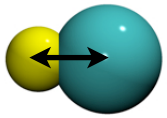
Generalized Born Implicit Solvent Algorithm in NAMD

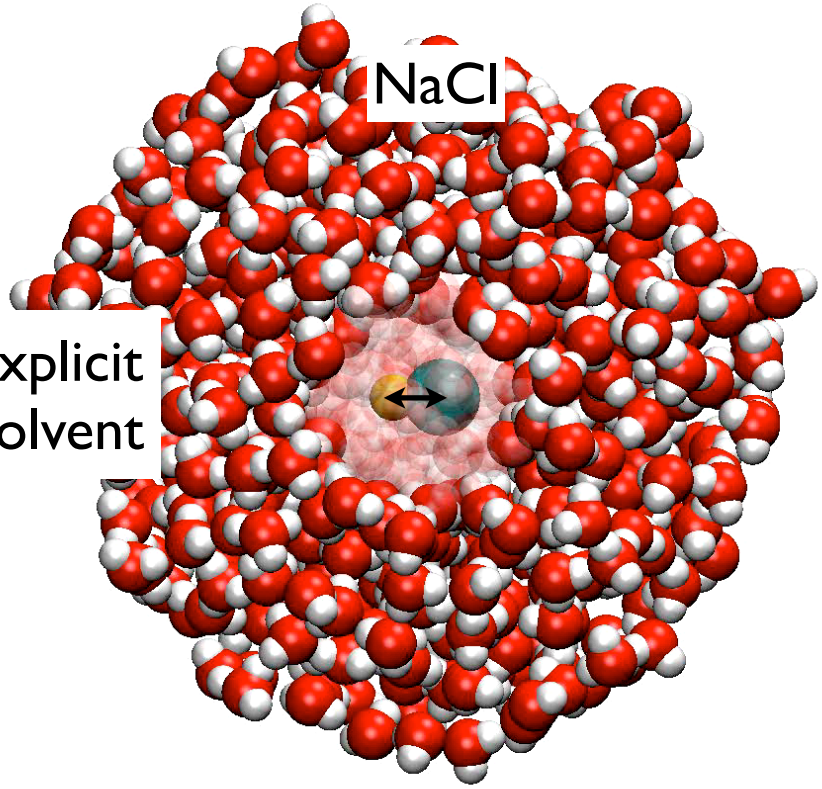
David Hardy
http://www.ks.uiuc.edu/Research/gpu/
NAIS: State-of-the-Art Algorithms for Molecular Dynamics

(Presenting the work of David Tanner.)

Generalized Born Implicit Solvent

Implicit Solvent





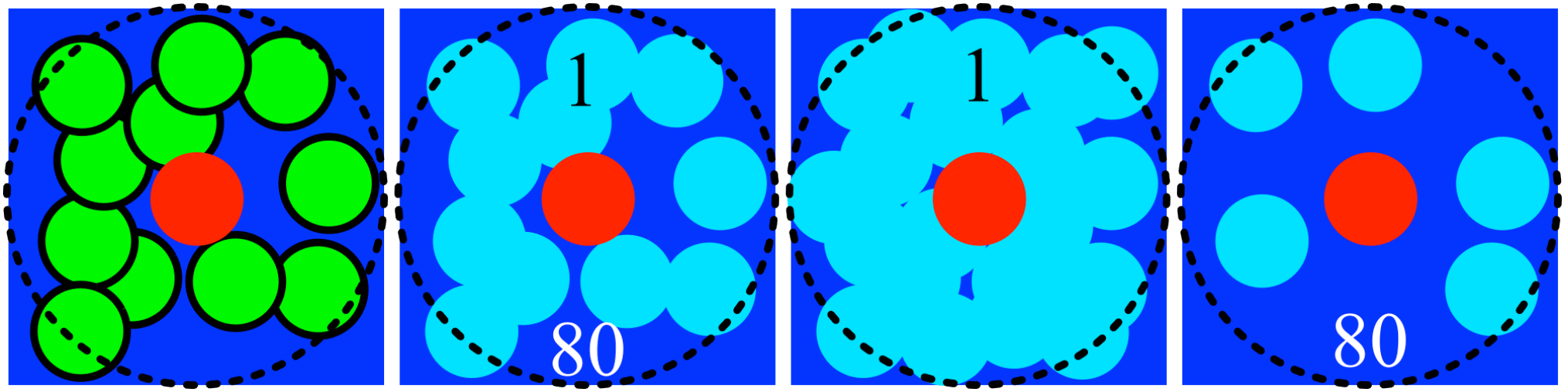
Explicit Solvent

- Benefits
  - instantaneous relaxation of water
  - low viscosity
  - increased diffusion
  - increased flexibility
- Caveats
  - describes bulk water only
  - no hydrogen bonding

Fewer atoms should be faster.

	Explicit	Implicit
Coulomb:	83 kcal/mol/Å	83 kcal/mol/Å
Solvation:	-32 kcal/mol/Å	-30 kcal/mol/Å
Total Elec:	51 kcal/mol/Å	53 kcal/mol/Å

Born Radius



Each atom exposed to other atoms and solvent	Surrounded by: Solvent (polar) ε=80 Protein (nonpolar) ε=1	Many Neighbors ε=1 low screening large Born radius	Few Neighbors ε=80 high screening small Born radius

GBIS is Expensive (7X)

Coulomb

$$\vec{F}_i = -\frac{k_e}{\epsilon_p}\sum_j \frac{q_i q_j}{r_{ij}^2}\hat{r}_{ij}$$

Generalized Born Implicit Solvent

$$\vec{F}_i^{GB} = -\sum_j \left[\frac{\partial E_{ij}^{GB}}{\partial r_{ij}} + \frac{\partial E_T^{GB}}{\partial \alpha_i}\frac{d\alpha_i}{dr_{ij}} + \frac{\partial E_T^{GB}}{\partial \alpha_j}\frac{d\alpha_j}{dr_{ij}}\right]\hat{r}_{ji}$$

$$E_{ij}^{\mathrm{GB}} = -k_e D_{ij} q_i q_j / f_{ij}^{\mathrm{GB}}$$

$$f_{ij}^{\mathrm{GB}} = \sqrt{r_{ij}^2 + \alpha_i\alpha_j \exp\left(-r_{ij}^2/4\alpha_i\alpha_j\right)}$$

$$\psi_i = \rho_{i0}\sum_{j\in N(i)} H_{ij}$$

$$\alpha_i = \left[1/\rho_{i0} - 1/\rho_i \tanh\left(\delta\psi_i - \beta\psi_i^2 + \gamma\psi_i^3\right)\right]^{-1}$$

$$\frac{\partial E_T^{GB}}{\partial \alpha_k} = \sum_i\sum_{j>i}\left[\frac{\partial E_{ik}^{GB}}{\partial \alpha_k} + \frac{\partial E_{kj}^{GB}}{\partial \alpha_k}\right] + \sum_i \frac{\partial E_{ii}^{GB}}{\partial \alpha_k}$$

$$\frac{\partial E_{ij}^{\mathrm{GB}}}{\partial r_{ij}} = -k_e\frac{q_i q_j r_{ij}}{f_{ij}^2}\left[1 - \frac{1}{4}\exp\left(\frac{-r_{ij}^2}{4\alpha_i\alpha_j}\right)\right]\left[\frac{\kappa}{\epsilon_s}\exp\left(-\kappa f_{ij}\right) - \frac{D_{ij}}{f_{ij}}\right]$$

$$\frac{\partial E_{ij}^{\mathrm{GB}}}{\partial \alpha_i} = -\frac{1}{\alpha_i}\frac{k_e q_i q_j}{2f_{ij}^2}\left(\frac{\kappa}{\epsilon_s}\exp\left(-\kappa f_{ij}\right) - \frac{D_{ij}}{f_{ij}}\right)\left(\alpha_i\alpha_j + \frac{r_{ij}^2}{4}\right)\exp\left(\frac{-r_{ij}^2}{4\alpha_i\alpha_j}\right)$$

$$\frac{\partial E_{ij}^{\mathrm{GB}}}{\partial \alpha_j} = -\frac{1}{\alpha_j}\frac{k_e q_i q_j}{2f_{ij}^2}\left(\frac{\kappa}{\epsilon_s}\exp\left(-\kappa f_{ij}\right) - \frac{D_{ij}}{f_{ij}}\right)\left(\alpha_i\alpha_j + \frac{r_{ij}^2}{4}\right)\exp\left(\frac{-r_{ij}^2}{4\alpha_i\alpha_j}\right)$$

$$\frac{d\alpha_i}{dr_{ij}} = \frac{\alpha_i^2\rho_{i0}}{\rho_i}\left(1 - \tanh^2\left(\delta\psi_i - \beta\psi_i^2 + \gamma\psi_i^3\right)\right)\left(\delta - 2\beta\psi_i + 3\gamma\psi_i^2\right)\frac{\partial H_{ij}}{\partial r_{ij}}$$

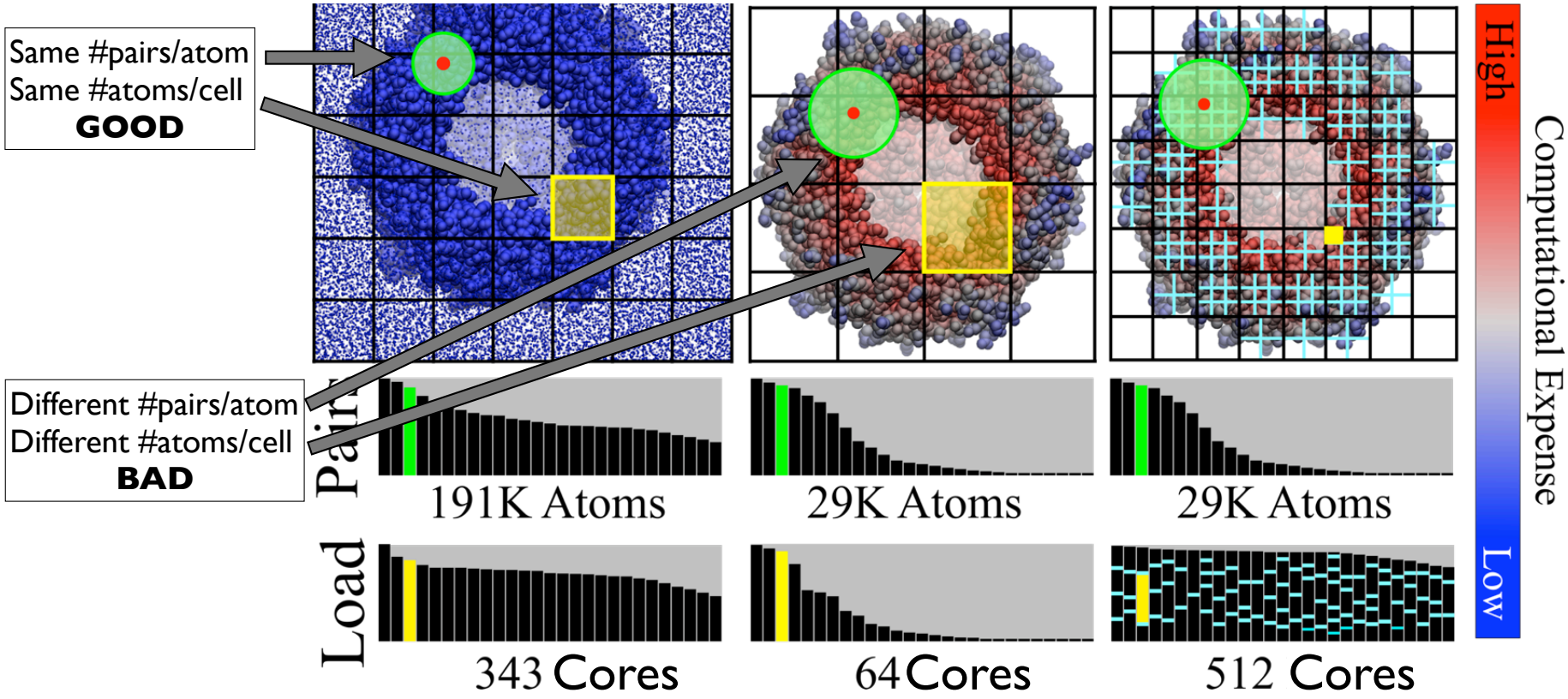
$$\frac{d\alpha_j}{dr_{ij}} = \frac{\alpha_j^2\rho_{j0}}{\rho_j}\left(1 - \tanh^2\left(\delta\psi_j - \beta\psi_j^2 + \gamma\psi_j^3\right)\right)\left(\delta - 2\beta\psi_j + 3\gamma\psi_j^2\right)\frac{\partial H_{ji}}{\partial r_{ij}}$$

$$H_{ij} = \begin{cases} \frac{1}{8r_{ij}}\left[1 + \frac{2r_{ij}}{r_{ij}-\rho_{js}} + \frac{1}{r_c^2}\left(r_{ij}^2 - 4r_c r_{ij} - \rho_{js}^2\right) + 2\ln\frac{r_{ij}-\rho_{js}}{r_c}\right] \\ \frac{\rho_{js}^2}{r_{ij}^2}\frac{\rho_{js}}{r_{ij}^2}\left[a + \frac{\rho_{js}^2}{r_{ij}^2}\left(b + \frac{\rho_{js}^2}{r_{ij}^2}\left(c + \frac{\rho_{js}^2}{r_{ij}^2}\left(d + \frac{\rho_{js}^2}{r_{ij}^2}e\right)\right)\right)\right] \\ \frac{1}{2}\left[\frac{\rho_{js}}{r_{ij}^2-\rho_{js}^2} + \frac{1}{2r_{ij}}\ln\frac{r_{ij}-\rho_{js}}{r_{ij}+\rho_{js}}\right] \\ \frac{1}{4}\left[\frac{1}{\rho_{i0}}\left(2 - \frac{1}{2r_{ij}\rho_{i0}}\left(r_{ij}^2 + \rho_{i0}^2 - \rho_{js}^2\right)\right) - \frac{1}{r_{ij}+\rho_{js}} + \frac{1}{r_{ij}}\ln\frac{\rho_{i0}}{r_{ij}+\rho_{js}}\right] \end{cases}$$

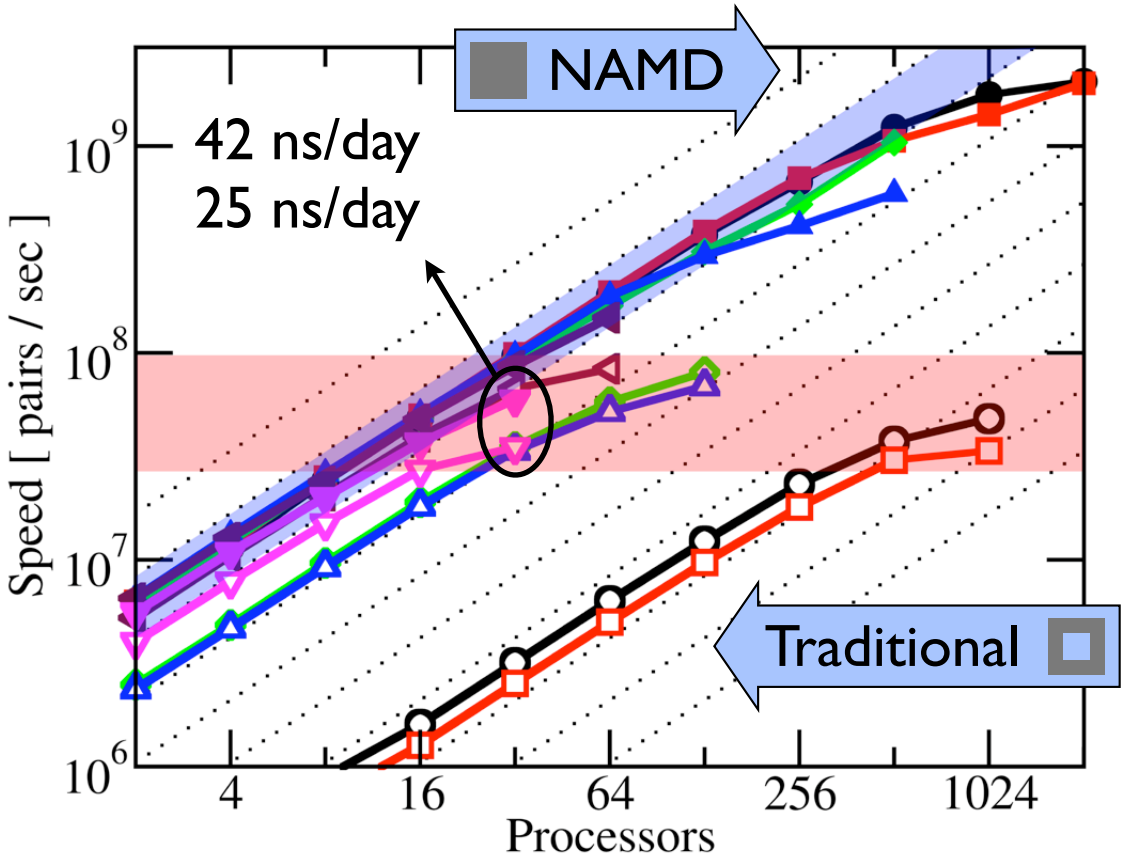
Onufriev, Bashford and Case. Exploring Protein Native States and Large-Scale Conformational Changes With a Modified Generalized Born Model. *Proteins: Struct., Funct., Bioinf.* 55:383–394 (2004).

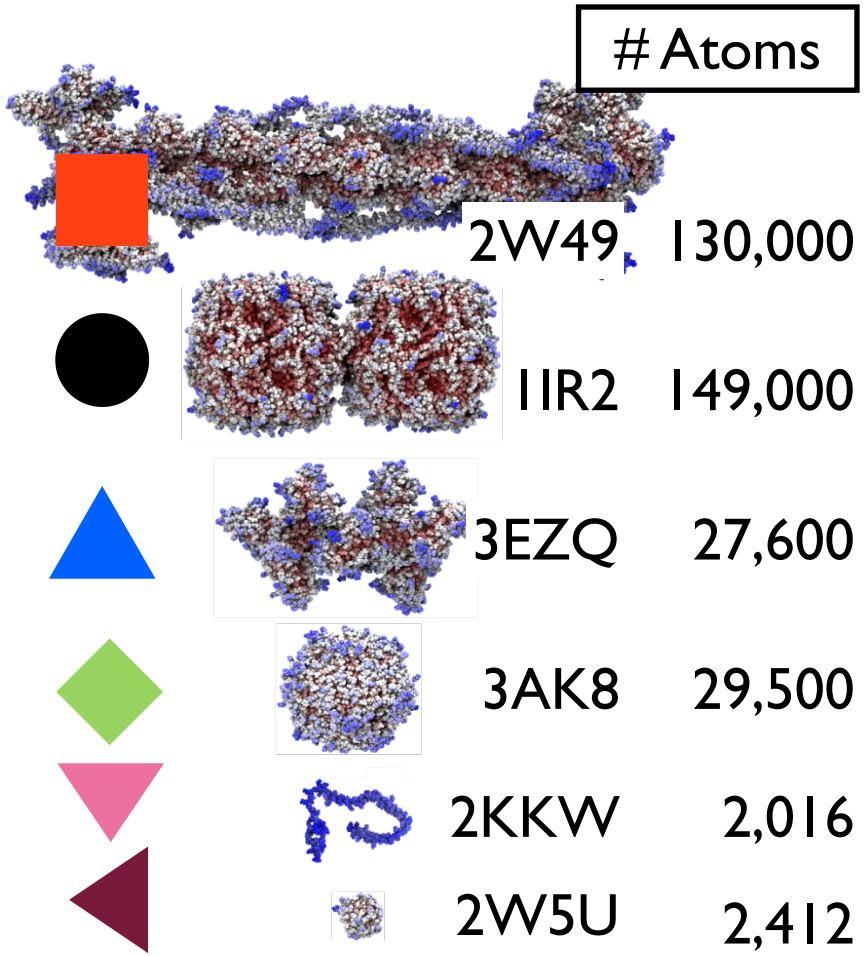
Parallel Computation

Explicit GB (Naive) GB (NAMD)



Excellent Parallel Performance





GB / SA Implicit Solvent

Generalized Born	Solvent-Accessible Surface Area

Hydrophilic Free Energy of Solvation
- polar effect
- Coulomb interactions of water
- screens electrostatics

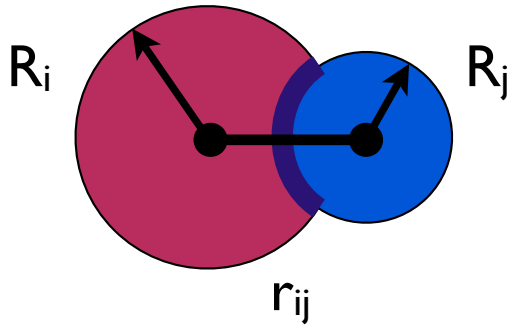
Hydrophobic Free Energy of Solvation
- nonpolar effect
- van der Waals interaction of water
- surface tension constricts protein

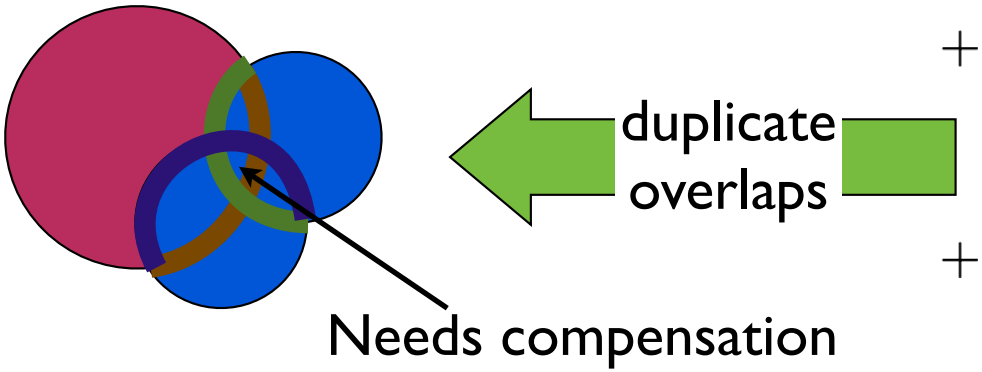
$$E = Surf\_Tension * Surf\_Area$$

$$[kcal/mol/Å^2] * [Å^2]$$

D. Sitkoff, K. Sharp, and B. Honig. Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem. 98:1978–1988 (1994).

Linear Combination of Pairwise Overlaps[8]





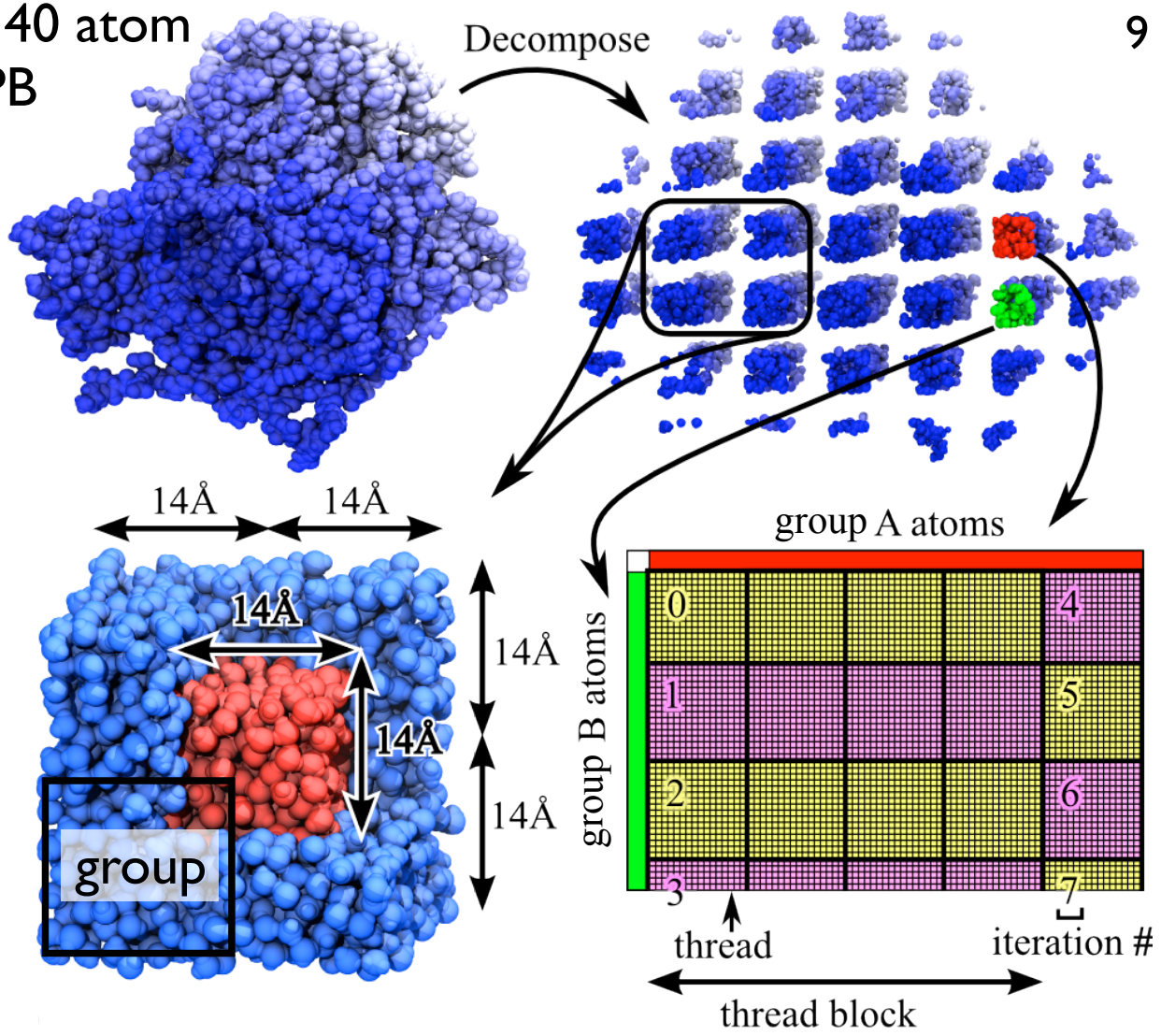
$$A_{ij} = 2\pi R_i \left[(R_i - r_{ij}/2 - (R_i^2 - R_j^2)/(2r_{ij}) \right]$$

$$\begin{aligned}
\mathrm{SA}_i &= P_{1,i} 4\pi R_i^2 \\
&+ P_{2,i} \sum_{j \in N(i)} A_{ij} \\
&+ P_{3,i} \sum_{j \in N(i)} A_{ij} \sum_{k \in N(i) \cap N(j)} A_{jk} \\
&+ P_{4,i} \sum_{j \in N(i)} \left[A_{ij} \sum_{j \in N(i)} \sum_{k \in N(i) \cap N(j)} A_{jk} \right]
\end{aligned}$$

J. Weiser et al. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J. Comp. Chem. 20:217–230 (1998).

Hybrid GB / SA ↓ on ↓ GPU / CPU

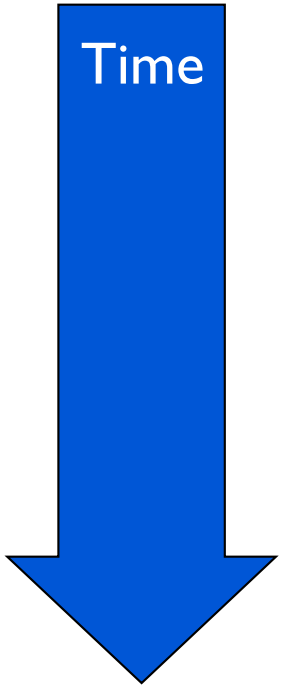
	Total	Valid	Fract
GB	1.7×10^8	6.5×10^6	3%
SA	2.3×10^{12}	5.2×10^6	0.0002%



Hybrid Calculation

The GPU requires a host CPU to copy data and submit work.

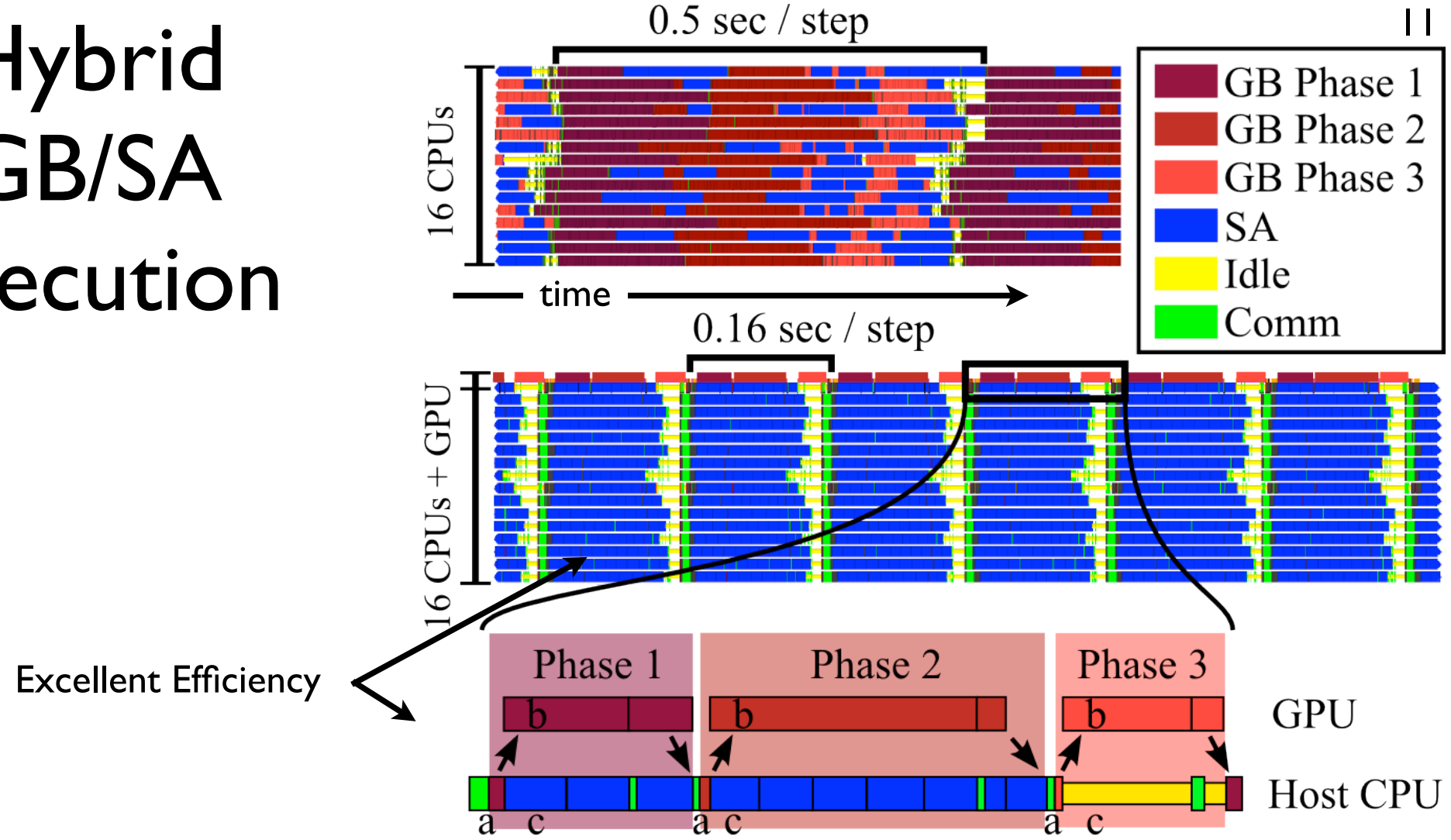
copy data to GPU
initialize work on GPU
SA
receive data from GPU
perform GB reduction
copy data to GPU
initialize work on GPU
SA
receive data from GPU
...



Solution

Allow switching by dividing SA calculation into small work units.

Hybrid GB/SA Execution



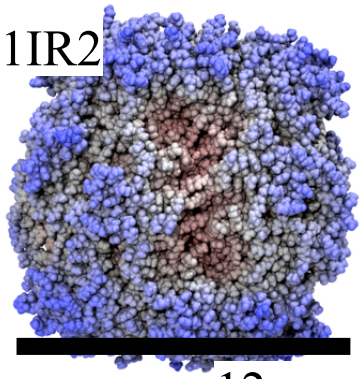
Performance

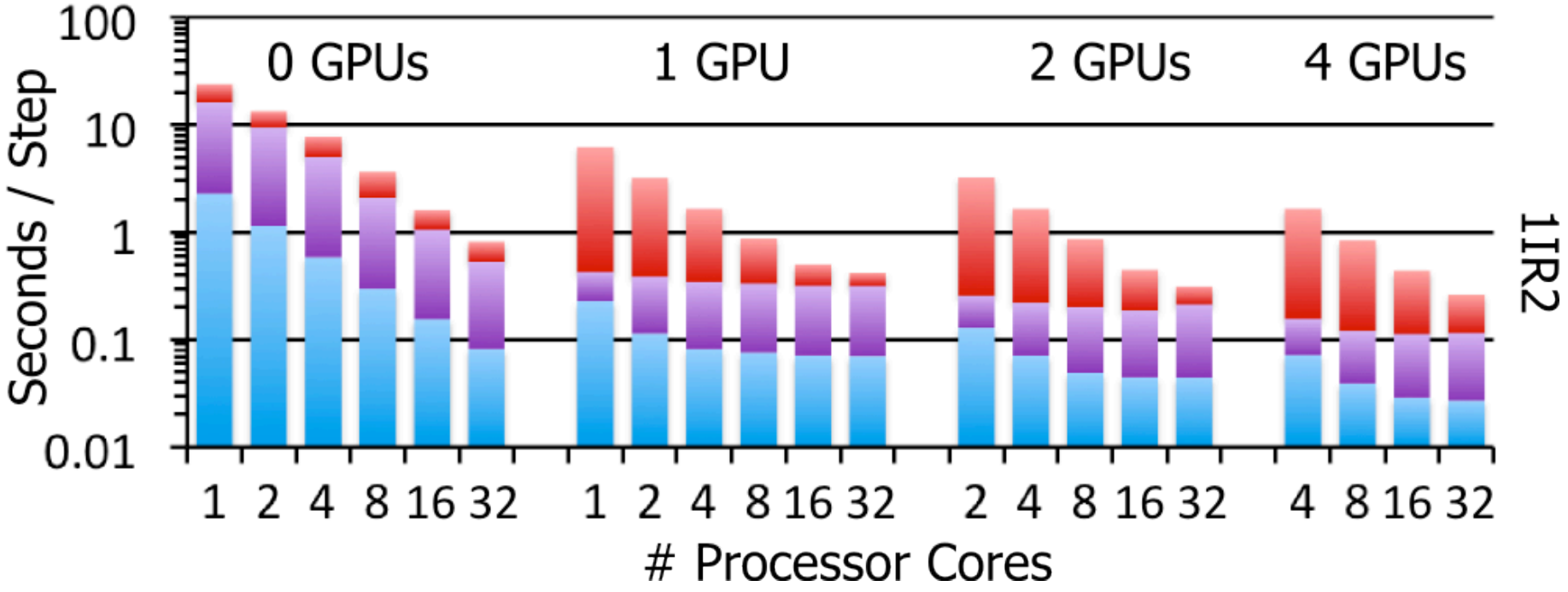
74,930 (1IR2):	10 ps/day (1c)	410 ps/day	(1g1c)	=	40X
13,340 (1GPB):	0.33 ns/day (1c)	9.4 ns/day	(1g16c)	=	28X

GB/SA

GB

in vacuo





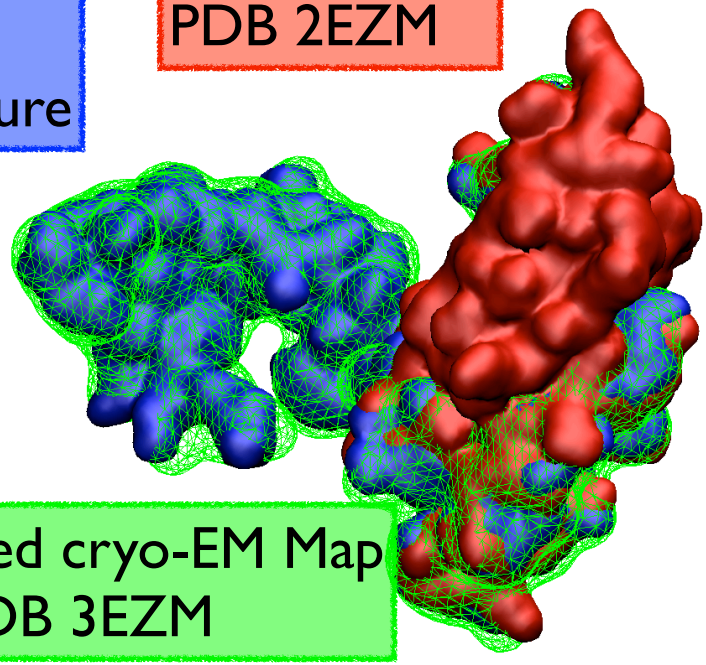
Molecular Dynamics Flexible Fitting

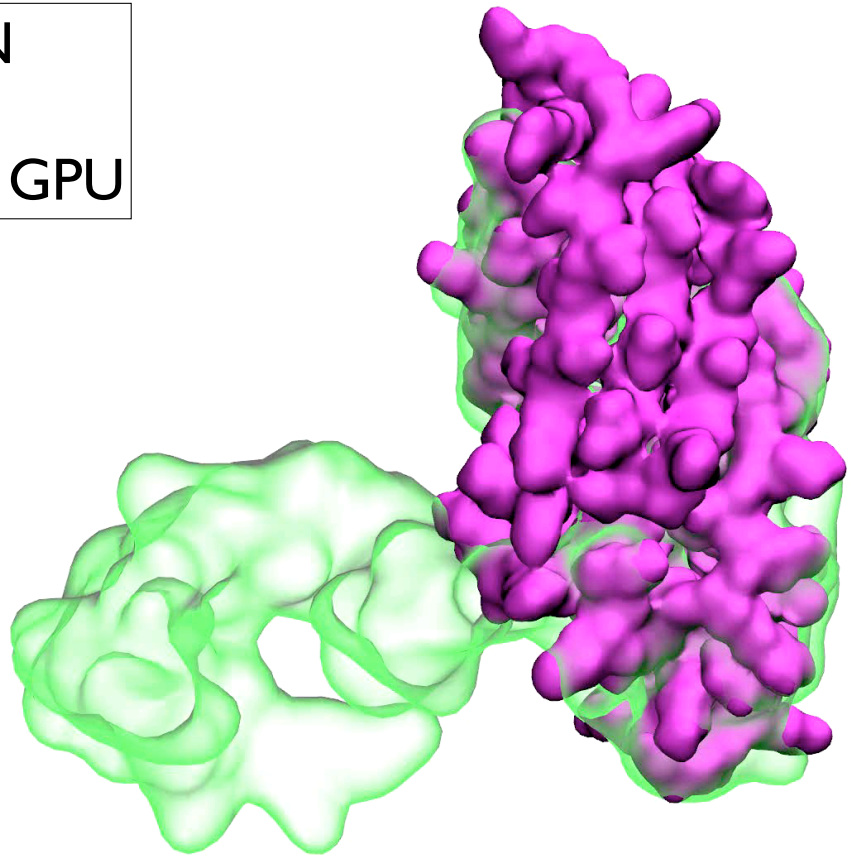
Cyanovirin-N
1,500 atoms
40 ns/day on GPU

Initial Crystal Structure PDB 2EZM

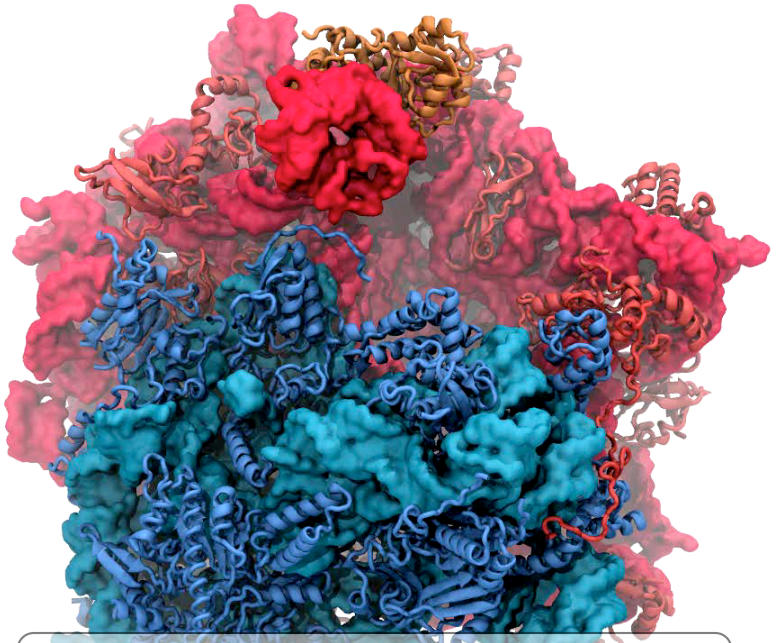
Fitted Structure

Simulated cryo-EM Map from PDB 3EZM

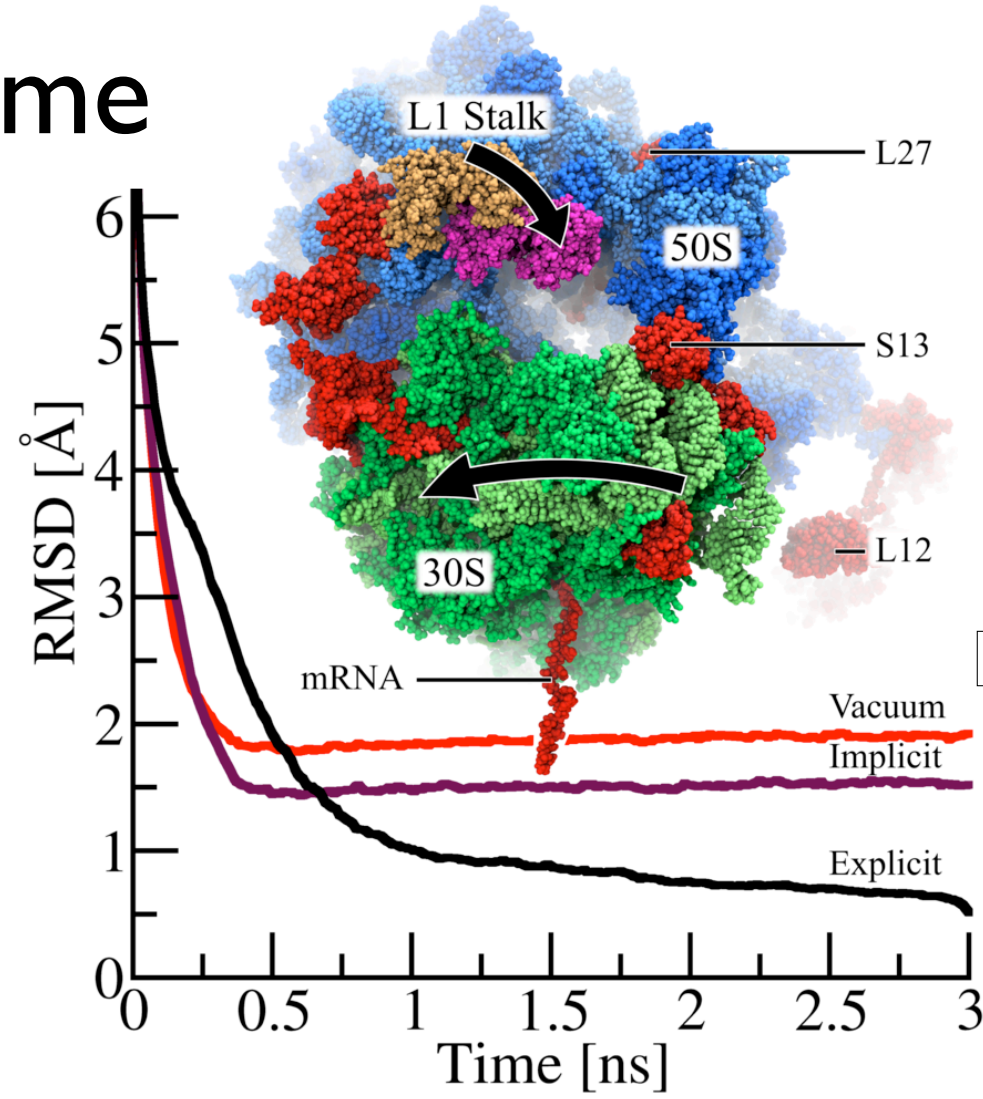




MDFF of Ribosome



fitting "classical-state" crystal structure
  50S (2I2V) is 3.2Å resolution
  30S (2I2U) is 3.2Å resolution
into "ratcheted-state" cryo-EM density map
  8.9Å resolution



1000 proc.
37 ns/day
5.2 ns/day
3.6 ns/day