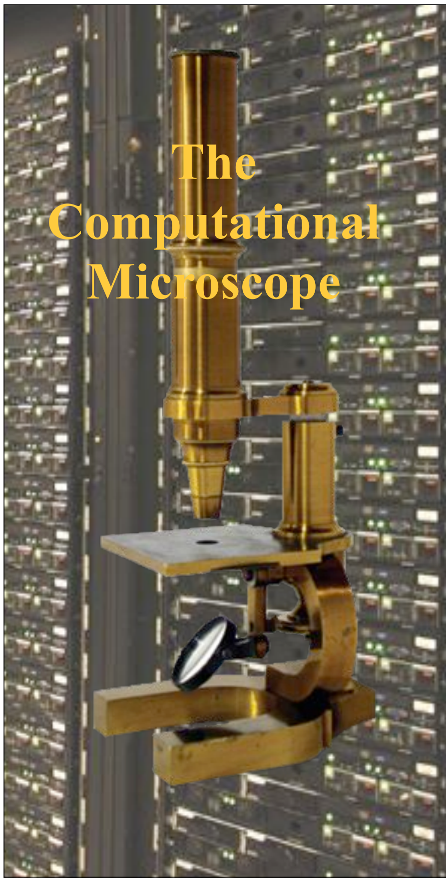
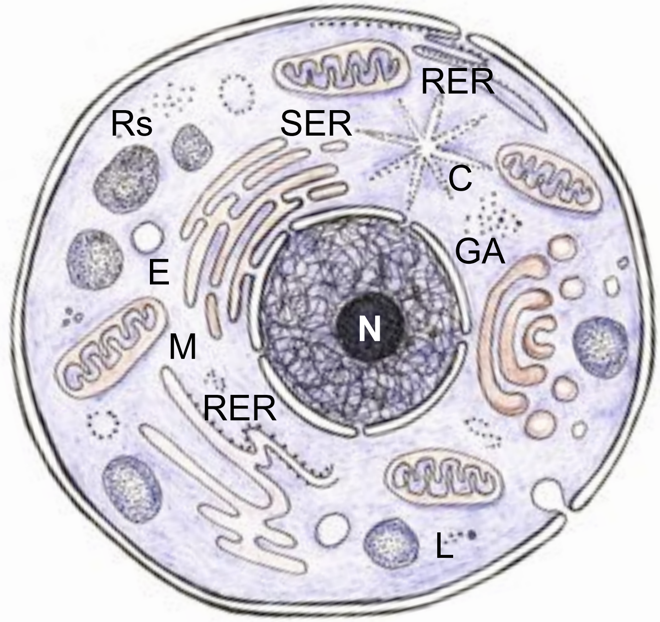


Computational microscope views at atomic resolution ...

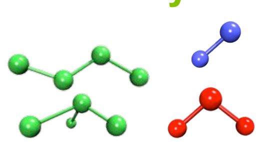


... how living cells maintain health and battle disease

John Stone

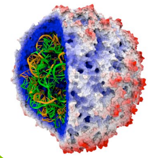
Our Microscope is Made of...

Chemistry

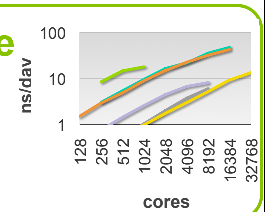


$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} + \underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

NAMD Software



Virus



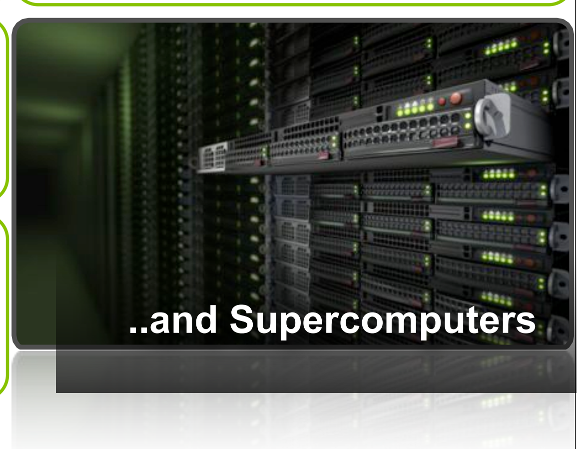
Physics

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla}U(\vec{R})$$

Math

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i}\vec{F}_i(t)$$

(repeat ***one billion times*** = microsecond)



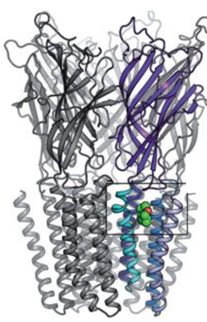
NAMD impact is broad and deep

- Comprehensive, industrial-quality software
  - Integrated with VMD for simulation setup and analysis
  - Portable extensibility through Tcl scripts (also used in VMD)
  - Consistent user experience from laptop to supercomputer
- Large user base – 51,000 registered users
  - 9,100 (18%) are NIH-funded; many in other countries
  - 14,100 have downloaded more than one version
- Leading-edge simulations
  - "most-used software" on NICS Cray XT5 (largest NSF machine)
  - "by far the most used MD package" at TACC (2nd and 3rd largest)
  - NCSA Blue Waters early science projects and acceptance test
  - Argonne Blue Gene/Q early science project

Outside researchers choose NAMD and succeed

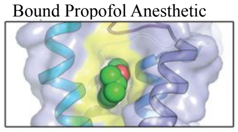
2100 external citations since 2007

Corringer, et al., *Nature*, 2011



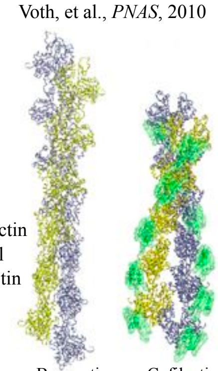
180K-atom 30 ns study of anesthetic binding to bacterial ligand-gated ion channel provided "complementary interpretations…that could not have been deduced from the static structure alone."

Bound Propofol Anesthetic



Voth, et al., *PNAS*, 2010

500K-atom 500 ns investigation of effect of actin depolymerization factor/cofilin on mechanical properties and conformational dynamics of actin filament.

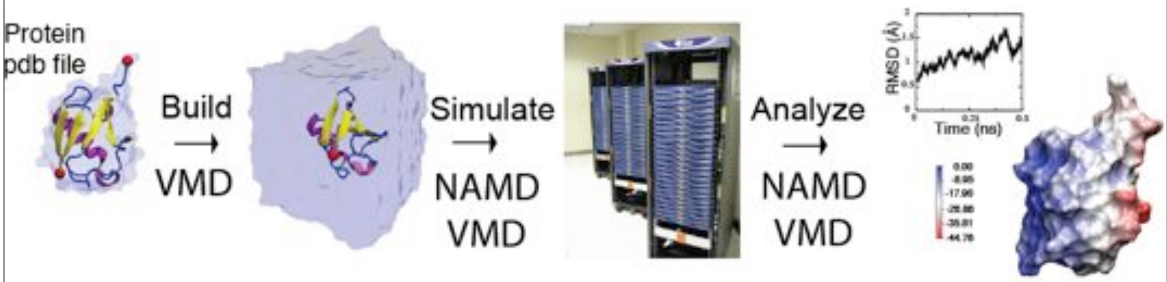


Bare actin   Cofilactin

Recent NAMD Simulations in *Nature*

- **M. Koeksal, et al.,** *Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis*. (2011)
- **C.-C. Su, et al.**, *Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli.* (2011)
- **D. Slade, et al.**, *The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase*. (2011)
- **F. Rose, et al.,** *Mechanism of copper(II)-induced misfolding of Parkinson's disease protein*. (2011)
- **L. G. Cuello, et al.**, *Structural basis for the coupling between activation and inactivation gates in K(+) channels*. (2010)
- **S. Dang, et al.,**, *Structure of a fucose transporter in an outward-open conformation.* (2010)
- **F. Long, et al.**, *Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport*. (2010)
- **R. H. P. Law, et al.**, *The structural basis for membrane binding and pore formation by lymphocyte perforin.* (2010)
- **P. Dalhaimer and T. D. Pollard**, *Molecular Dynamics Simulations of Arp2/3 Complex Activation.* (2010)
- **J. A. Tainer, et al.**, *Recognition of the Ring-Opened State of Proliferating Cell Nuclear Antigen by Replication Factor C Promotes Eukaryotic Clamp-Loading*. (2010)
- **D. Krepkiy, et al.,**, *Structure and hydration of membranes embedded with voltage-sensing domains*. (2009)
- **N. Yeung, et al.,**, *Rational design of a structural and functional nitric oxide reductase*. (2009)
- **Z. Xia, et al.**, *Recognition Mechanism of siRNA by Viral p19 Suppressor of RNA Silencing: A Molecular Dynamics Study*. (2009)

The Molecular Dynamics Simulation Process



For textbooks see:

M.P. Allen and D.J. Tildesley. ***Computer Simulation of Liquids.***Oxford University Press, New York, 1987.
D. Frenkel and B. Smit. ***Understanding Molecular Simulations. From Algorithms to Applications.*** Academic Press, San Diego, California, 1996.
A. R. Leach. ***Molecular Modelling. Principles and Applications.***Addison Wesley Longman, Essex, England, 1996.
More at http://www.biomath.nyu.edu/index/course/99/textbooks.html

Classical Dynamics

at 300K

Energy function: $U(\vec{r}_1, \vec{r}_2, \cdots \vec{r}_N) = U(\vec{R})$

used to determine the force on each atom:

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})$$

yields a set of 3N coupled 2^{nd}-order differential equations that can be propagated forward (or backward) in time.

Initial coordinates obtained from crystal structure, velocities taken at random from Boltzmann distribution.

Maintain appropriate temperature by adjusting velocities.

Classical Dynamics

discretization in time for computing

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})$$

Use positions and accelerations at time t and the positions from time t-δt to calculate new positions at time t+δt.

$$\mathbf{r}(t+\delta t) \approx \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2$$

$$+$$

$$\mathbf{r}(t-\delta t) \approx \mathbf{r}(t) - \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2$$

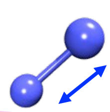
"Verlet algorithm"

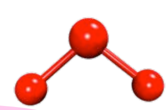
$$-\vec{\nabla} U(\vec{R}) / m_i$$

$$\boxed{\mathbf{r}(t+\delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \mathbf{a}(t)\delta t^2}$$

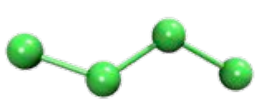
Potential Energy Function of Biopolymer

- Simple, fixed algebraic form for every type of interaction.
- Variable parameters depend on types of atoms involved.





heuristic

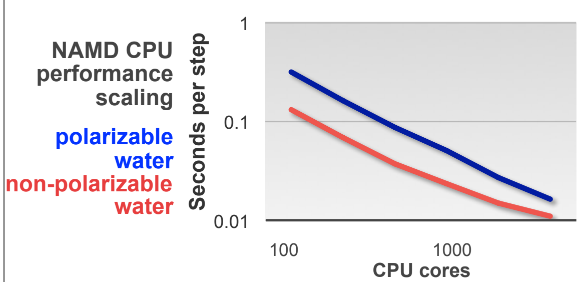


$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} + \underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

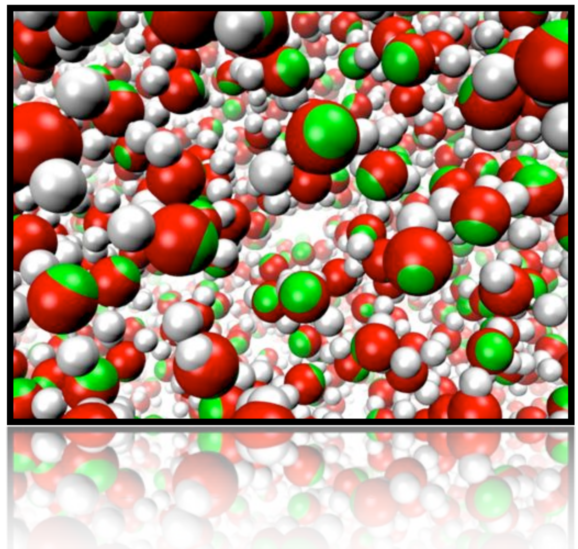
from physics

Improving the Force Field

- Atomic polarizability increases computation by 2x…
- …but, the additional computations are perfectly suited to the GPU!
- For now, NAMD calculates atomic polarizability on CPUs only...soon we will also use GPUs



Atomic polarizability of water, highly accurately simulated through additional particles (shown in green)



Molecular Dynamics Ensembles

Constant energy, constant number of particles (NE)

Constant energy, constant volume (NVE)

Constant temperature, constant volume (NVT)

Constant temperature, constant pressure (NPT)

Choose the ensemble that best fits your system and start the simulations, but use NE to check on accuracy of the simulation.

Langevin Dynamics

for temperature control

Langevin dynamics deals with each atom separately, balancing a small friction term with Gaussian noise to control temperature:

$$m\ddot{\vec{r}} \;=\; \vec{F}(\vec{r}) \;-\; \gamma\, m\,\dot{\vec{r}} \;+\; \vec{R}(t)$$

$$\langle \vec{R}(t)\cdot\vec{R}(t')\rangle \;=\; 6k_BT\gamma\,\delta(t\;-\;t')$$

Langevin Dynamics

for pressure control

Underlying Langevin-Hoover barostat equation for all atoms: Equations solved numerically in NAMD

$$\frac{d^2V(t)}{dt^2} \;=\; \frac{1}{W_{bs}}\,[P(t)\;-\;P_{\rm target}]\;-\;\frac{1}{\tau_{\rm bs}}\frac{dV(t)}{dt}\;+\;R_{\rm bs}(t)$$

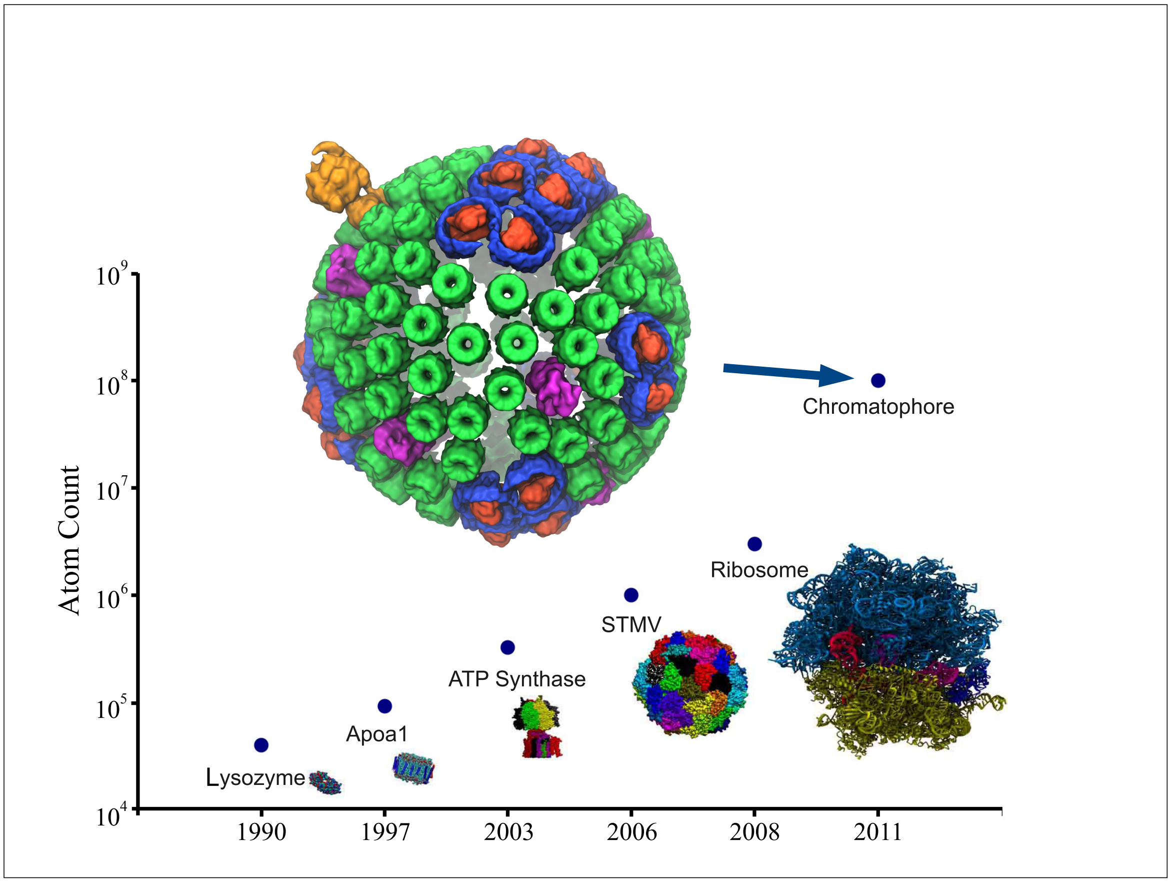
$$P = \rho k_B T + \frac{1}{Vd}\sum_{i<j}\langle r_{ij}\frac{dU_{tot}(r_{ij})}{dr_{ij}}\rangle \qquad d \;=\; \text{dimension}$$

$$\langle R_{\rm bs}(t)\,R_{\rm bs}(t')\rangle \;=\; \frac{2\,k_{\rm B}\,T_{\rm target}\;\delta(t\;-\;t')}{W_{\rm bs}\,\tau_{\rm bs}} \qquad W_{\rm bs} \;=\; d\,N_{\rm atoms}\,k_{\rm B}T_{\rm target}\,\tau^2_{period}$$

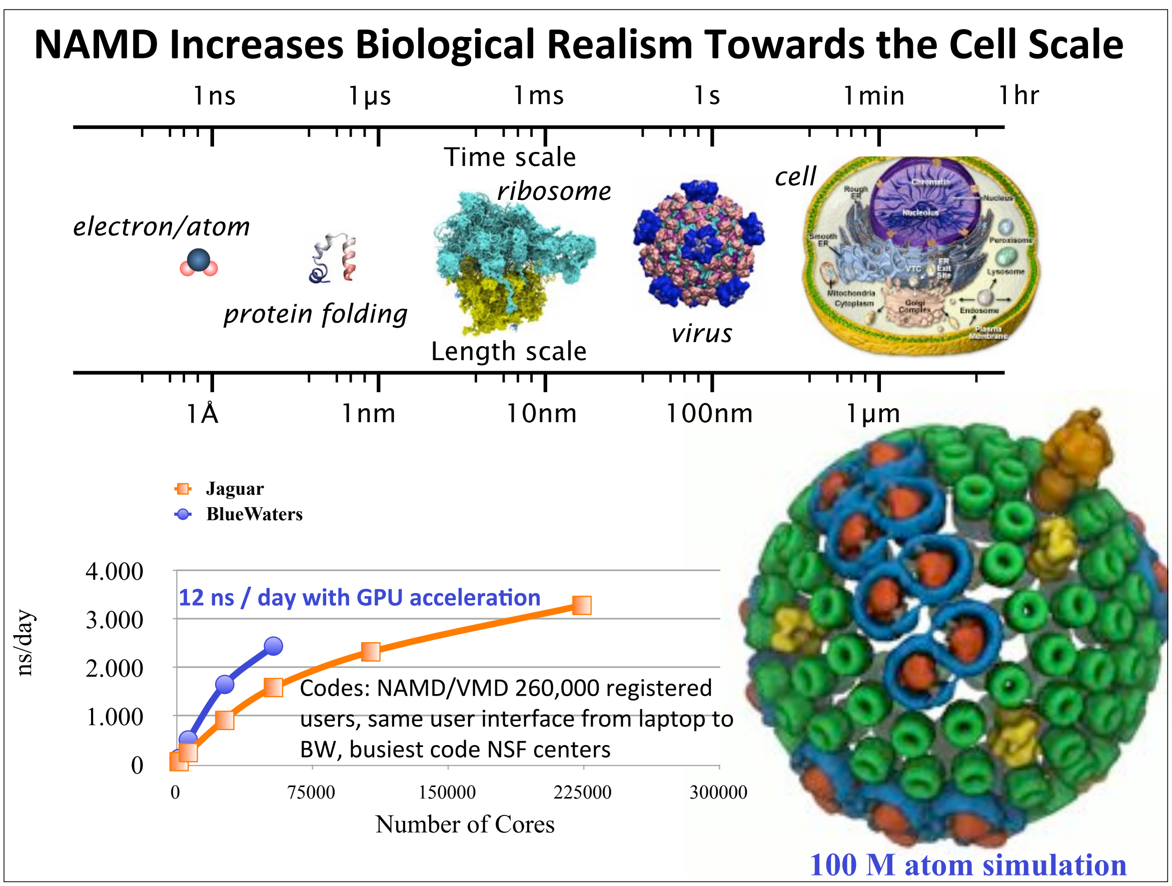
$$\dot{\mathbf{r}}_i = \mathbf{v}_i + s\mathbf{r}_i \qquad \dot{\mathbf{v}}_i = \mathbf{F}_i\,/\,m_i - s\mathbf{v}_i$$

$$\dot{V} = dVs \qquad \dot{s} = dV(P - P_{\rm target})\,/\,W - s\,/\,\tau_{\rm bs} \quad + R(t)$$

d - dimension

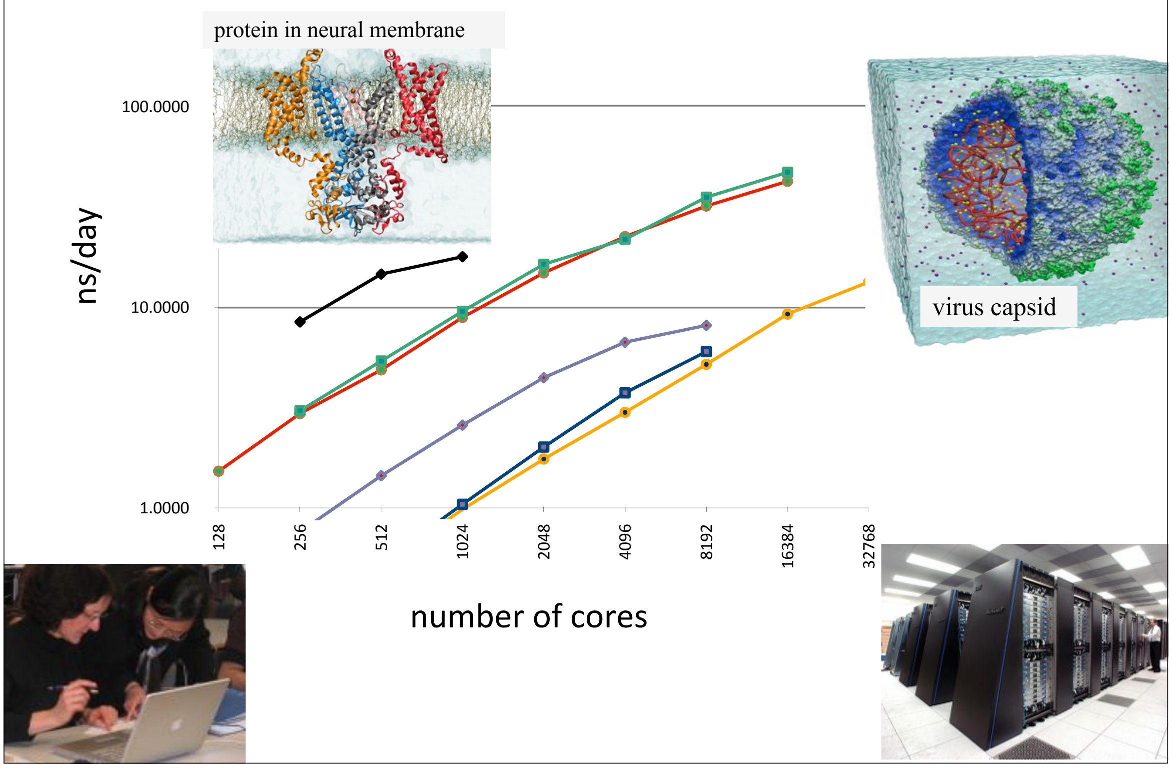


Atom Count
10^9
10^8
10^7
10^6
10^5
10^4
1990
1997
2003
2006
2008
2011
Lysozyme
Apoa1
ATP Synthase
STMV
Ribosome
Chromatophore



NAMD Increases Biological Realism Towards the Cell Scale
1ns
1µs
1ms
1s
1min
1hr
Time scale
electron/atom
protein folding
ribosome
virus
cell
Length scale
1Å
1nm
10nm
100nm
1µm
Jaguar
BlueWaters
ns/day
4.000
3.000
2.000
1.000
0
12 ns / day with GPU acceleration
Codes: NAMD/VMD 260,000 registered users, same user interface from laptop to BW, busiest code NSF centers
0
75000
150000
225000
300000
Number of Cores
100 M atom simulation

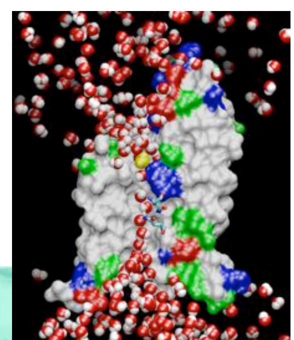
NAMD Scalability



From 10,000 to 100,000 Atom MD in 2000

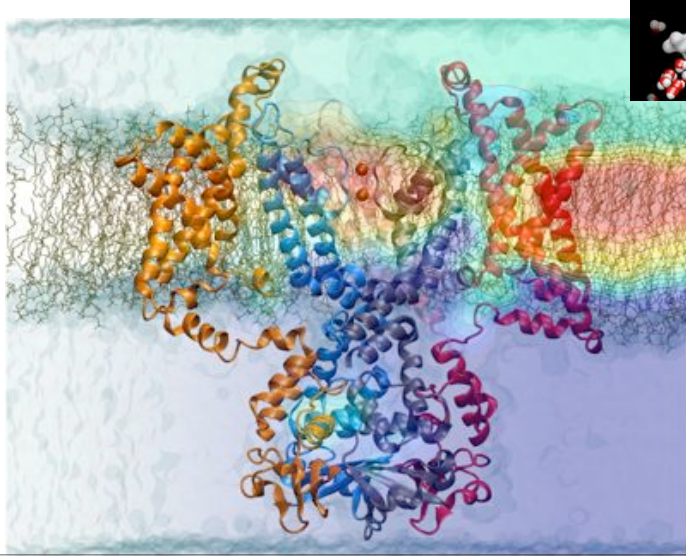
100k atom MD reached in 2000

- then a factor 10 increase in computation;
- **needed to describe membrane processes;**
- was achieved through cluster computing;
- produced good quality results for aquaporin;
- **is now standard.**



E.Tajkhorshid, P. Nollert, M. Jensen, L. Miercke, J. O'Connell, and K. Schulten. *Science*, **296:**525-530, 2002.

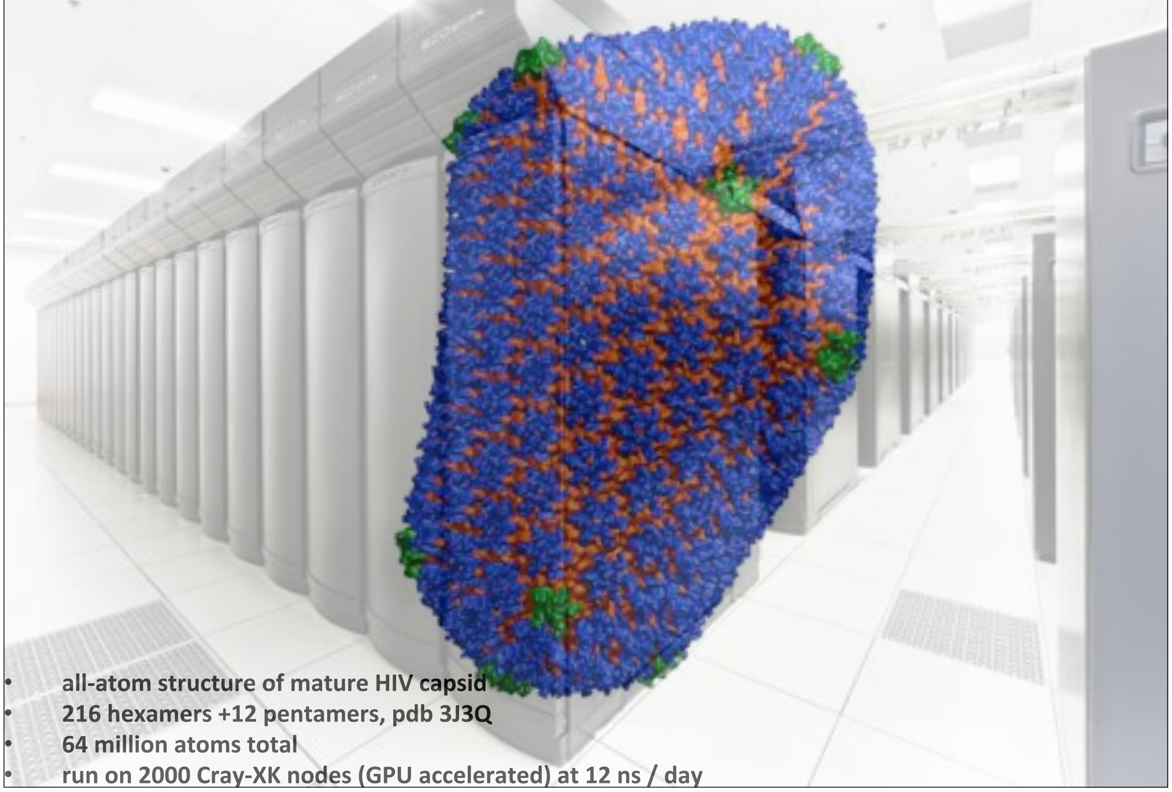
100,000 atoms, 12 ns



F. Khalili-Araghi, V. Jogini, V. Yarov-Yarovoy, E. Tajkhorshid, B. Roux, and K. Schulten. **Calculation of the gating charge for the Kv1.2 voltage-activated potassium channel.** *Biophysical Journal*, **98:**2189-2198, 2010.

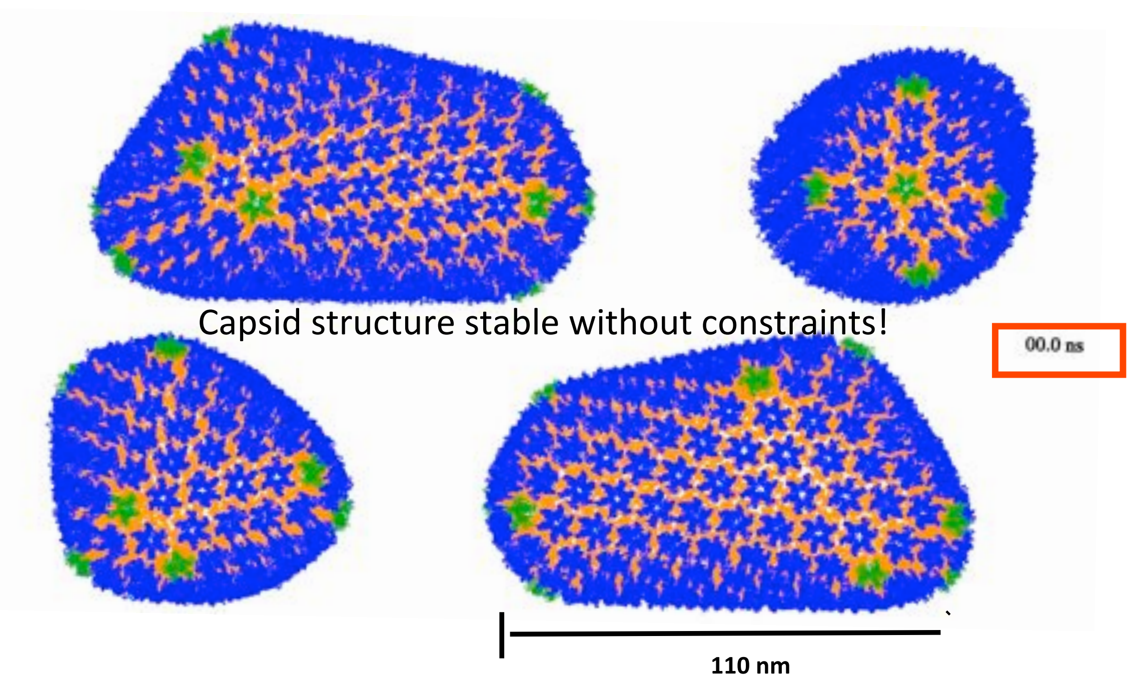
350,000 atoms, 0.5 µs

From 100,000 to 64,000,000 Atom MD Now



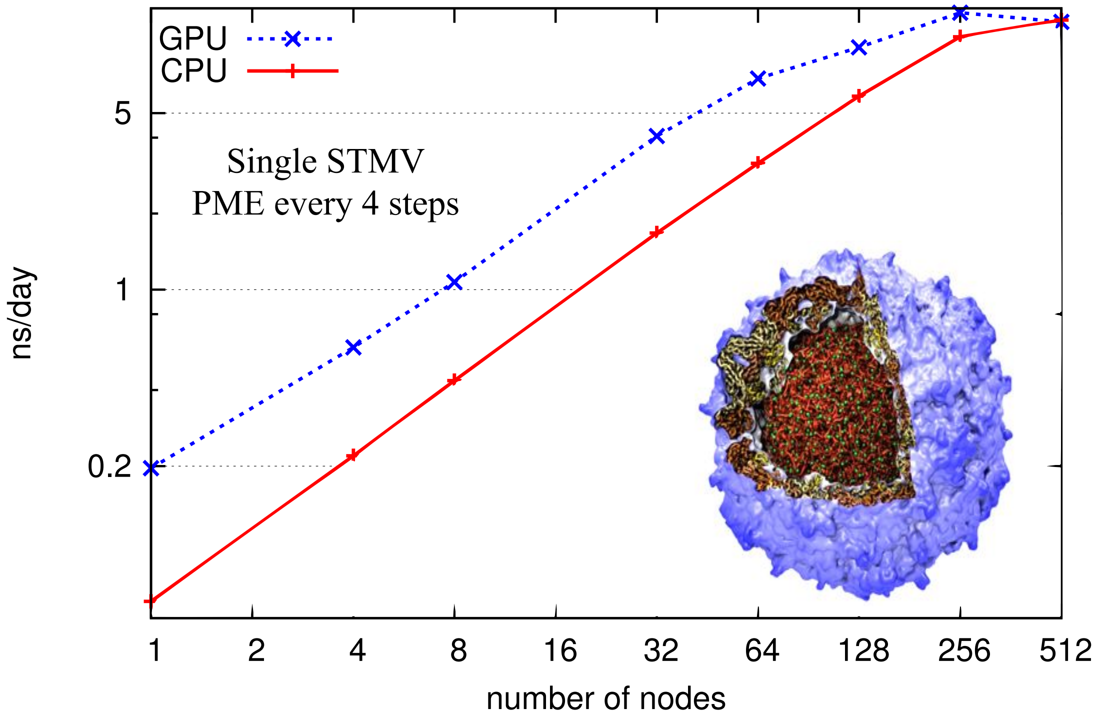
- all-atom structure of mature HIV capsid
- 216 hexamers +12 pentamers, pdb 3J3Q
- 64 million atoms total
- run on 2000 Cray-XK nodes (GPU accelerated) at 12 ns / day

All-atom MD Simulation of HIV-1 Capsid

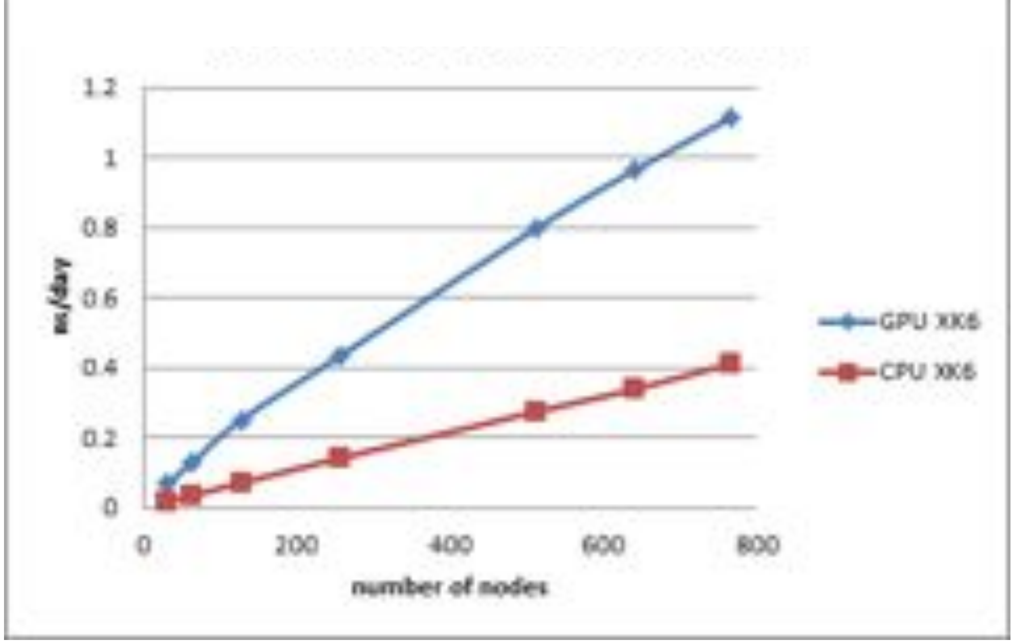


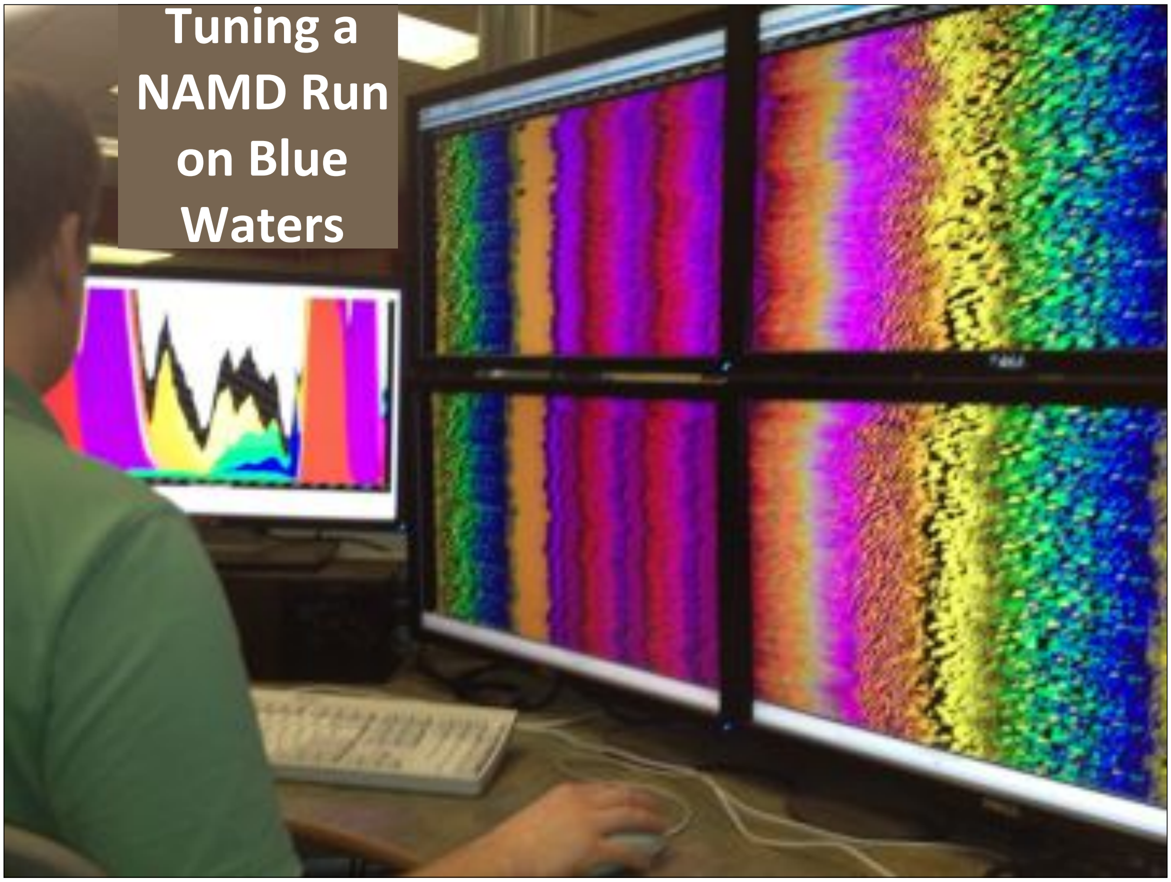
- 216 hexamers +12 pentamers, pdb 3J3Q (available May 29)
- 64 million atoms total
- Over 100 ns of MD on NSF Blue Waters – 5000 Nodes, 160,000 cores - 10ns/day

1M Atom Virus on TitanDev GPU

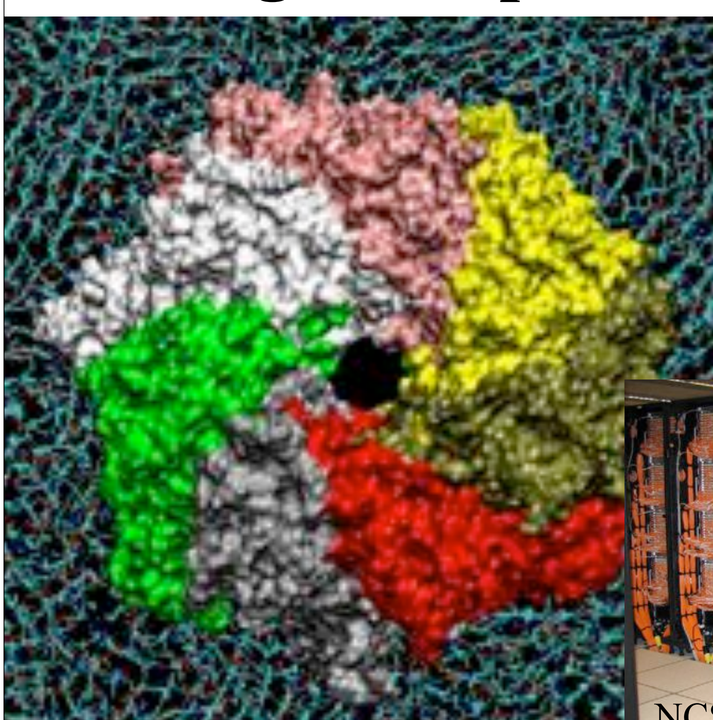


100M Atoms on TitanDev





Large is no problem. But …



Molecular dynamics simulation of alpha-hemolysin with about 300,000 atoms; 1 million atom simulations are routine today, 20 million atom simulations are possible.

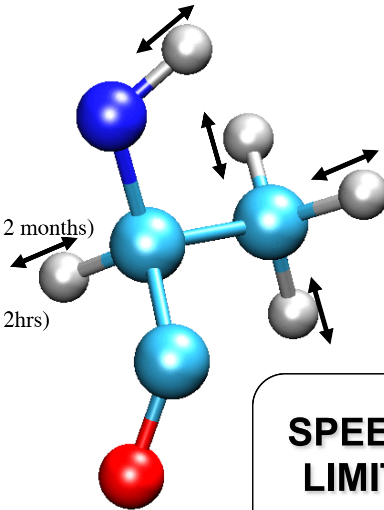
NCSA machine room

But long is still a problem!

biomolecular timescale and timestep limits

		steps
Rotation of buried sidechains; Local denaturations; Allosteric transitions (s – ms)	s	10^{15}
	ms	10^{12} (30 years, 2 months)
small protein folding	μs	10^{9} (10 days, 2hrs)
Hinge bending (μs – ps)	ns	10^{6} (15 min)
Rotation of surface sidechains; Elastic vibrations	ps	10^{3}
Bond stretching		
Molecular dynamics timestep →	fs	10^{0}

(NSF center, Shaw Res.)



SPEED LIMIT

δt = 1 fs