Large-Scale Biomolecular Simulations: Biomedical and Bioenergy Applications

Kevin Sanbonmatsu (PI)

Los Alamos National Laboratory

www.t10.lanl.gov/kys

Thank you organizers

High Speed Computing, Glenenden, Oregon, April 28, 2009

Outline

1. Biomolecular simulations - overview
2. Biomedical applications – antibiotics
3. Bioenergy applications – cellulosic ethanol

Overview of biomolecular simulation

Computational biology

- Sequence analysis / bioinformatics
- Systems biology – coupled ODEs.
- Quantum calculations – reaction mechanism
- Molecular dynamics with electrostatics – molecular machines and binding

Typically exert largest demand on CPU resources - 10^8-10^{11} time steps, 10^5-10^7 atoms, 1000-10,000 cores for 6-18 months per project.

Molecular Dynamics Simulation

For each atom, solve $\mathbf{F}_i = m_i \mathbf{a}_i$, where $\mathbf{F}_i = -\nabla U$, repeat 10^7-10^{10} times

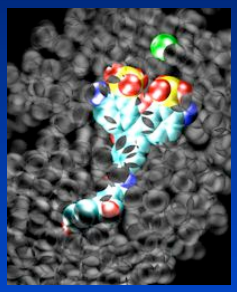
$U = \Sigma \, ½ \, K_b (r-r_0)^2$ Bonds

$+ \Sigma \, ½ \, K_\theta (\theta-\theta_0)^2$ Angles

$+ \Sigma \, K_\varphi (1 - \cos (n\varphi + \delta))$ Torsional Angles

$+ \Sigma \, \varepsilon ((r/r_0)^{-12} - 2(r/r_0)^{-6})$ Van der Waals

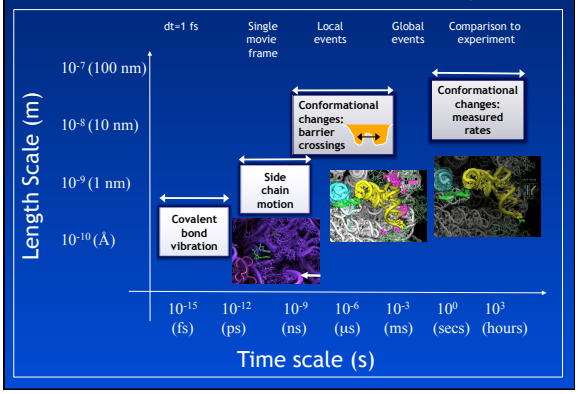
$+ \Sigma \, \zeta \, q_i \, q_j / r$ Electrostatic

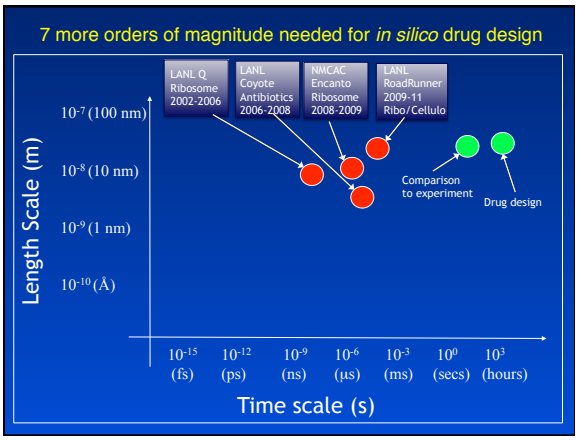


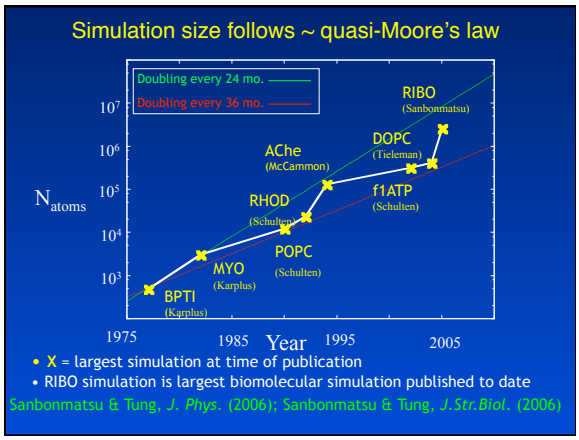
ALEXA 488

Validation: Garcia and Sanbonmatsu, PNAS 2002

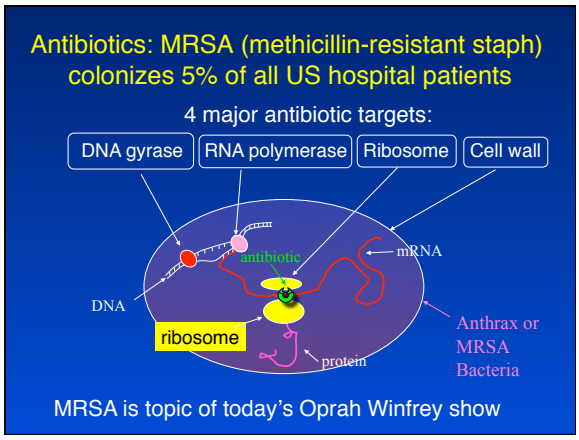
Biomolecular time scales span > 15 orders of magnitude

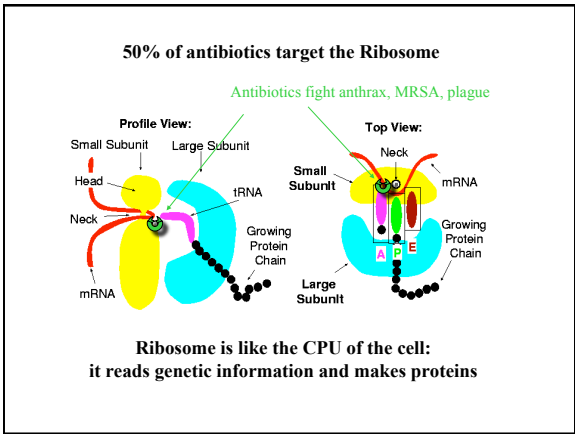


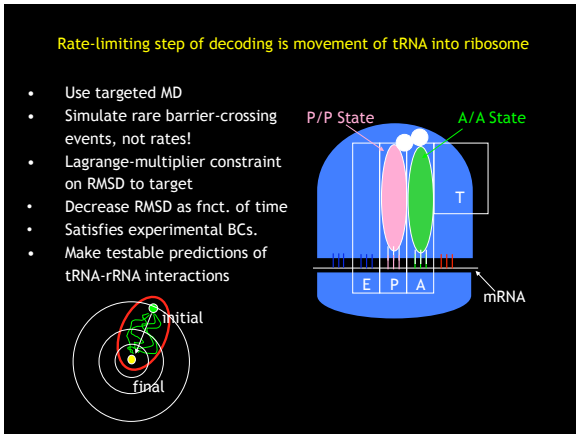




Biomedical applications: antibiotics





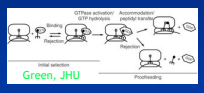


Available data on accommodation

Chemical Protection and mutation

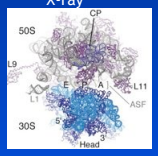
Noller UCSC

Rapid Kinetics



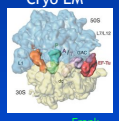
Green, JHU

X-ray



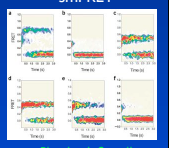
Cate, UCB

Cryo EM



Frank HHMI

smFRET

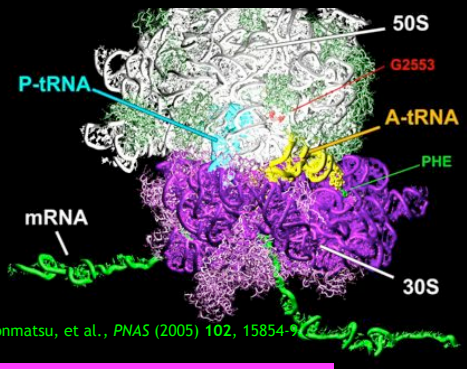


Blanchard, Cornell
Steve Chu, UCB

Simulation Set-up: accommodation

- Explicit Solvent
- Particle-mesh Ewald electrostatics
- NAMD scalable MD code
- AMBER force field
- 1.6 ns equilibration time
- 22 ns production (new runs 500 ns)
- 2.64 x 10^6 atoms
- Outstanding dynamic load balancing

Explicit solvent accommodation simulations:



Sanbonmatsu, et al., *PNAS* (2005) **102**, 15854-9

Water and ions (0.1 M KCl; 7mM $MgCl_2$) not shown.

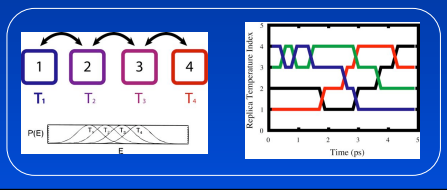
Replica method produces enhanced sampling

(Sugita and Okamoto, 1999; Garcia and Sanbonmatsu PNAS 2002)

- N replicas are simulated in parallel at different temperatures
- Replicas are allowed to swap temperatures providing thermal 'kick'

$$P(\text{exchange}) = \exp\left[\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)(E_i - E_j)\right]$$

- A 48 replica simulation with 15 μs total sampling (312 ns/replica) samples more than a 15 μs standard MD simulation (Sanbonmatsu and Garcia PSFG 2002).
- Estimates range between 25-75 fold increase in sampling (conservative estimate: 15 μs total sampling ~ 0.375 ms effective sampling).



Entropy shuttling facillitates gentamicin dissociation from the ribosome

Andrea C. Vaiana and Kevin Y. Sanbonmatsu

Los Alamos National Lab
T-10 Theoretical Biology and Biophysics

Replica Exchange Molecular Dynamics Simulation

Total sampling: 15 microseconds

LANL Coyote Supercomputer

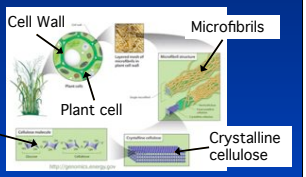
Bioenergy applications: cellulosic ethanol

Bioenergy

- Idea: produce ethanol from simple sugars via fermentation.
- Sugar-cane ethanol: requires tropical climate, fertile soil
- Corn-based ethanol: use enzyme to convert starch to sugar. Not sufficient, increases food prices.
- Cellulosic ethanol: recycles agricultural waste; can use sawdust, woodchips, switchgrass (grown on wastelands).
- Cellulosic ethanol: potential to satisfy 30% of transportation fuel demand.

Cellulose degradation is a bottleneck in ethanol production

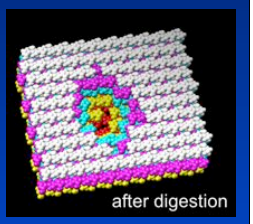
Ethanol production:
1. Pre-treatment
2. Cellulose degradation by cellulase enzymes
3. Fermentation



Cellulose
- Resides in plant cell walls
- Extremely tough, resisting treatment by acid and steam explosion.
- Exists in the form of 2-D crystalline sheets.

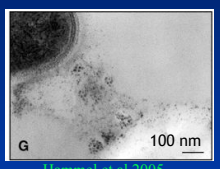
Cellulose

- Cellulose exists in the form of stacked layers of two-dimension crystalline sheets.
- Each sheet consists of long polysaccharide chains connected in a lattice by hydrogen bonds .
- A pre-treatment step is necessary to make the cellulose susceptible to breakdown by the cellulosome.

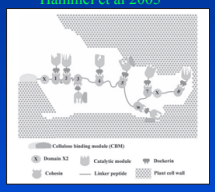


The Cellulosome

- Bacteria have evolved extremely efficient ways of degrading cellulose.
- The "Cellulosome" is a molecular machine that degrades cellulose.
- Acts like molecular paper shredder.
- "Pac-men" subunits degrade single strands of cellulose.
- The mechanism is poorly understood.
- Idea: make designer cellulosomes with customized subunits.



Hammel et al 2005

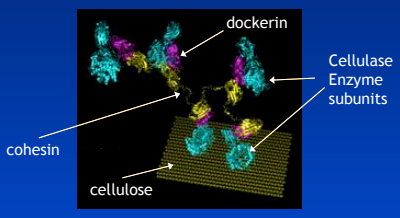


Bayer et al 1998

Simulation Set-up

- Simulate movement of cellulose strand through cellulosome subunits
- Steered MD (restrain end of cellulose chain, apply force on c-o-m of subunits.
- Particle-mesh Ewald electrostatics
- GROMACS code
- AMBER force field

Cellulosome model



Los Alamos RoadRunner

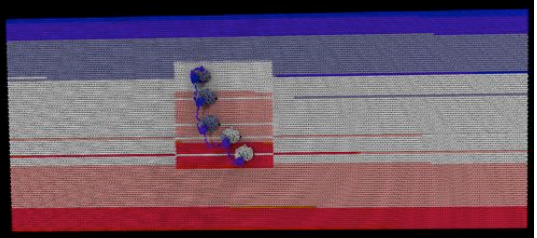
- New "Hybrid" architecture based on SONY PlayStation 3 "Cell" chip
- Cell has 7 cores (1 PPU, 6 SPUs) - 200 Gflops per cell
- >2x faster than BG/L LLNL
- 12,960 cells, 6,948 dual-core AMD, 80 terabytes RAM (2cell,2 dual amd/node).



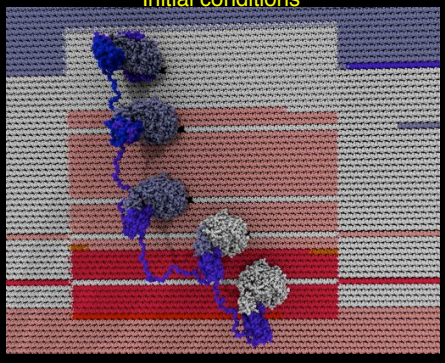
Porting to RoadRunner

- Gromacs modified: IBM DaCS libraries for nonbonded calculations on the cell processors.
- Other modifications: launching the cell processes, aligned memory buffers, demand DMA transfers, and the port of the water-water nonbonded kernel on the cell broadband accelerator.

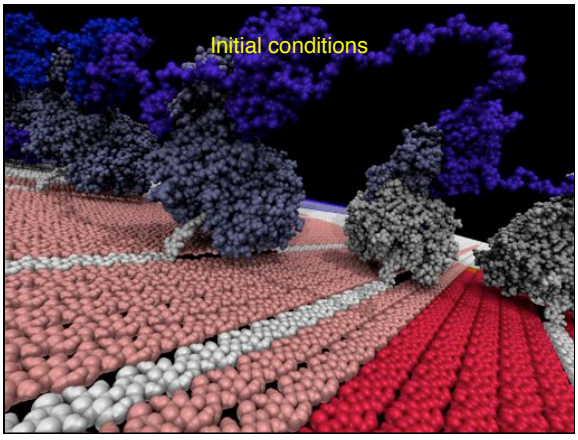
Initial conditions



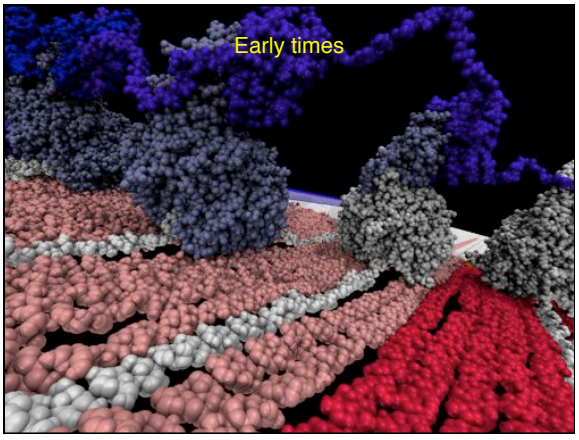
Initial conditions

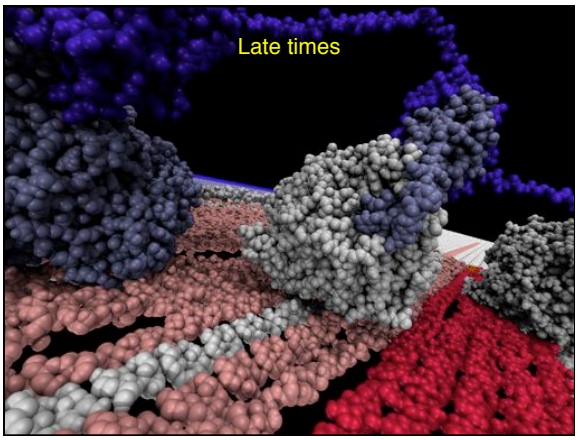


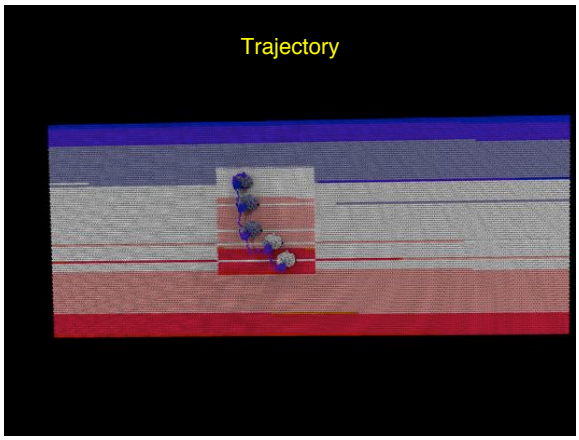
Initial conditions



Early times







Conclusions

- Biomolecular simulations require time scale range of 15 orders of magnitude
- Simulations of ribosome uncover potential antibiotic targets
- Simulating movement of cellulosome through cellulose during degradation.

kys@lanl.gov www.t10.lanl.gov/kys

Acknowledgements

Sanbonmatsu Team

Paul Whitford
Andrea Vaiana
Scott Hennelly
Yanan Yu
Chang-Shung Tung

Financial Support

NIH
LANL LDRD

Lab space

Cliff Unkefer (LANL)

RoadRunner Bioenergy Team

Mark Vernon
Christine Ahrens
Matt Sheats
Joel Berendzen
Sriram Swaminarayan

Computing Resources

LANL Inst. Computing
(Andy White, Ken Koch)

Collaborators

Scott Blanchard (Cornell), Jamie Cate (UCB), Jose Onuchic (UCSD)