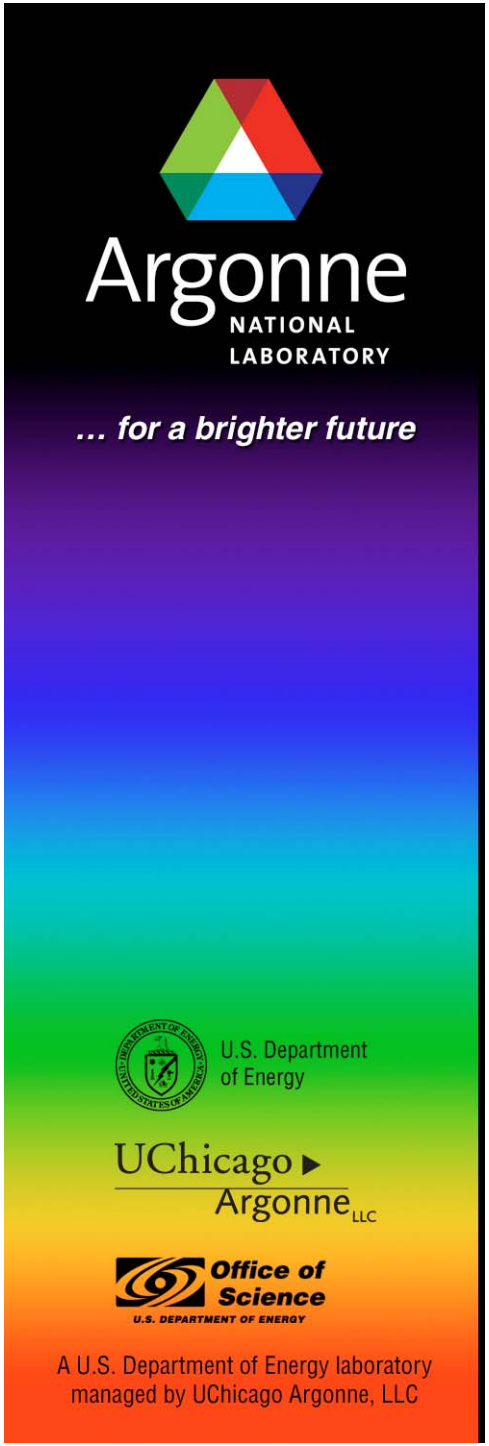


Electrostatic Interactions Between Biomolecules: A PDE-Constrained Approach

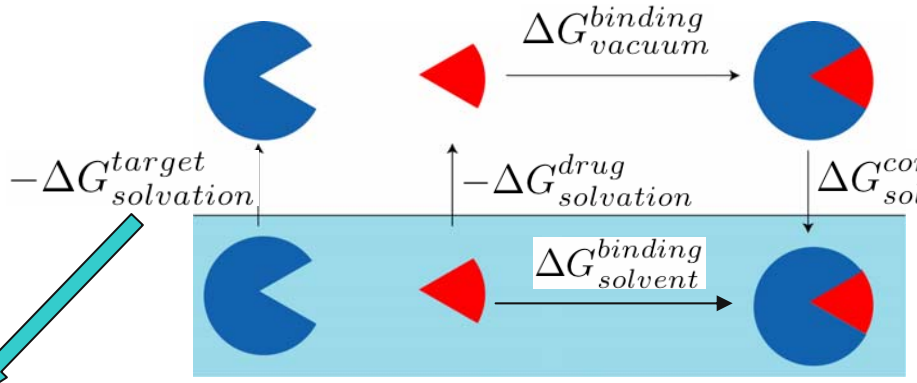
Jay Bardhan

Mathematics and Computer Science Division

Argonne National Laboratory

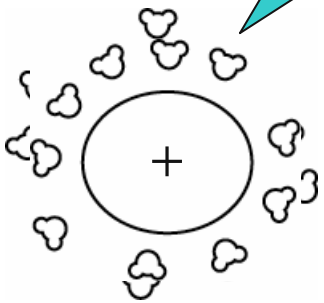
June 19, 2007

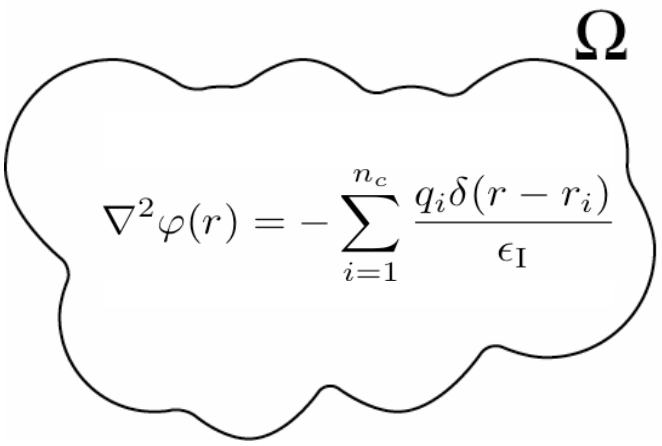
Water strongly mediates biomolecule interactions!



(A rigid binding model is shown here, but it is not essential)

$$\Delta G^0 = \Delta G^0_{\text{non es}} + \Delta G^0_{\text{es}}$$





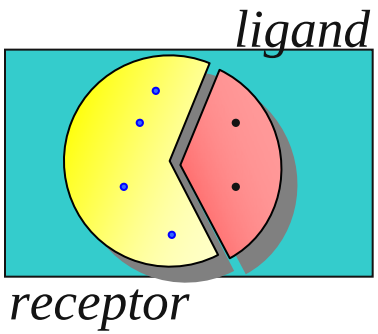
$$\nabla^2\varphi(r) = \kappa^2\varphi(r)$$

$$\nabla^2\varphi(r) = 0$$

Outline:

- Electrostatic optimization as a tool for molecular analysis and design
  - Mathematical model and assumptions
  - Applications of the theory
- Numerical methods for simulating electrostatics
  - Integral equation formulations of the problem
  - The boundary-element method
- PDE-constrained optimization
  - A spectrum of approaches exist; not all are equal
  - Subtle implementation details
- (Epilogue) A report from one year out of the CSGF program
  - How can we grow the interdisciplinary community?
  - What responsibilities do we have?

The Objective Function: The Ligand-Dependent Component of the Electrostatic Binding Free Energy



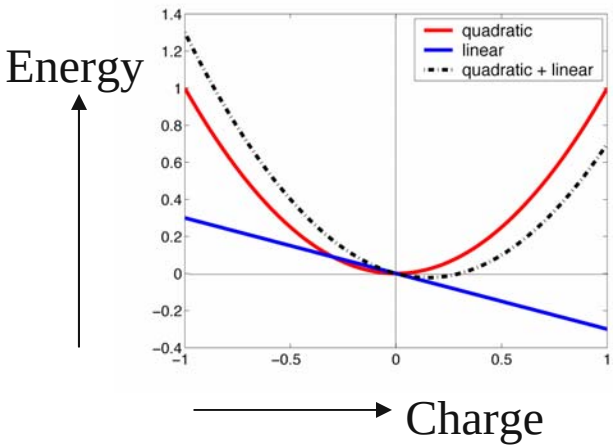
$$\Delta G^0_{\text{var,es}} = \frac{1}{2} q_L^T \varphi^{bound} - \frac{1}{2} q_L^T \varphi^{unbound}$$

Using some sort of numerical method:

$$A_u \sigma = B_u q_L$$

$$\varphi^{unbound} = C_u A_u^{-1} B_u q_L$$

- Proven: This energy function is convex!
- The idea: It always *costs energy* to remove the water from the receptor volume!

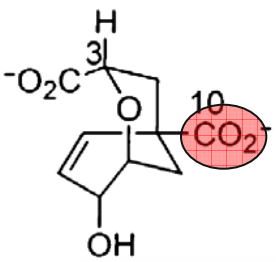


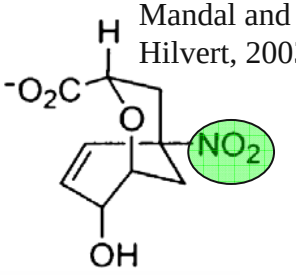
Electrostatic Optimization of Biomolecules:
Applications in Analysis and Design

- *E. coli* chorismate mutase inhibitors:
  - Analyzed by Kangas and Tidor
  - Suggested substitution experimentally verified: result is the tightest-binding inhibitor yet known

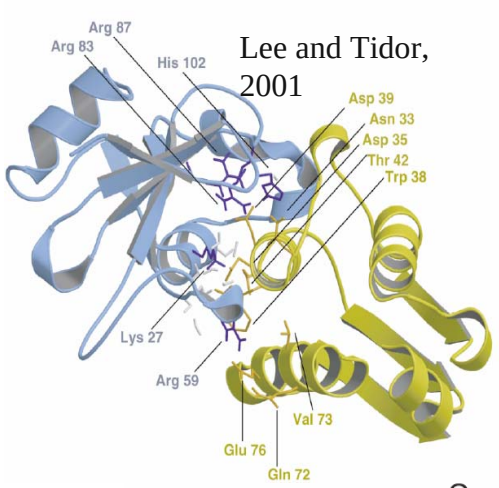
- Barnase/barstar protein complex:
  - Tight-binding complex
  - Optimal charge distribution closely matches "wild-type" charge distribution
  - Mutations do not improve binding

- Glutaminyl-tRNA synthetase:
  - Recognition of proper amino acid is critical for error-free protein synthesis

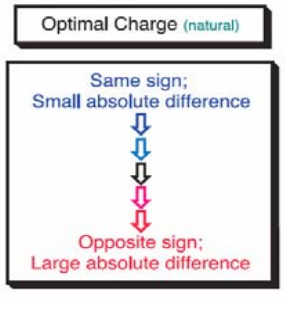




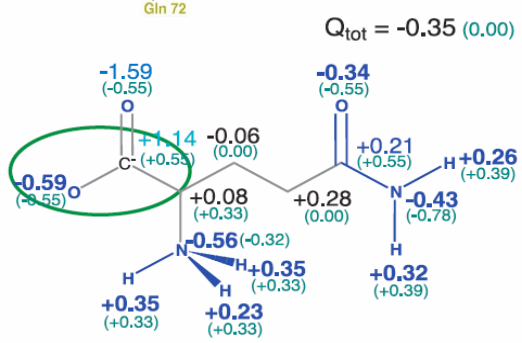
Mandal and Hilvert, 2003



Lee and Tidor, 2001



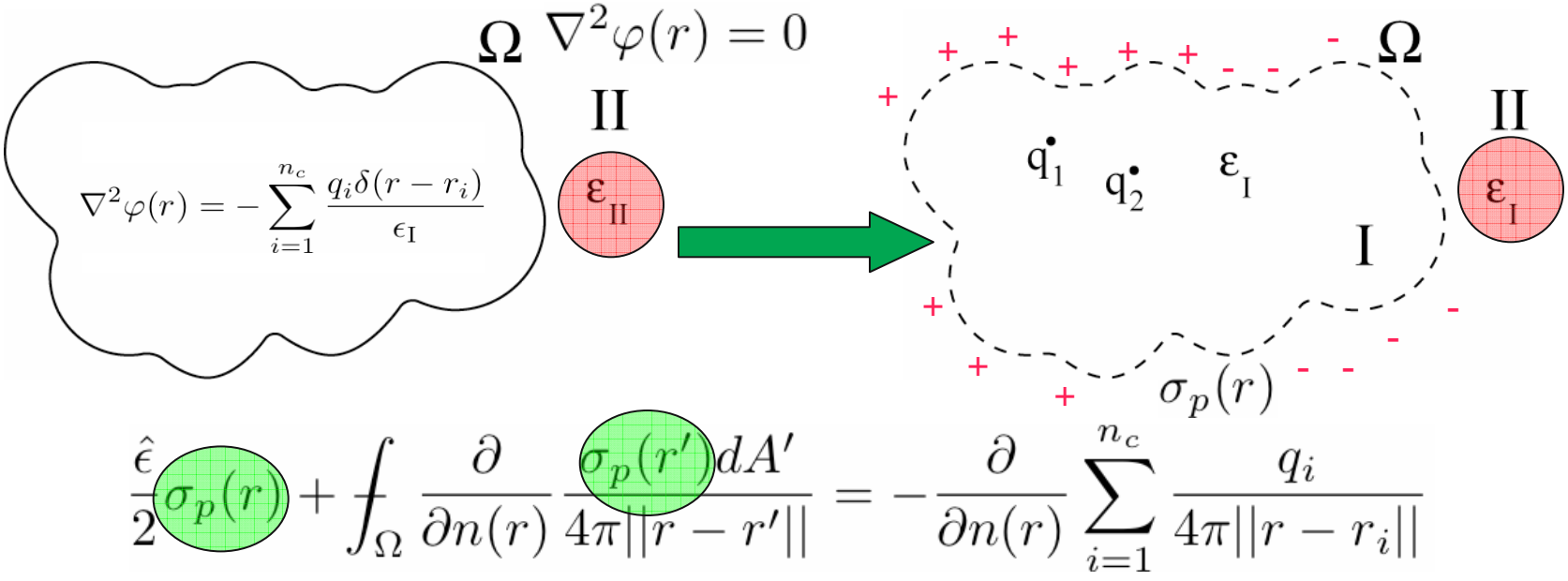
Green and Tidor, 2004



Outline:

- ✓ Electrostatic optimization as a tool for molecular analysis and design
  - Mathematical model and assumptions
  - Applications of the theory
- ❑ Numerical methods for simulating electrostatics
  - Integral equation formulations of the problem
  - The boundary-element method
- ❑ PDE-constrained optimization
  - A spectrum of approaches; not all are equal
  - Subtle implementation details
- ❑ (Epilogue) A report from one year out of the CSGF program
  - How can we grow the interdisciplinary community?
  - What responsibilities do we have?

Numerical Methods for Electrostatic Simulation



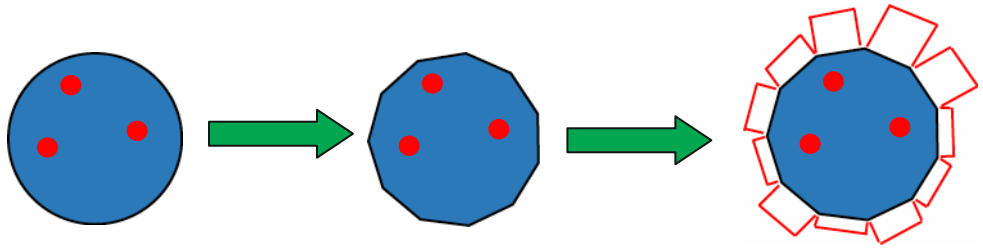
Another formulation:

$$\frac{1}{2}\varphi_I(r_\Omega) + \int_\Omega dr' \varphi_I(r') \frac{\partial G_I}{\partial n}(r_\Omega; r') - \int_\Omega dr' \frac{\partial \varphi_I}{\partial n}(r') G_I(r_\Omega; r') = \sum_{i=1}^{n_c} \frac{q_i}{\epsilon_I} G_I(r_\Omega; r_i)$$

$$\frac{1}{2}\varphi_I(r_\Omega) - \int_\Omega dr' \varphi_I(r') \frac{\partial G_{II}}{\partial n}(r_\Omega; r') + \frac{\epsilon_I}{\epsilon_{II}} \int_\Omega dr' \frac{\partial \varphi_I}{\partial n}(r') G_{II}(r_\Omega; r') = 0$$

- Integral equation methods have some advantages for these problems:
  - Boundary conditions at infinity are treated exactly (not truncated)
  - Discrete point charges are treated exactly (not spread to grid)
  - The dielectric boundary can be treated exactly (hard work not described here!)

Numerical Methods for Electrostatic Simulation: The Boundary-Element Method



piecewise constant basis functions

$$\int_\Omega K(r;r')\sigma_p(r')dA = f(r) \qquad \hat{\sigma}_p(r) = \sum_{panels} x_i b_i(r)$$

$$R(r) = A\sigma_p = Bq_L r')\sigma_p(r')$$
$$\varphi_R = C\sigma_p$$

- Krylov-subspace iterative methods are used to solve the linear systems:

$$\hat{x} \in span\{b, Ab, \ldots, A^{i-1}b\}$$

- Preconditioning is commonly employed to reduce # of required iterations:

$$PAx = Pb$$

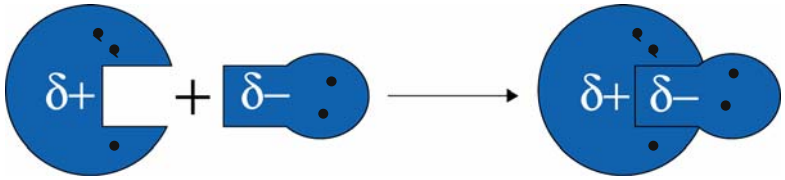
Outline:

- ✓ Electrostatic optimization as a tool for molecular analysis and design
  - Mathematical model and assumptions
  - Applications of the theory
- ✓ Numerical methods for simulating electrostatics
  - Integral equation formulations of the problem
  - The boundary-element method
- ❑ PDE-constrained optimization
  - A spectrum of approaches; not all are equal
  - Subtle implementation details
- ❑ (Epilogue) A report from one year out of the CSGF program
  - How can we grow the interdisciplinary community?
  - What responsibilities do we have?

PDE-Constrained Optimization:

- The electrostatic binding free energy is determined by solving two PDEs:

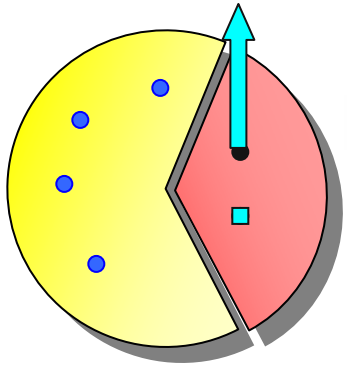
Receptor and ligand are infinitely far apart; energy due to ligand charges is purely quadratic



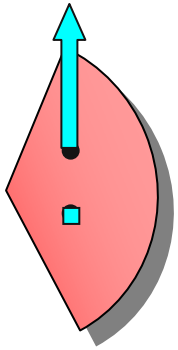
Ligand charges produce another (different!) quadratic reaction energy *and* interact directly with the receptor charges

$$E_{unbound} = q^T L_u q \qquad E_{bound} = q^T L_b q + q^T c$$

- One approach: find explicit representation of the objective function (invert the PDE constraints)



$$q_i = 1$$
$$q_j = 0 \ \text{if} j \neq i$$



$$A_b q_{diel.b} = B_b q_L \qquad A_u q_{diel,u} = B_u q_L$$
$$\varphi_b = C_b q_{diel,b} \qquad \varphi_u = C_u q_{diel,u}$$

PDE-Constrained Optimization: Implicit Hessian Methods

$$\text{minimize} \frac{1}{2} q^T (L_b - L_u) q + c^T q$$

- Alternatives:
  - "Nested analysis and design" methods
    - PDE solver and optimization solver are "black boxes"
    - Every gradient evaluation requires two PDE solves:

$$\hat{q}^i \in \text{span}\{-c, -(L_b - L_u)c, \ldots, -(L_b - L_u)^{i-1}c\}$$

  - "Simultaneous analysis and design" methods
    - PDE solver black box is broken; optimization is still a black box
    - PDE state variables are introduced as optimization variables:

$$\text{minimize} \quad \frac{1}{2} \begin{bmatrix} q \\ y_b \\ y_u \end{bmatrix}^T \begin{bmatrix} 0 & \frac{1}{2}M_{3,b} & -\frac{1}{2}M_{3,u} \\ \frac{1}{2}M_{3,b}^T & & \\ -\frac{1}{2}M_{3,u}^T & & \end{bmatrix} \begin{bmatrix} q \\ y_b \\ y_u \end{bmatrix} + \begin{bmatrix} c \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} q \\ y_b \\ y_u \end{bmatrix} \quad \text{subj. to} \quad \begin{bmatrix} -M_{1,b} & M_{2,b} & \\ -M_{1,u} & & M_{2,u} \end{bmatrix} \begin{bmatrix} q \\ y_b \\ y_u \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- Our "reverse-Schur" method breaks both black boxes

The "Reverse-Schur" PDE-Constrained Method

- The problem:

$$\begin{aligned} \text{minimize} \quad & q^T (L_{bound} - L_{unbound}) q + c^T q \\ \text{subject to} \quad & A_c q = b \\ \text{and} \quad & m_i \leq q_i \leq M_i, \forall i \in \{1, \ldots, n_c\}, \end{aligned}$$

- The "reverse-Schur" approach:

$$CA^{-1}Bx = b \longrightarrow \begin{bmatrix} & C \\ B & -A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

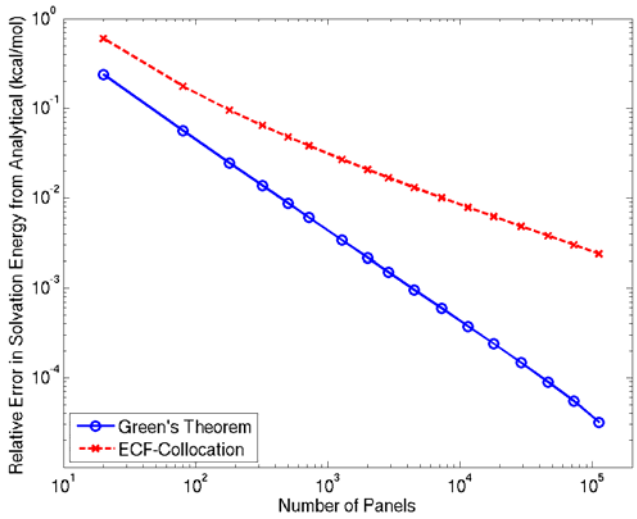
- The unconstrained optimization problem is therefore:

$$\begin{bmatrix} & C_b & -C_u \\ B_b & -A_b & \\ B_u & & -A_u \end{bmatrix} \begin{bmatrix} q_L \\ q_{diel,b} \\ q_{diel,u} \end{bmatrix} = \begin{bmatrix} -c \\ 0 \\ 0 \end{bmatrix}$$

- The constrained problems (linear equality constraints, bound constraints on charges) can be solved using this approach as well!

Some Implementation Issues

1. Why do some integral equation methods converge more slowly as we use more basis functions?

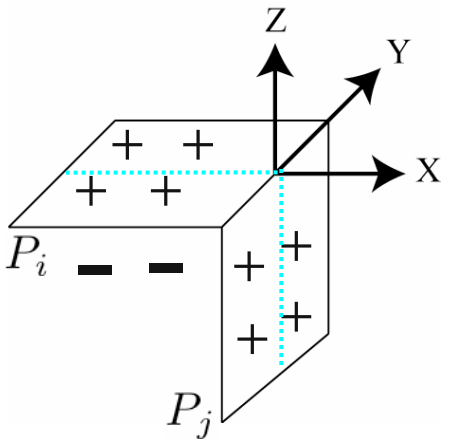


2. The “explicit Hessian” approach, although slow, lets scientists regularize the problem carefully
   - *Can implicit methods allow this?*
   - We will need some cheaply computed *approximate* Hessian
   - Choice of PDE formulation has a strong impact on our ability to obtain one!

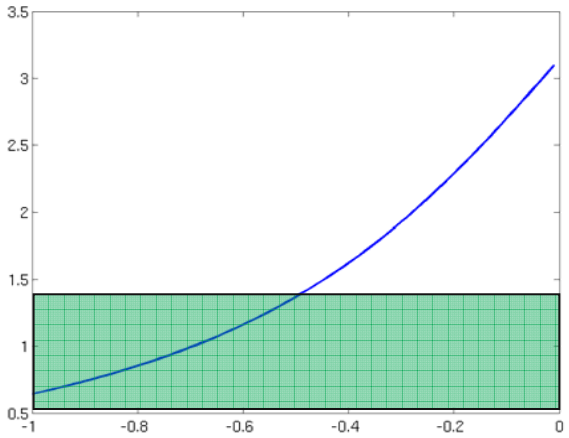
One Implementation Issue: Discretization of the Integral Equations

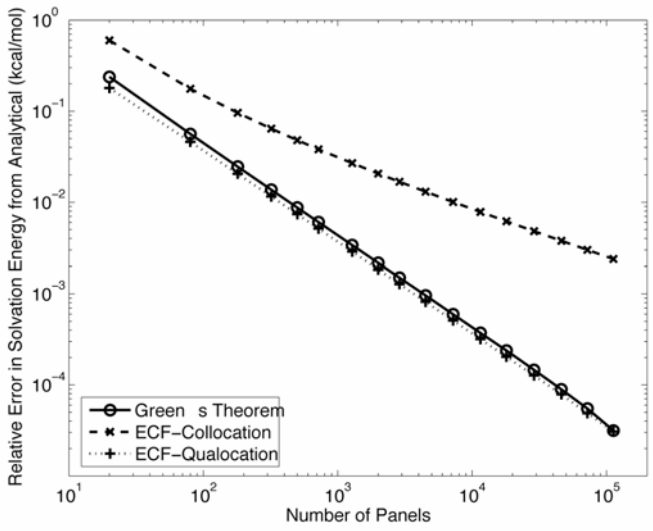
- The double integrals must be evaluated carefully!

$$A_{ij} = \int_{P_j} \left[\int_{P_i} \frac{\partial}{\partial n(r)} G(r;r')dA \right] dA'$$



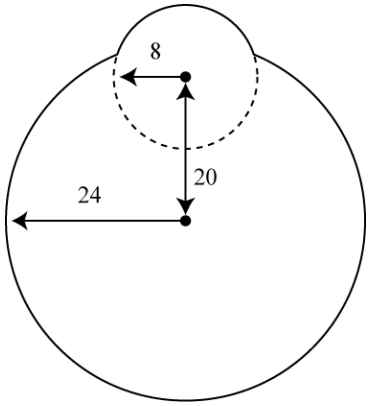
$$\int_{P_i} \frac{\partial}{\partial n(r)} G(r;r')dA$$





Implementation Issue Two: Regularizing the Optimization Problem

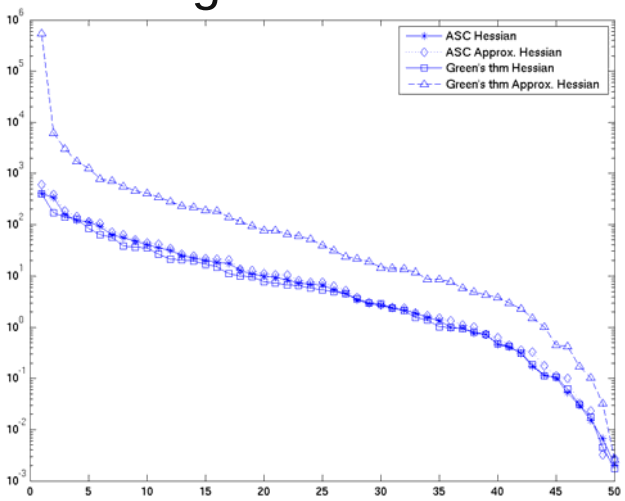
- Generate Hessian approximations for a model geometry:



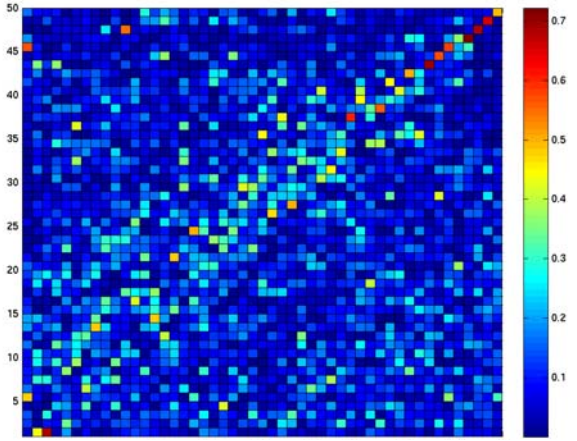
Ligand: 50 charges separated by at least 1.5 A from the sphere and each other

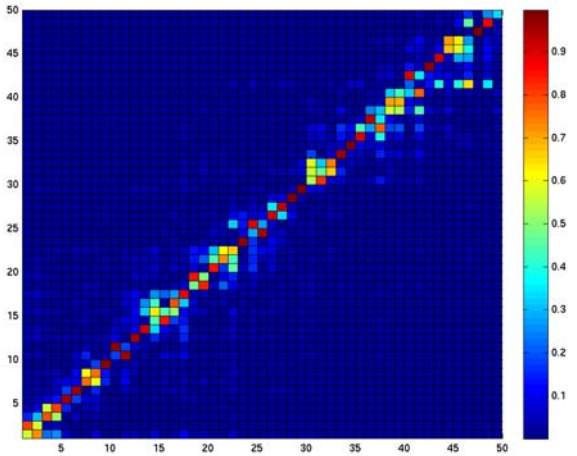
Receptor: 2000 ligand charges separated by at least 1.5 A from the sphere and each other

Eigenvalues



Eigenvectors

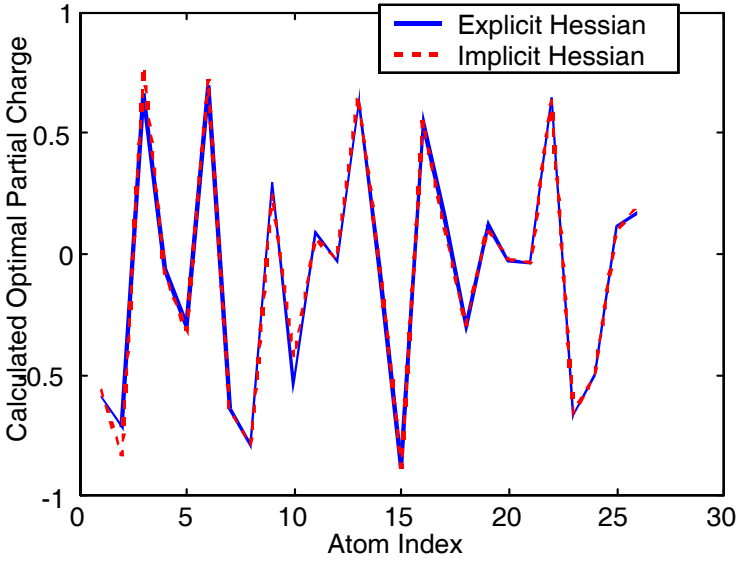




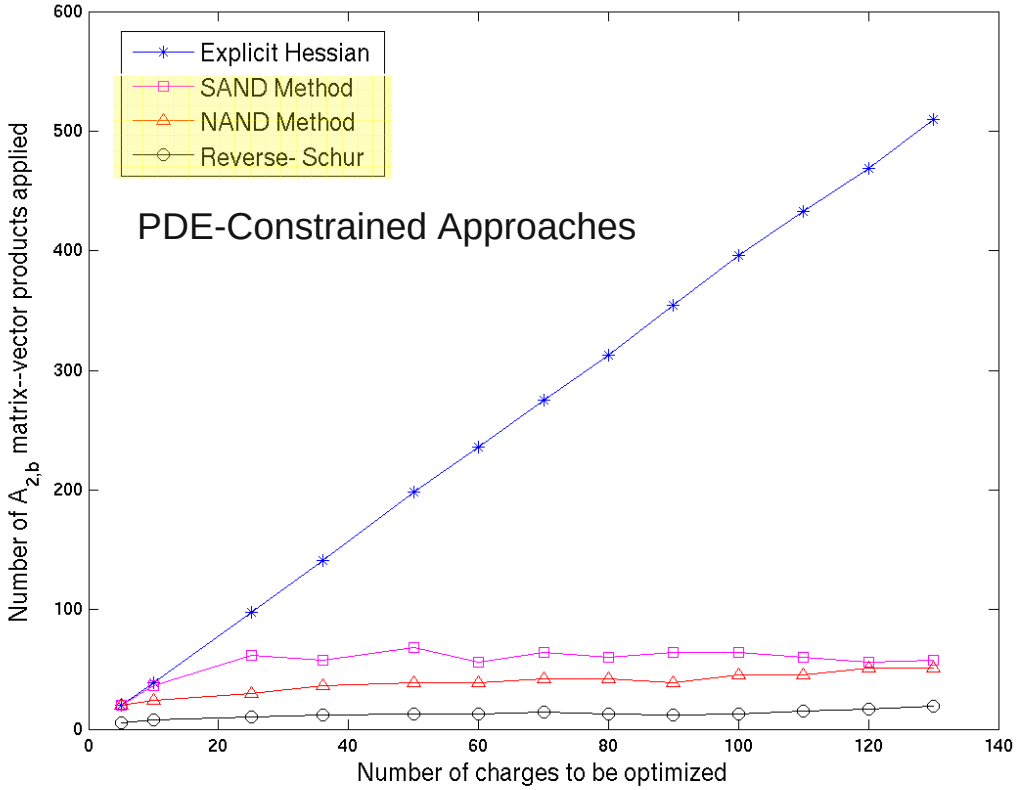
Performance of the "Reverse-Schur" Approach

- Implementations using PETSc, precorrected-FFT; FFTSVD

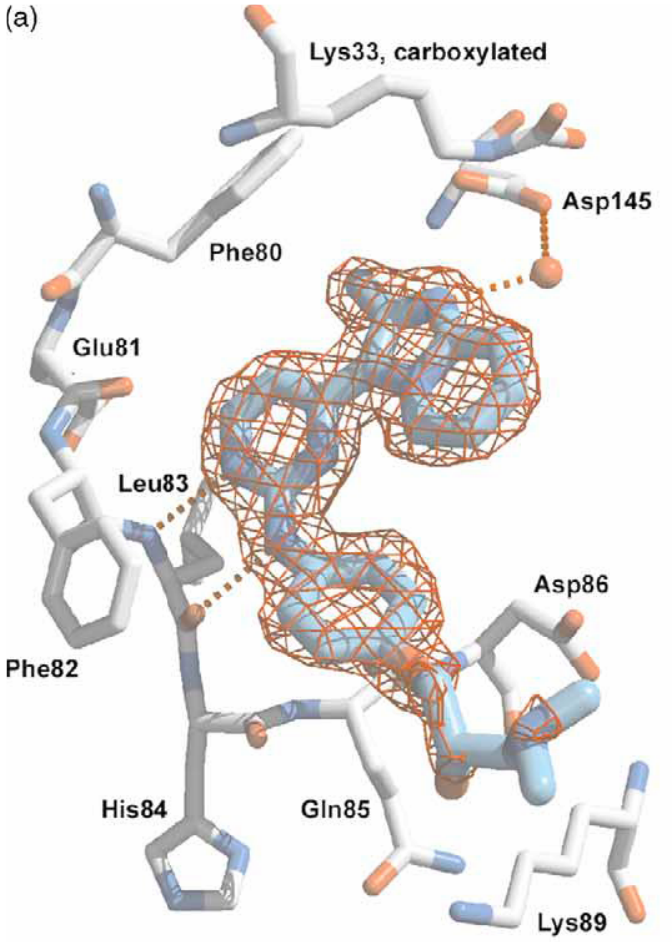
Computed charges agree closely

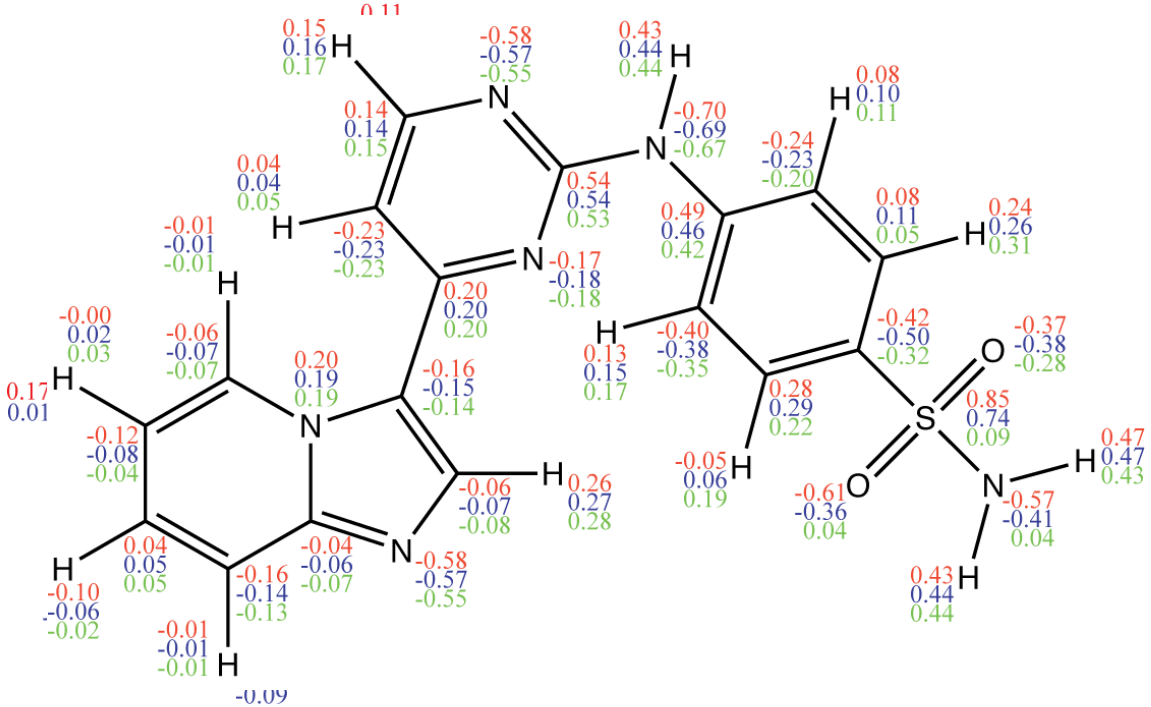


PDE-constrained approaches scale better



Application I: An Inhibitor of Cyclin-Dependent Kinase 2





Anderson, et al. 2003

Outline:

- ✓ Electrostatic optimization as a tool for molecular analysis and design
  - Mathematical model and assumptions
  - Applications of the theory
- ✓ Numerical methods for simulating electrostatics
  - Integral equation formulations of the problem
  - The boundary-element method
- ✓ PDE-constrained optimization
  - A spectrum of approaches; not all are equal
  - Subtle implementation details
- ❑ (Epilogue) A report from one year out of the CSGF program
  - How can we grow the interdisciplinary community?
  - What responsibilities do we have?

Strategies for Making "Interdisciplinary" Happen

1. Enable specialists to interact more effectively with scientists from other disciplines
   - "When in Rome…"
   - Stay focused on the big *shared* picture
   - Traditional fields' balances of "pragmatic" vs. "idealistic" will need to adapt!
     - Analyzing others' work
     - Assembling multidisciplinary teams will require compromises

2. Promote interdisciplinary pedagogy
   - The basics show up time and time again
   - Where should specialized curricula diverge? Should they diverge at all?
   - Two key lessons: humility and patience

Acknowledgements:

- Krell Institute for administering the CSGF program
- DOE for its continued support
- ANL Wilkinson Fellowship in the Mathematics and Computer Science Division
- Collaborators have been funded by: the National Institutes of Health, the Singapore-MIT Alliance, and the MARCO Interconnect Focus Center
- Fantastic advisors (Jacob K. White, Bruce Tidor)