Parallelized Successive Over Relaxation (SOR) Method and Its Implementation to Solve the Poisson-Boltzmann (PB) Equation

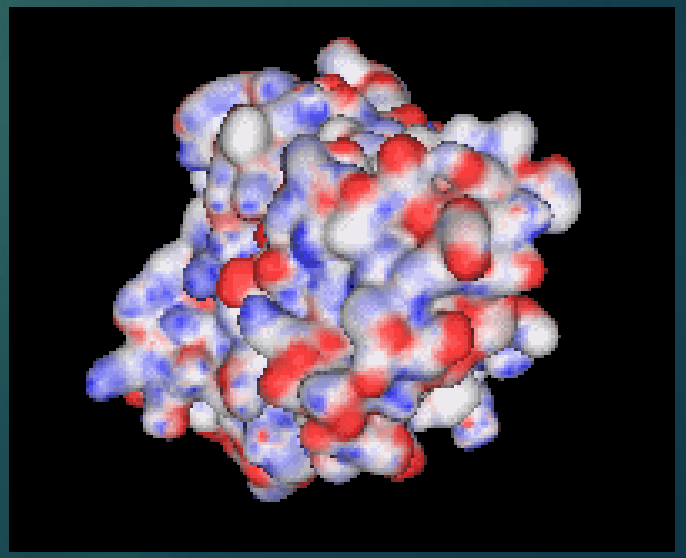
XIAOJUAN YU & DR. CHUAN LI

EPaDel Spring 2017 Section Meeting
Kutztown University
April 1, 2017

Importance of Electrostatics

- Electrostatic interactions are among the most important factors to be considered when analyzing the function of biological molecules

- The ONLY long-range force to provide guidance of long distance



Mathematical Model: The PB Equation

$\nabla \cdot [\varepsilon(x)\nabla\Phi(x)] - k(x)^2 \sinh(\Phi(x)) = -4\pi\rho(x)$ (1)

Where $\Phi(x)$: the electrostatic potential

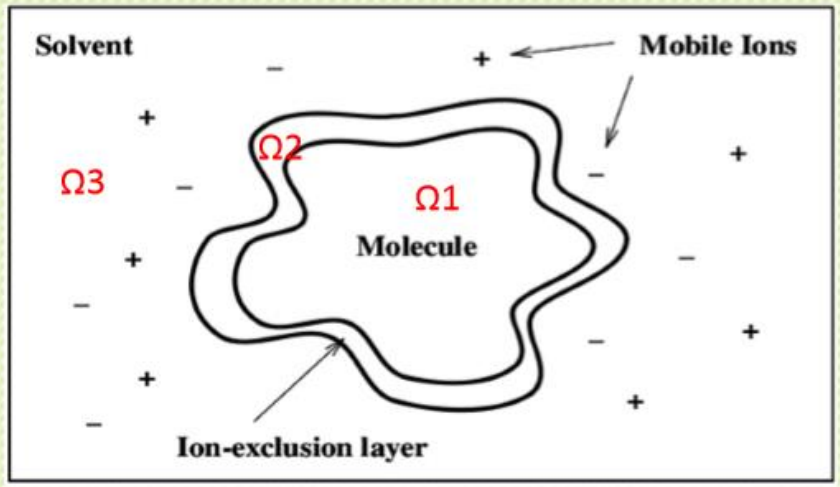
$\varepsilon(x)$: the spatial dielectric function

$$\varepsilon(x)=\begin{cases}\varepsilon_1 & r\in\Omega_1\\ \varepsilon_2=(\varepsilon_3) & r\in\Omega_2 \text{ and } \Omega_3\end{cases}$$

k(x): a modified Debye-Huckel parameter

$$k(x)=\begin{cases}0 & r\in\Omega_1 \text{ and } \Omega_2\\ \sqrt{\varepsilon_3}\,k & r\in\Omega_3\end{cases}$$

$\rho(x)$: the charge distribution function



Ω1: molecule with fixed ions, ε_1

Ω2: exclusion layer (no ions), $\varepsilon_2=\varepsilon_3$

Ω3: solution with mobile ions, ε_3

Assume that all ions in solution are univalent: $q=\pm e_c$

On $\Gamma_{12} = \Omega_1 \cap \Omega_2, \Phi_1(r) = \Phi_2(r), \varepsilon_1\nabla\Phi_1(r)\cdot n = \varepsilon_2\nabla\Phi_2(r)\cdot n$

On $\Gamma_{23} = \Omega_2 \cap \Omega_3, \Phi_2(r) = \Phi_3(r), \varepsilon_2\nabla\Phi_2(r)\cdot n = \varepsilon_3\nabla\Phi_3(r)\cdot n$

The appropriate boundary conditions for the infinite domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_1 = R^3$ are $\Phi(\infty)=0$.

Numerical Method

Let $K_0(x_0, y_0, z_0)$ be an arbitrary grid point away from the boundary of Ω.
Applying finite difference formulation yields an iteration equation for eqn.1:

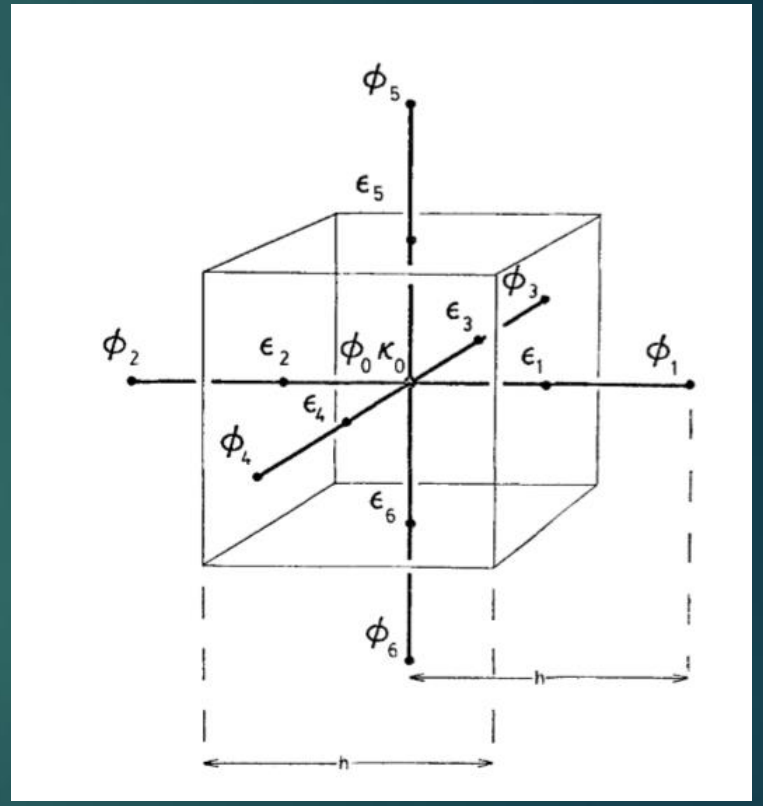
$$\Phi_0 = \frac{\sum_{i=1}^{6} \varepsilon_i \Phi_i + 4\pi q_0/h}{\sum_{i=1}^{6} \varepsilon_i + (kh)^2 \frac{\sinh(\Phi_0)}{\Phi_0}} \qquad (2)$$

Φ_0 is the potential at K_0;
Φ_i is the potential at the six nearest neighboring grids of K_0;
ε_i is the dielectric constant at the midpoint between Φ_0 and Φ_i;
q_0 is the charge assigned to K_0;
h is the grid spacing between grid points.

Eqn.2 can be written in matrix form as

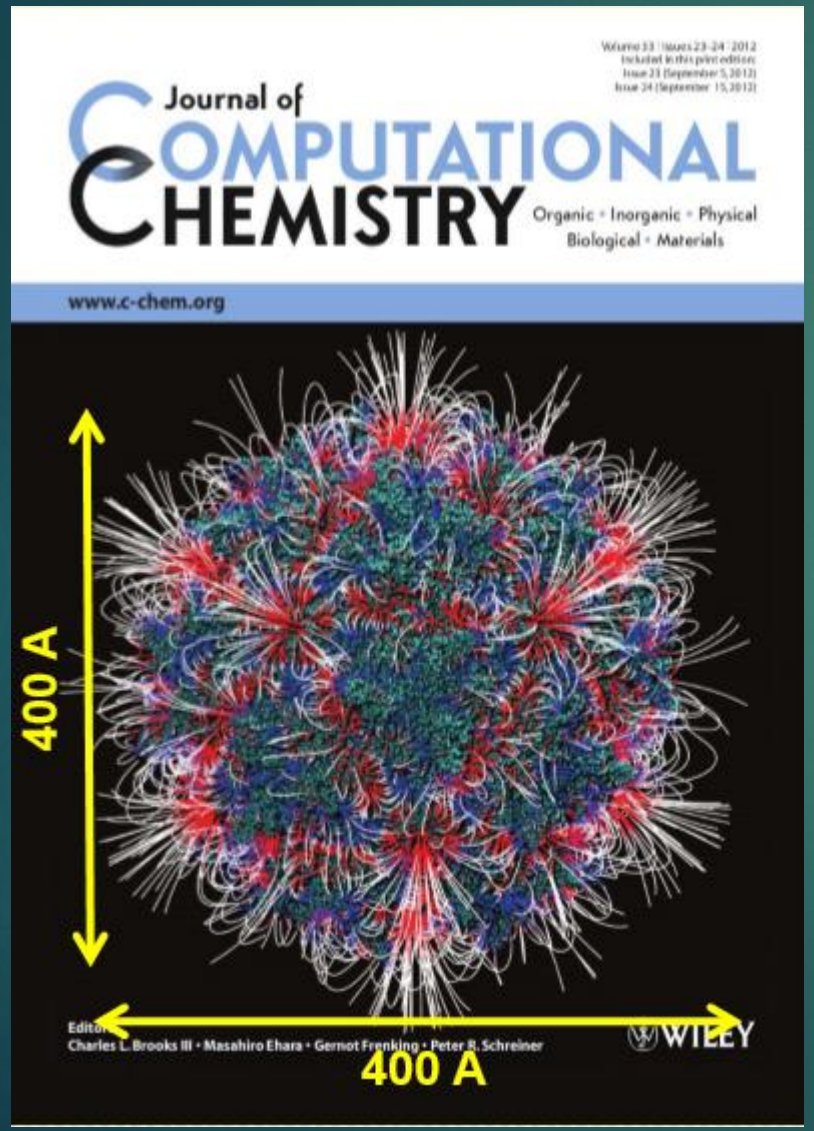
$$\phi = T\phi + Q \qquad (3)$$

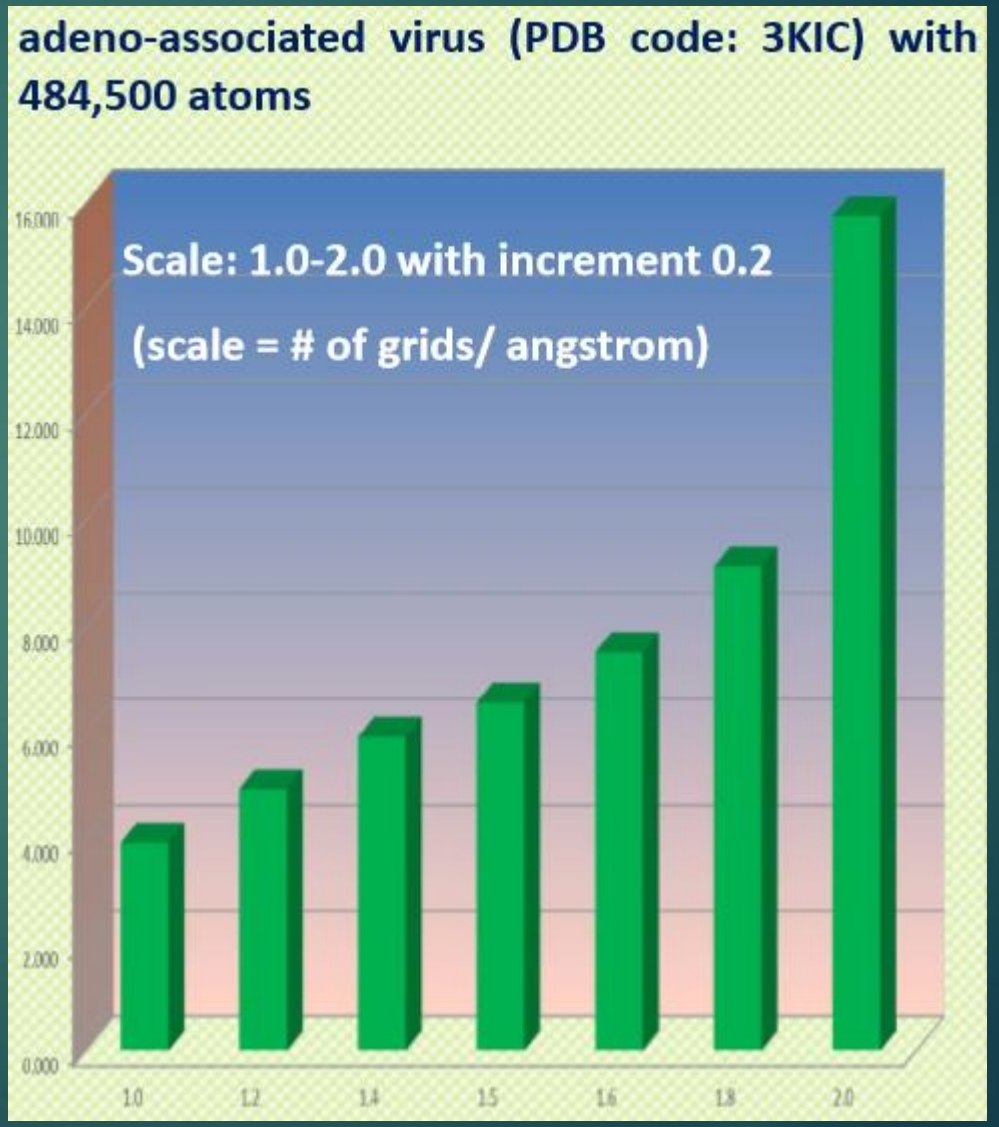
Where T is the coefficient matrix;
ϕ and Q are column vectors.



Klapper, I., et al., Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effects of ionic strength and amino-acid modification.
Proteins: Structure, Function, and Bioinformatics, 1986. 1(1): p. 4759.

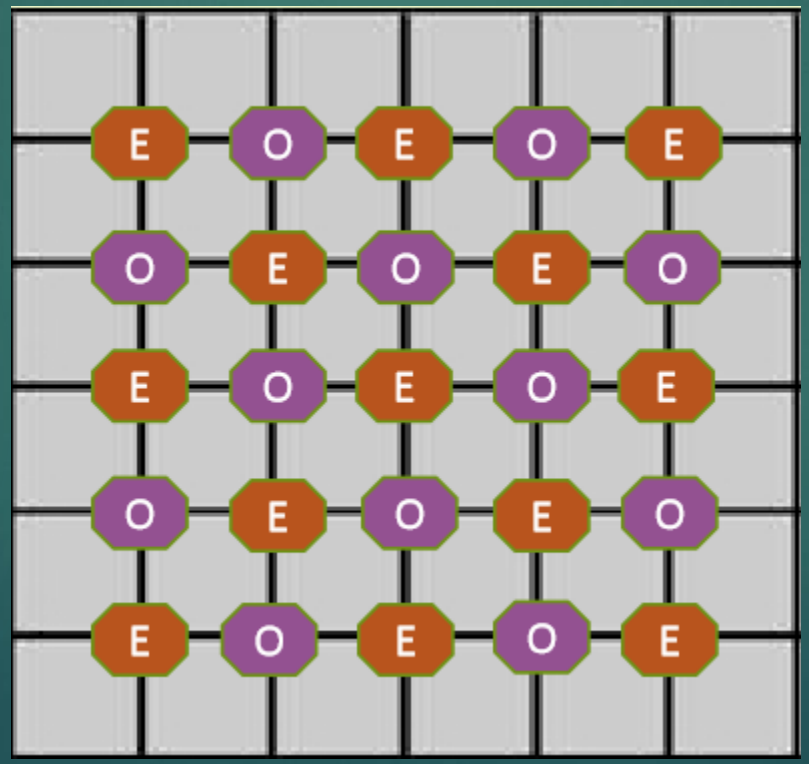
An Example of Large Proteins





“Checkboard” Ordering

- Each grid point is assigned as odd or even by the sum of its grid coordinates.
- The six nearest neighbors to any grid point must be of opposite nature.
- Every even point is surrounded by odd points and vice verse.
- This technique was presented in a paper written by Anthony Nicholls and Barry Honig.



“Checkboard” Ordering, cont.

Provided L an odd number,

$$\Phi = \begin{bmatrix} \Phi_{even} \\ \Phi_{odd} \end{bmatrix} \qquad Q = \begin{bmatrix} Q_{even} \\ Q_{odd} \end{bmatrix} \quad (5)$$

$$T = \begin{bmatrix} \mathbf{0} & T_{odd} \\ T_{even} & \mathbf{0} \end{bmatrix} \quad (6)$$

After reordering Φ and Q,

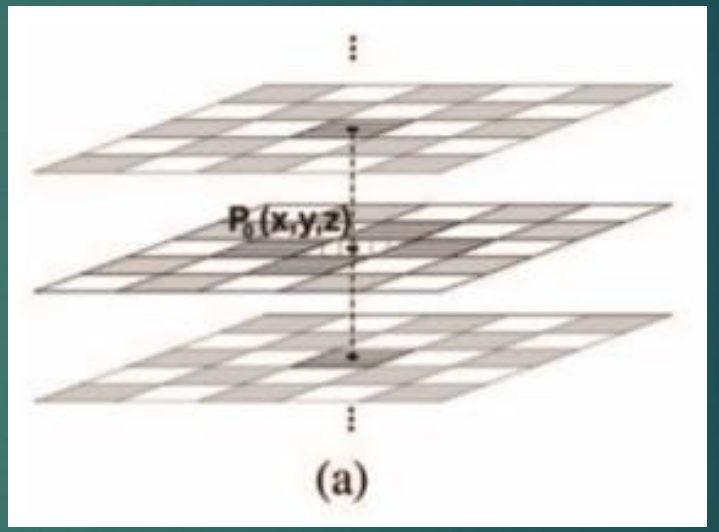
$\Phi = T\Phi + Q$ (3) $\Longleftrightarrow$

$$\begin{cases} \Phi_{Even}^{n+1} = T_{Odd}\Phi_{Odd}^{n} + Q_{Even} \\ \Phi_{Odd}^{n+1} = T_{Even}\Phi_{Even}^{n+1} + Q_{Odd} \end{cases} \quad (7)$$

Contiguous Memory Mapping

- Map the odd and even points separately into two contiguous memory/arrays.
- This technique was presented in the paper written by Anthony Nicholls and Barry Honig

Mapping

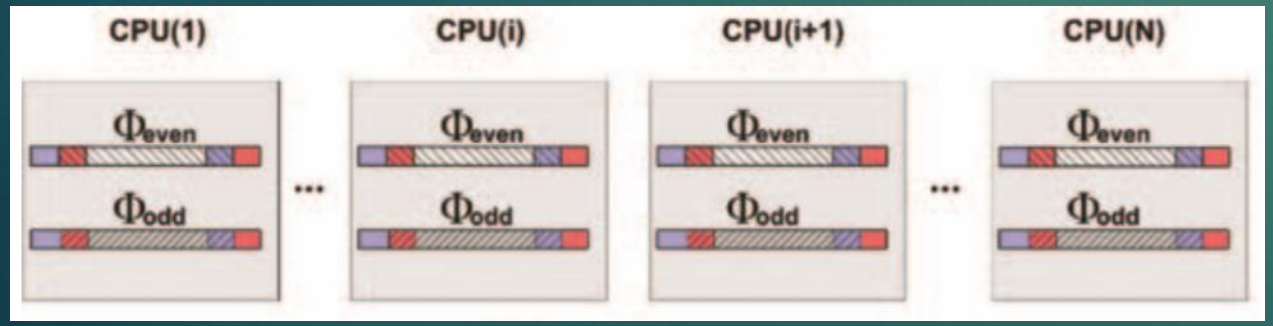


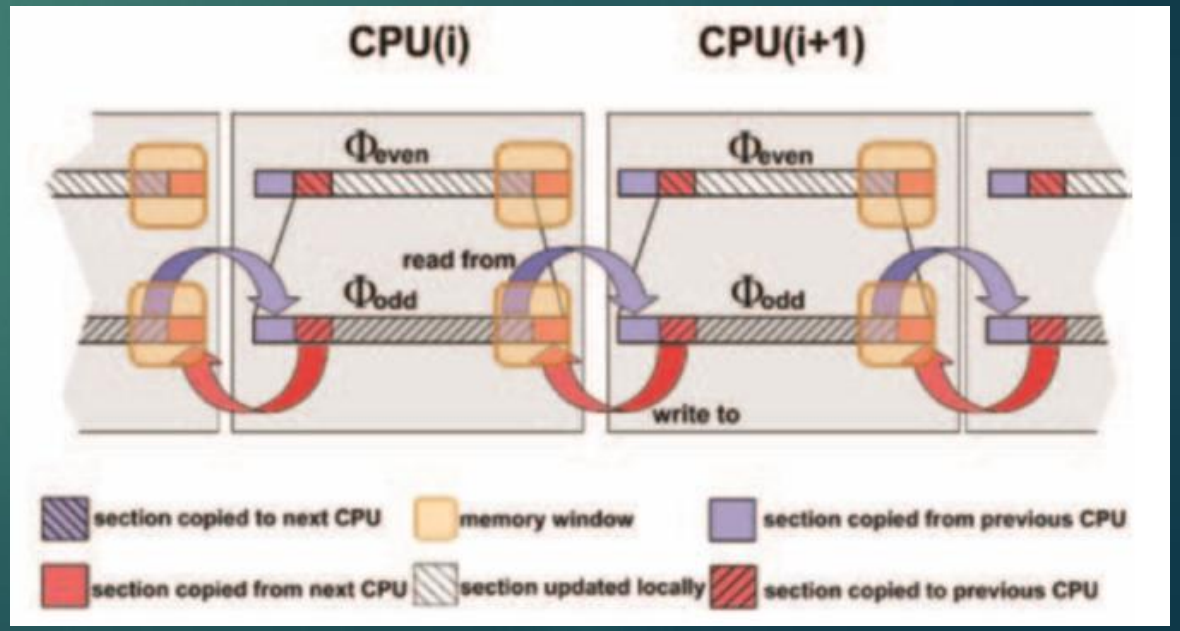
Odd points	...	33	35	37	39	41	43	45	47	...
Even points	...	34	36	38	40	42	44	46	48	...

Parallelizing Iteration:
An Algorithmic Parallelization Technique for GS/SOR Iterations

- Parallelizing eqn.7

$$\begin{cases} \Phi_{Even}^{n+1} = T_{Odd}\Phi_{odd}^{n} + Q_{Even} \\ \Phi_{Odd}^{n+1} = T_{Even}\Phi_{Even}^{n+1} + Q_{Odd} \end{cases} \qquad (7)$$





Introduction to DelPhi Program

- Open-source finite difference solver utilizing GS and SOR iterations to model electrostatics in molecular biology.
- DelPhi was developed in Dr. Barry Honig's lab at Columbia University in 1986 and has been continuously improved in Dr. Emil Alexov's lab at Clemson University

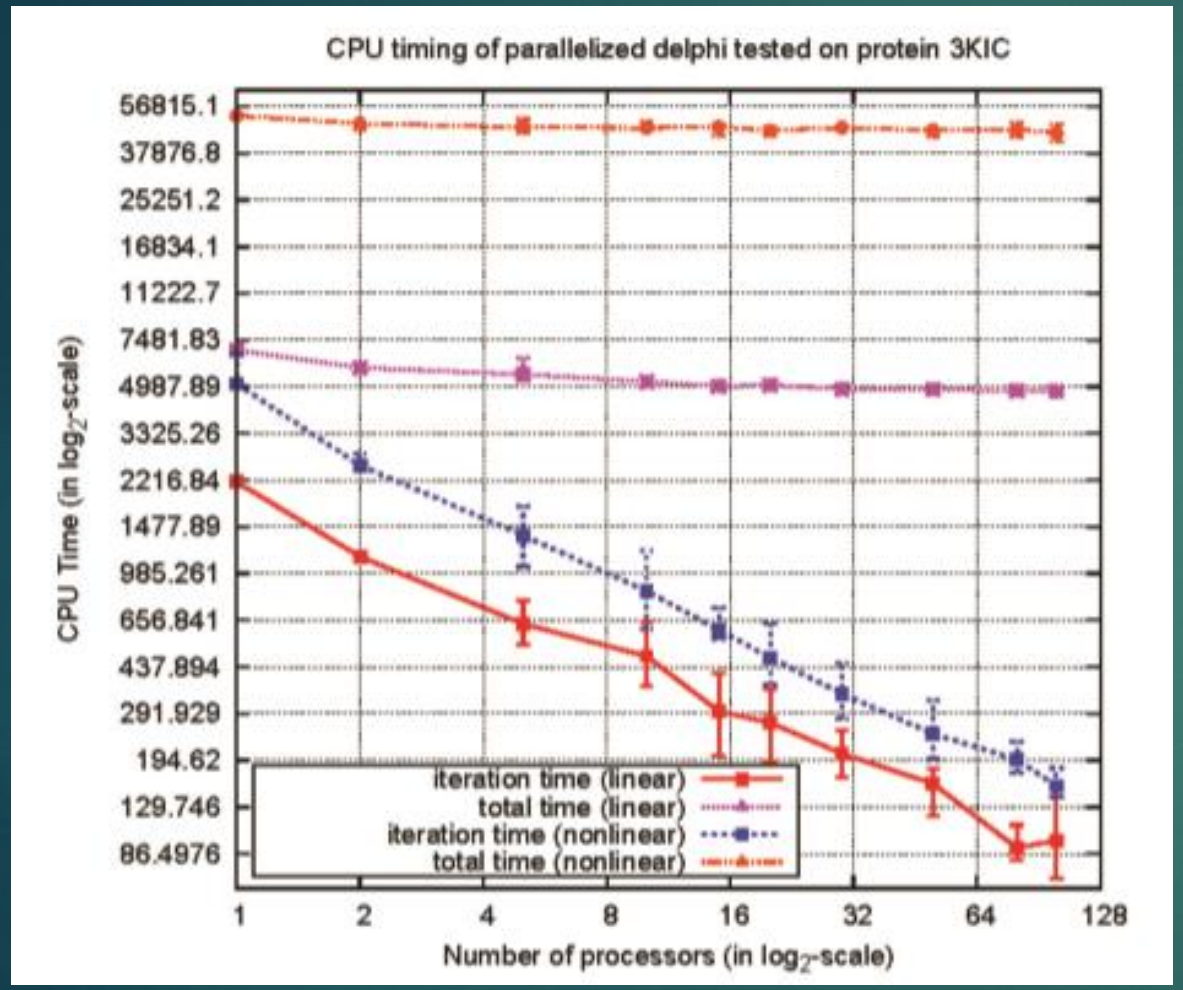
http://compbio.clemson.edu/delphi

Numerical Experiments

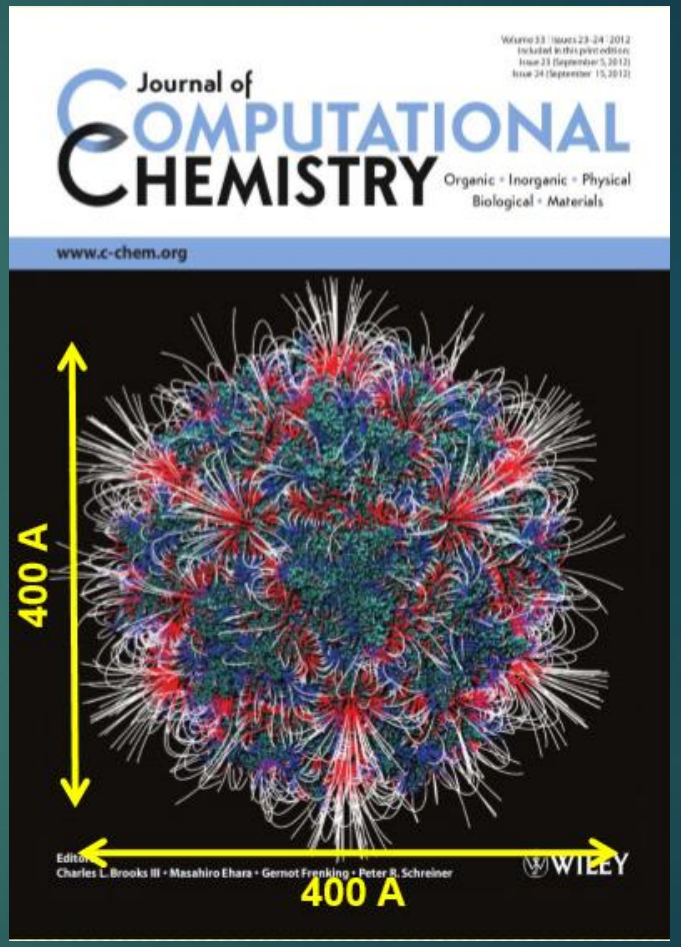
- Performed using a dedicated queue on Palmetto cluster at Clemson University
- 1 node with 100 GB memory (master process)
- 100 nodes with 30 GB memory (slave process)
- Used one CPU per node in order to avoid potential memory competition on one node.
- gcc version 4.5.1
- MPICH2 version 1.4
- All identical runs were repeated 5 times and their averages are reported here in order to reduce random fluctuations caused by system workload and network communication.
- Both sequential and parallel experiments were given enough memory in order to avoid possible data exchange between memory and hard disk.

Implementation Results

Object: the protein of human adeno-associated virus 3KIC



Execution (purple) and iteration (red) time for solving the linear PBE, compared to execution (orange) and iteration (blue) time for solving the nonlinear PBE.



Resulting Electrostatic Field

Special thanks to Dr. Emil Alexov and Dr. Chuan Li!

Grant: "New Generation DelPhi: large systems and beyond electrostatics" (NIH grant #: 5R01GM093937-07) for developing in DelPhi parallel computing algorithms and numerical methods for solving time-dependent differential equations with numerous applications in molecular biology.

Thank you!

Any questions?

References

Klapper, Isaac, Ray Hagstrom, Richard Fine, Kim Sharp, and Barry Honig. "Focusing of Electric Fields in the Active Site of Cu-Zn Superoxide Dismutase: Effects of Ionic Strength and Amino-acid Modification." *Proteins: Structure, Function, and Bioinformatics*. Wiley Subscription Services, Inc., A Wiley Company, 03 Feb. 2004. Web. 31 Mar. 2017.

Li, Chuan, Lin Li, Jie Zhang, and Emil Alexov. "Highly Efficient and Exact Method for Parallelization of Grid-based Algorithms and Its Implementation in DelPhi." *Journal of Computational Chemistry* 33.24 (2012): 1960-966. Web.

Nicholls, Anthony, and Barry Honig. "A Rapid Finite Difference Algorithm, Utilizing Successive Over-relaxation to Solve the Poisson-Boltzmann Equation." *Journal of Computational Chemistry* 12.4 (1991): 435-45. Web.

"Why Is Molecular Modeling Important?" *Why Is Molecular Modeling Important?* N.p., n.d. Web. 25 Mar. 2017.