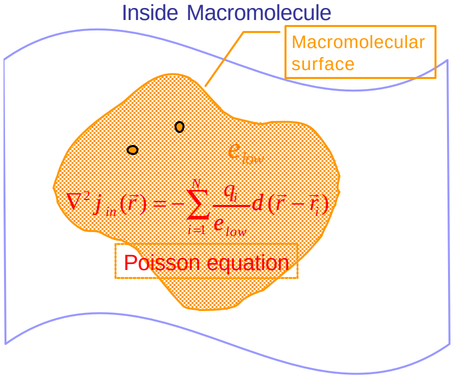


$$\nabla \bullet (e(\vec{r})\nabla j(\vec{r})) - e(\vec{r})k^2(\vec{r})j(\vec{r}) + r(\vec{r}) = 0$$

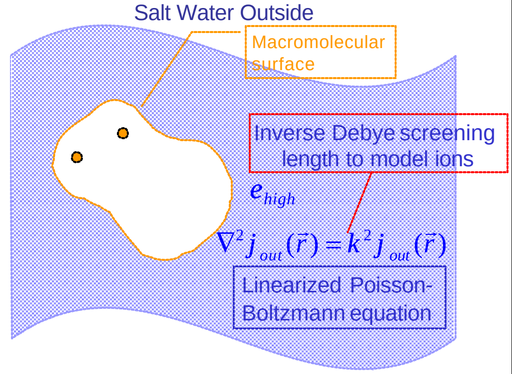
Molecular Surface Representation



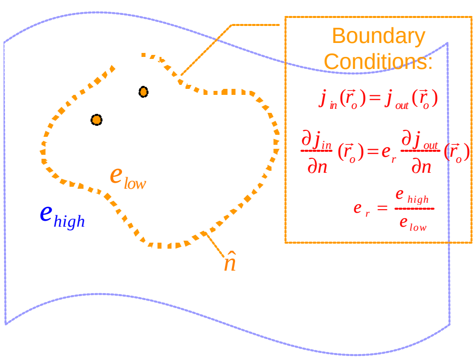
Simplified Mathematical Model: Inside Macromolecule



Simplified Mathematical Model: Salt Water Outside



Interface Condition

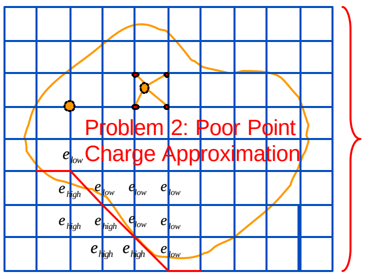


Why Use this Simplified Model?

- Atomistic Level Simulation is too expensive
  - Salt ions and water molecules treated individually
- Continuum Model Matches Well with Experimental Data

Standard Finite-Difference Method

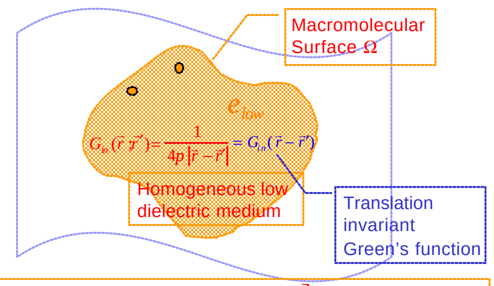
set up boundary conditions and
solve for grid potentials



Problem 3: Inexact Boundary Conditions

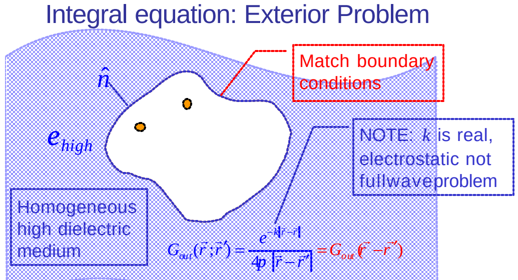
Problem 1: Inaccurate Molecular Surface

Integral equation: Interior Problem



$$\int_\Omega \left[j_{in}(\vec r')\frac{\partial G_{in}}{\partial n}(\vec r;\vec r') - G_{in}(\vec r;\vec r')\frac{\partial j_{in}}{\partial n}(\vec r')\right] d\vec r' = \sum_{k=1}^{N}\frac{q_k}{e_{low}} G_{in}(\vec r;\vec r_k)$$

Integral equation: Exterior Problem



$$\int_\Omega \left[-\underbrace{j_{out}(\vec r')}_{j_{in}(\vec r')}\frac{\partial G_{out}}{\partial n}(\vec r;\vec r') + G_{out}(\vec r;\vec r')\underbrace{\frac{\partial j_{out}}{\partial n}(\vec r')}_{\frac{1}{e_r}\frac{\partial j_{in}}{\partial n}(\vec r')}\right] d\vec r' = 0$$

Advantages For Integral Equation Formulation

- Directly discretize surfaces
- Point charges treated exactly
- Handles infinite exterior

$$\int_\Omega \left[j_{in}(\vec r')\frac{\partial G_{in}}{\partial n}(\vec r;\vec r') - G_{in}(\vec r;\vec r')\frac{\partial j_{in}}{\partial n}(\vec r')\right] d\vec r' = \sum_{k=1}^{N}\frac{q_k}{e_{low}} G_{in}(\vec r;\vec r_k)$$

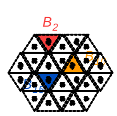
$$\int_\Omega \left[-j_{in}(\vec r')\frac{\partial G_{out}}{\partial n}(\vec r;\vec r') + G_{out}(\vec r;\vec r')\frac{1}{e_r}\frac{\partial j_{in}}{\partial n}(\vec r')\right] d\vec r' = 0$$

Standard piecewise constant collocation discretization method

$$j_{in}(\vec r) \approx \sum_j a_j B_j(\vec r)$$

$$\frac{\partial j_{in}}{\partial n}(\vec r) \approx \sum_j b_j B_j(\vec r)$$

$$\vec r \in \Omega$$



- Piecewise constant basis functions
- Collocation points at panel centroids

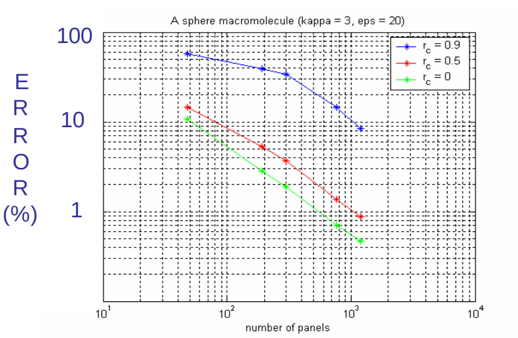
Matrix Equation

$$\begin{bmatrix} D^{in} & S^{in} \\ D^{out} & S^{out} \end{bmatrix}\begin{bmatrix} a_j \\ b_j \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{N}\frac{q_k}{4p|\vec r_i - \vec r_k|} \\ 0 \end{bmatrix}$$

$$D_{ij}^{in} = \int_{panel_j} \frac{\partial}{\partial n'}\left(\frac{1}{4p|\vec r_i - \vec r'|}\right) d\vec r' \qquad S_{ij}^{in} = -\int_{panel_j} \frac{1}{4p|\vec r_i - \vec r'|} d\vec r'$$

$$D_{ij}^{out} = -\int_{panel_j} \frac{\partial}{\partial n'}\left(\frac{e^{-k|\vec r_i - \vec r'|}}{4p|\vec r_i - \vec r'|}\right) d\vec r' \qquad S_{ij}^{out} = \frac{1}{e_r}\int_{panel_j} \frac{e^{-k|\vec r_i - \vec r'|}}{4p|\vec r_i - \vec r'|} d\vec r'$$

A sphere molecule: comparison with analytical result



Iterative solver

$$j(\vec{r}) \equiv \int K(\vec{r};\vec{r}')\, s(\vec{r}')\, d\vec{r}'$$

Discretization

$$\left[K_{i,j}\right]\left[s_j\right] = \left[j_i\right]$$

Gaussian Elimination

$$\left[s_j\right] = \left[K_{i,j}\right]^{-1}\left[j_i\right]$$

```
for i = 1:n-1
  for j = i+1:n
    K_j,i = K_j,i / K_i,i
    for k = i+1:n
      K_j,k = K_j,k - K_j,i / K_i,k
    end
  end
end
```

$O(N^3)$

Iterative Solver

$[s_j]_{guess}$ → matrix-vector multiply black box → $[K_{i,j}][s_j]_{guess}$ → improve guess solution

$O(N^2)$ per iteration

Use Fast Integral Equation Solver

$O(N \log N)$ Matrix-vector multiply

- Multiple Green's functions
- Translation Invariant kernel

$$G_{in}(\vec{r};\vec{r}') = \frac{1}{4p\,|\vec{r}-\vec{r}'|}$$

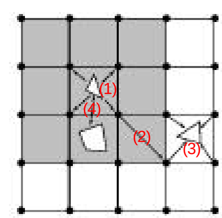
$$G_{out}(\vec{r};\vec{r}') = \frac{e^{-k|\vec{r}-\vec{r}'|}}{4p\,|\vec{r}-\vec{r}'|}$$

Pre-corrected FFT algorithm

$[s_j]_{guess}$ — matrix-vector multiply black box — $[K_{i,j}][s_j]_{guess}$

charge distribution — pre-corrected FFT — potential due to space-invariant kernel

1) projection of panel charges onto grid charges
2) grid potentials due to grid charges are computed by FFT
3) potentials on panel centroids are interpolated from grid potentials
4) direct interaction and correction among near neighbors



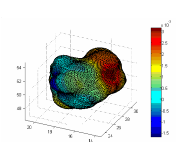
Picture courtesy of J. Phillips

Preconditioner on Two Examples

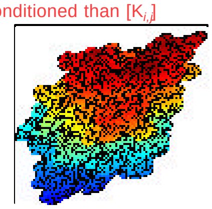
Need to find a good preconditioner $[P] \approx [K_{i,j}]$

And solve $[P]^{-1}[K_{i,j}][s_j] = [P]^{-1}[j_i]$

hopefully better conditioned than $[K_{i,j}]$

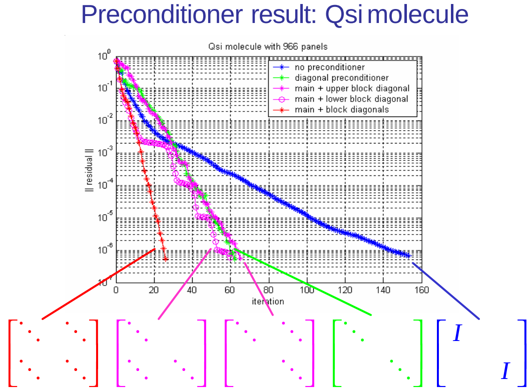


Qsi molecule

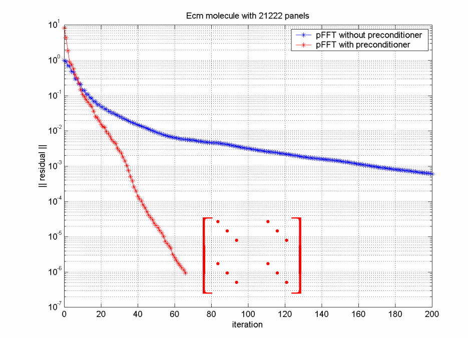


Ecm protein

Preconditioner result: Qsi molecule



Preconditioner result: Ecm protein

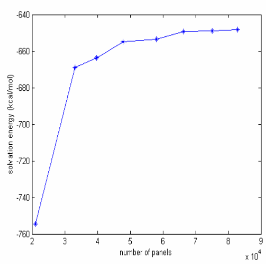


Accuracy comparison with DelPhi

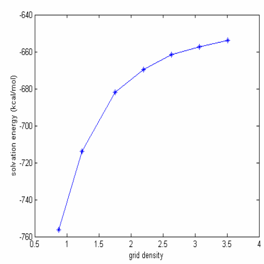
			$E_{solvation}$ (kcal/mol)	
	# of dielectric panels	# of salt panels	pFFT	DelPhi
Water	17204	9330	-3.14	-3.17
TSA	34114	5842	-34.62	-34.75
ECM	82868	18596	-646.42	-653.88

Convergence Results of Ecm Protein

pFFT



DelPhi



Binding energy calculation of a protein-peptide complex

	Energy calculated (kcal/mol)			
	$R_{desolvation}$	$L_{desolvation}$	$(R\text{->}L)_{interaction}$	$(L\text{->}R)_{interaction}$
pFFT	14.52	24.47	130.80	130.91
DelPhi	14.51	24.47	131.03	131.03

Conclusions and Future work

- Carefully selected Integral Formulation results in Fast Solver for Biomolecule Electrostatics
- Working on coupling to charge optimization problem in drug design
- Extending formulation to include more complicated geometry (inner cavities in macromolecule)
- Fine tuning existing pre-corrected FFT code