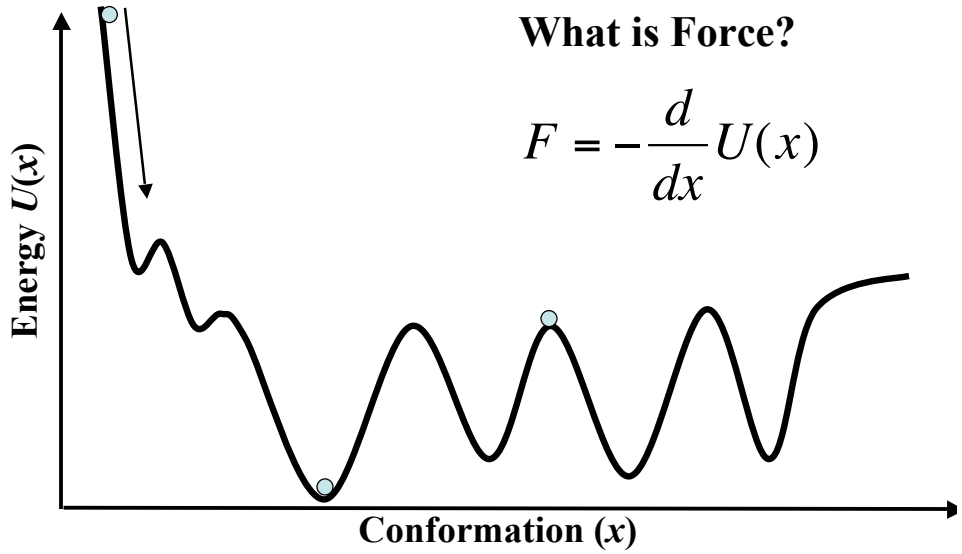
Don’t forget to bring your MD tutorial

Lab session starts at 1pm

You will have to finish an
MD/SMD exercise on α-conotoxin
in oxidized and reduced forms

Potential Energy (hyper)Surface



Classical Molecular Dynamics

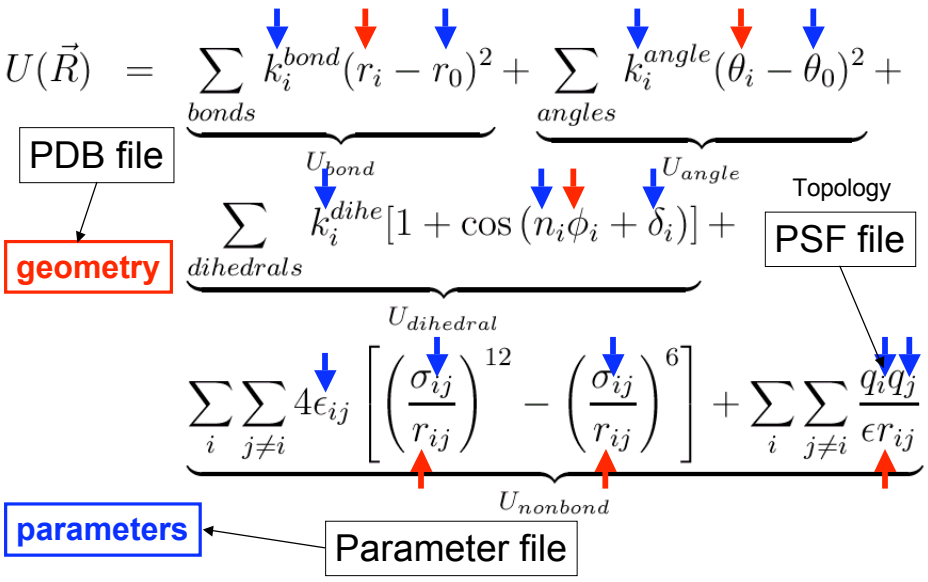
$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t)\delta t$$

$$\boldsymbol{v}(t+\delta t) = \boldsymbol{v}(t) + \boldsymbol{a}(t)\delta t$$

$$\boldsymbol{a}(t) = \boldsymbol{F}(t)/m$$

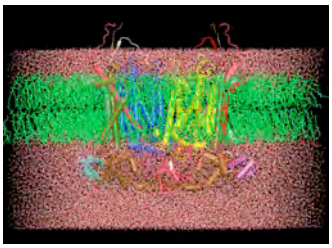
$$\boldsymbol{F} = -\frac{d}{d\boldsymbol{r}}U(\boldsymbol{r})$$

CHARMM Potential Function



Preparing Your System for MD
Solvation

Biological activity is the result of interactions between molecules and occurs at the interfaces between molecules (protein-protein, protein-DNA, protein-solvent, DNA-solvent, etc).



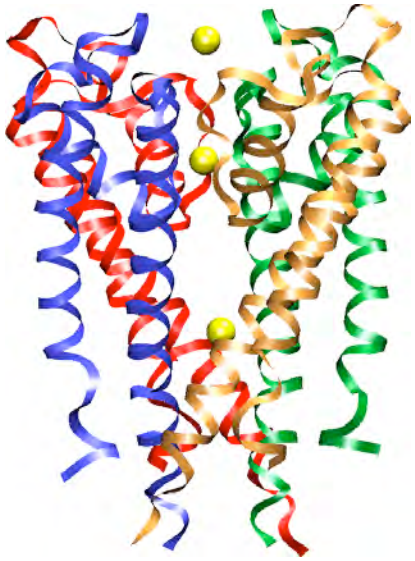
Why model solvation?
- many biological processes occur in aqueous solution
- solvation effects play a crucial role in determining molecular conformation, electronic properties, binding energies, etc

How to model solvation?
- explicit treatment: solvent molecules are added to the molecular system
- implicit treatment: solvent is modeled as a continuum dielectric

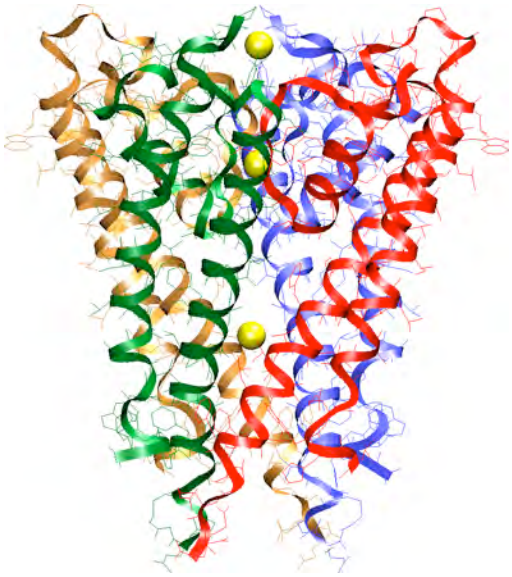
Example: MD Simulations of the K^+ Channel Protein

Ion channels are membrane - spanning proteins that form a pathway for the flux of inorganic ions across cell membranes.

Potassium channels are a particularly interesting class of ion channels, managing to distinguish with impressive fidelity between K^+ and Na^+ ions while maintaining a very high throughput of K^+ ions when gated.

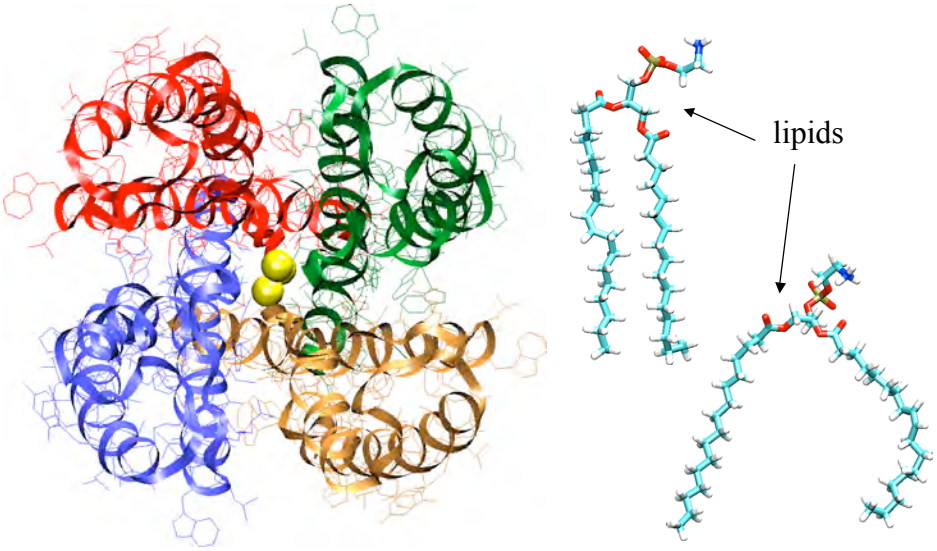


Setting up the system (1)



- retrieve the PDB (coordinates) file from the Protein Data Bank
- add hydrogen atoms using PSFGEN
- use topology and parameter files to set up the structure
- minimize the protein structure using NAMD2

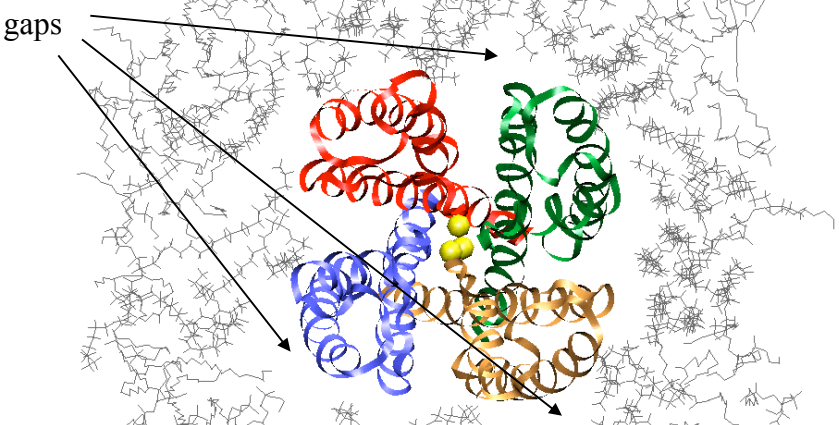
Setting up the system (2)



Simulate the protein in its natural environment: solvated lipid bilayer

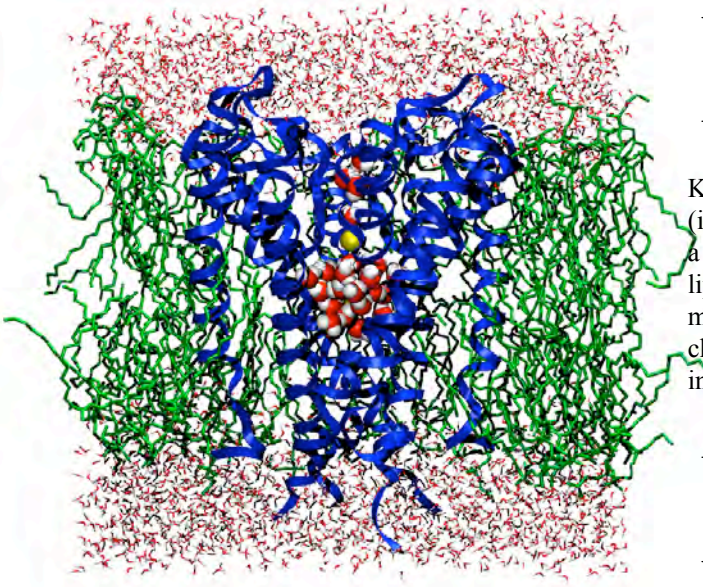
Setting up the system (3)

Inserting the protein in the lipid bilayer



Automatic insertion into the lipid bilayer leads to big gaps between the protein and the membrane => long equilibration time required to fill the gaps.
Solution: manually adjust the position of lipids around the protein

The system

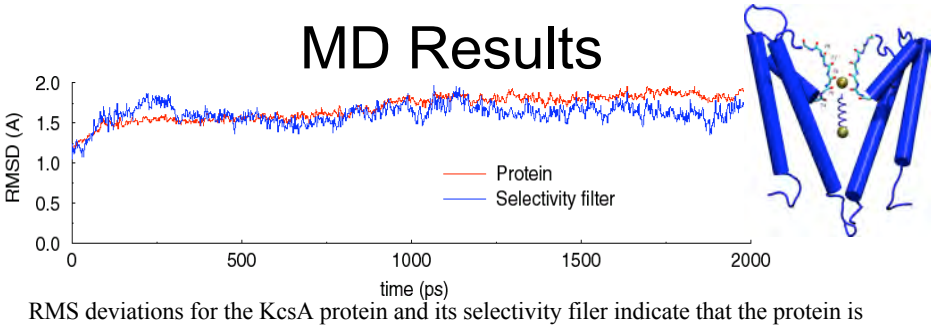


solvent

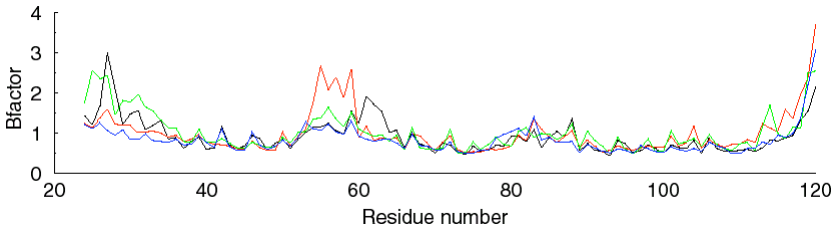
Kcsa channel protein (in blue) embedded in a (3:1) POPE/POPG lipid bilayer. Water molecules inside the channel are shown in vdW representation.

solvent

MD Results



RMS deviations for the KcsA protein and its selectivity filer indicate that the protein is stable during the simulation with the selectivity filter the most stable part of the system.



Temperature factors for individual residues in the four monomers of the KcsA channel protein indicate that the most flexible parts of the protein are the N and C terminal ends, residues 52-60 and residues 84-90. Residues 74-80 in the selectivity filter have low temperature factors and are very stable during the simulation.

Simulating the system: Free MD

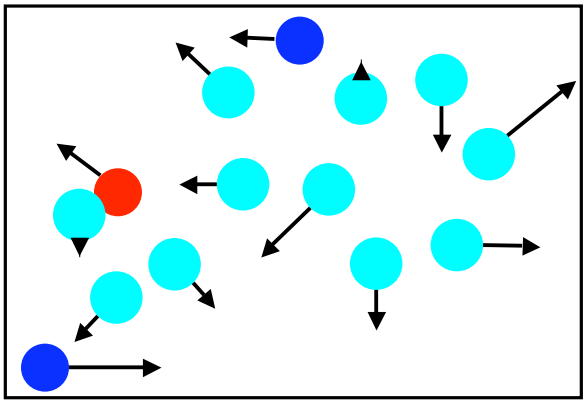
Summary of simulations:

- protein/membrane system contains 38,112 atoms, including 5117 water molecules, 100 POPE and 34 POPG lipids, plus K^+ counterions
- CHARMM26 forcefield
- periodic boundary conditions, PME electrostatics
- 1 ns equilibration at 310K, NpT
- 2 ns dynamics, NpT

Program: NAMD2

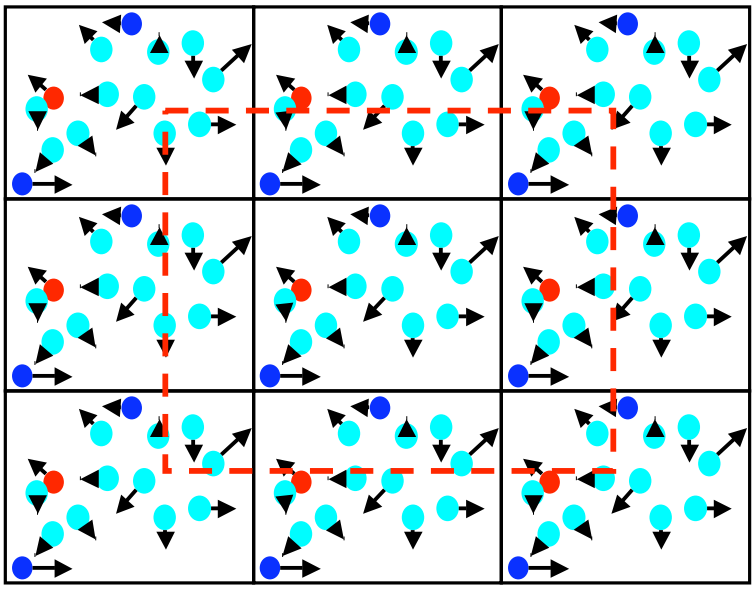
Platform: Cray T3E (Pittsburgh Supercomputer Center)

Boundary Conditions?



What happens if you put water under vacuum!?
Problems: Density, pressure, boundary effects, …
One solution: reflective boundaries, not quite good.

Periodic Boundary Conditions

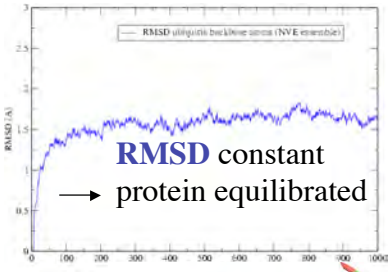


Equilibrium Properties of Proteins

Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

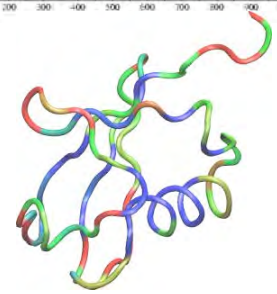
$$RMSD_\alpha = \sqrt{\frac{\sum_{j=1}^{N_t}\sum_{\alpha=1}^{N_\alpha}(\vec{r}_\alpha(t_j)-\langle\vec{r}_\alpha\rangle)^2}{N_\alpha}}$$



RMSD constant
→ protein equilibrated

NMR structures
aligned together to see flexibility

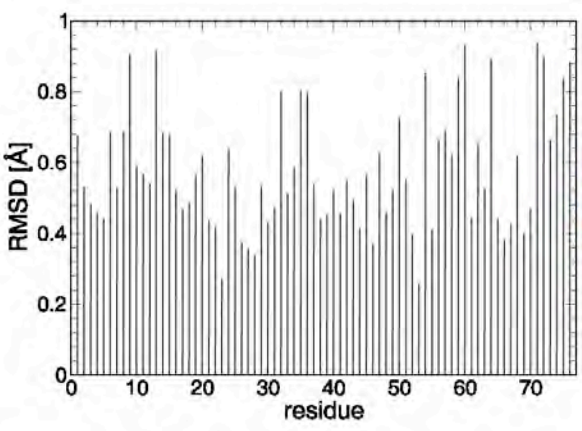
Protein sequence exhibits characteristic permanent flexibility!



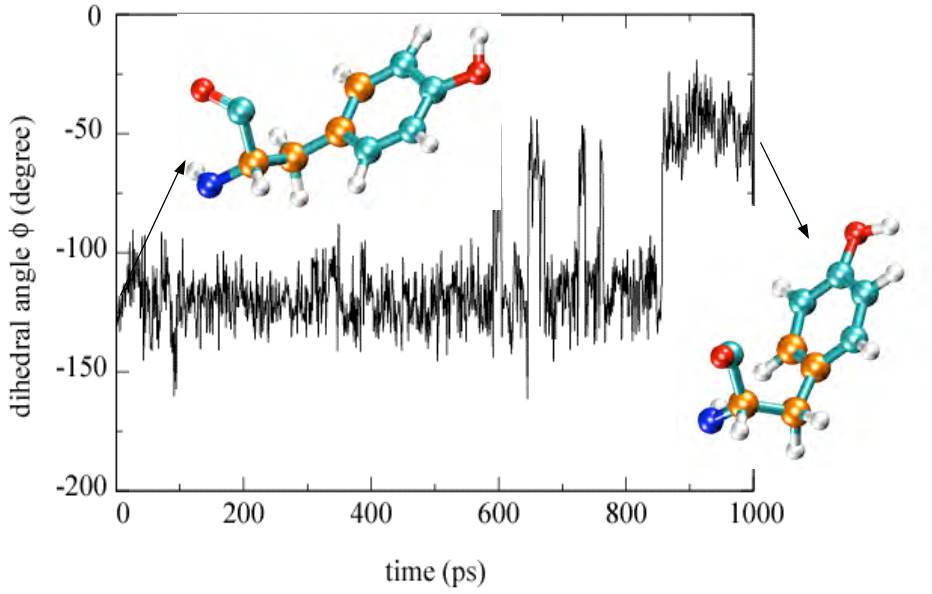
MD simulation
The color represents mobility of the protein through simulation (red = more flexible)

Thermal Motion of Ubiquitin from MD

RMSD values per residue

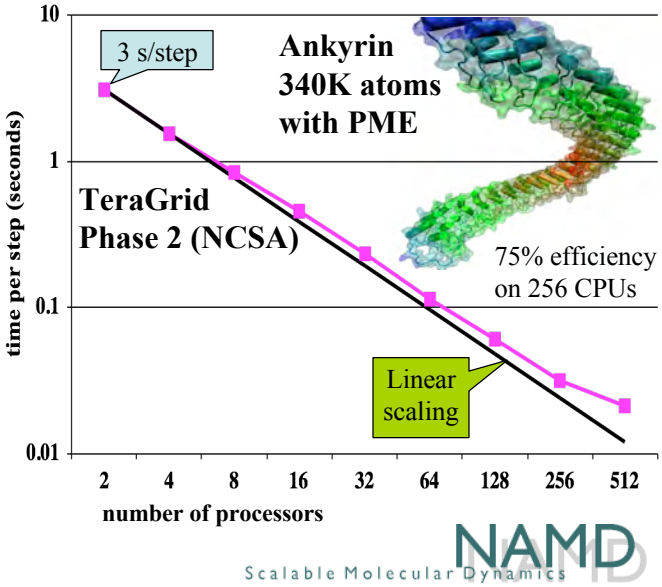


Patience is required to observe Molecular Events



NAMD: The Program we will Use

J. Phillips Ph.D.



Simulation of large biomolecular systems

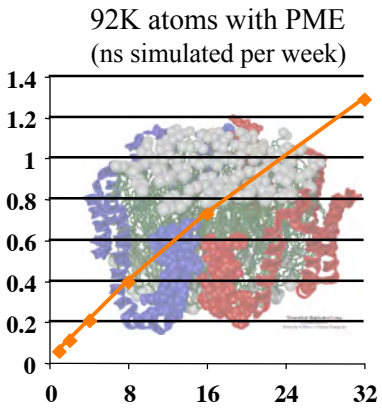
Runs at NSF centers, on clusters, and on desktop.

Available for **FREE** as precompiled binaries; includes source code.

NAMD
Scalable Molecular Dynamics

Linux Clusters

parallel computing on a professor's salary



The NAMD Configuration File / 1

Files needed:

```
structure          mypsf.psf
coordinates        mypdb.pdb
```

Define temperature

```
set temperature    310
     ;# target temperature used several times below
```

Starting simulation with random velocities

```
# starting from scratch
temperature         $temperature
       ;# initialize velocities randomly
```

The NAMD Configuration File / 2

Continuing a simulation with positions and velocities from previous run

```
# continuing a run
set inputname      myinput          ;# only need to edit this in one place!
binCoordinates     $inputname.coor  ;# coordinates from last run (binary)
binVelocities      $inputname.vel   ;# velocities from last run (binary)
extendedSystem    $inputname.xsc   ;# cell dimensions from last run
firsttimestep      50000            ;# last step of previous run
numsteps           100000           ;# run stops when this step is reached
```

The NAMD Configuration File / 3

Organizing output

```
outputName          myoutput
        ;# base name for output from this run

restartfreq         500     ;# 500 steps = every 1ps
dcdfreq             500
xstFreq             500

outputEnergies      100     ;# 100 steps = every 0.2 ps
outputTiming       1000
          ;# shows time per step and time to completion
```

The NAMD Configuration File / 4

```
# Force-Field Parameters
paraTypeCharmm     on
parameters          par_all27_prot_lipid.inp

# These are specified by CHARMM
exclude             scaled1-4
1-4scaling          1.0
switching           on

# You have some freedom choosing the cutoff
cutoff              12. ;# may use smaller, maybe 10., with PME
switchdist          10. ;# cutoff - 2.

# Promise that atom won't move more than 2A in a cycle
pairlistdist        14. ;# cutoff + 2.
stepspercycle       10   ;# redo pairlists every ten steps
```

12A cutoff is official standard for CHARMM force field but smaller is OK when using full electrostatics

```
# Integrator Parameters
timestep            2.0  ;# 2fs/step
rigidBonds          all  ;# needed for 2fs steps
nonbondedFreq       1    ;# nonbonded forces every step
fullElectFrequency  2    ;# PME only every other step
```

Energy drifts if too large, but smaller requires more steps per ns.

The NAMD Configuration File / 5

Controlling temperature

γ

```
{\small \begin{verbatim}
# Constant Temperature Control
langevin            on              ;# langevin dynamics

langevinDamping     5.              ;# damping coefficient of 5/ps
langevinTemp        $temperature    ;# random noise at this level
langevinHydrogen    no              ;# don't couple bath to hydrogens
```

Underlying Langevin equation for all atoms

$$m_i \frac{d^2 x_i(t)}{dt^2} = F_{i,\mathrm{ff}} - \gamma m_i \frac{dx_i(t)}{dt} + R_i(t)$$

$$\langle R_i(t)\, R_i(t') \rangle \;=\; 2 k_B T_{\mathrm{target}} \gamma_i \, \delta(t - t')$$

The NAMD Configuration File / 6

Using periodic boundary conditions
(avoids surface effects; permits Particle-Mesh-Ewald (PME) electrostatics; permits pressure control)

```
# Periodic Boundary conditions
cellBasisVector1    31.2   0.   0.  ;# vector to the next image
cellBasisVector2     0.   44.8  0.
cellBasisVector3     0.    0   51.3
cellOrigin           0.    0.   0.  ;# the *center* of the cell

wrapWater           on              ;# wrap water to central cell
wrapAll             on              ;# wrap other molecules too
wrapNearest         off             ;# use for non-rectangular cells
```

The NAMD Configuration File / 7

Particle-Mesh-Ewald electrostatics
(avoids cut-off of long-range Coulomb forces)

```
#PME (for full-system periodic electrostatics)
PME                 yes
PMEGridSizeX        32  ;# 2^5, close to 31.2
PMEGridSizeY        45  ;# 3^2 * 5, close to 44.8
PMEGridSizeZ        54  ;# 2 * 3^3, close to 51.3
```

The NAMD Configuration File / 9

Fix atoms

```
fixedAtoms          on
fixedAtomsFile      myfixedatoms.pdb  ;# flags are in this file
fixedAtomsCol       B                 ;# set beta non-zero to fix an atom
```

The NAMD Configuration File / 10

Energy-minimize structure (T=0) , reset temperature T, run:

```
minimize            1000          ;# lower potential energy for 1000 steps
reinitvels          $temperature  ;# since minimization zeros velocities
run 50000 ;# 100ps
```

The NAMD Output File / 1

Preamble

```
Info: NAMD 2.5b2ss03 for Linux-i686-Clustermatic
Info:
Info: Please visit http://www.ks.uiuc.edu/Research/namd/
Info: and send feedback or bug reports to namd@ks.uiuc.edu
Info:
Info: Please cite Kale et al., J. Comp. Phys. 151:283-312 (1999)
Info: in all publications reporting results obtained with NAMD.
Info:
Info: Built Fri May 30 13:09:06 CDT 2003 by jim on umbriel
Info: Sending usage information to NAMD developers via UDP.
Info: Sent data is: 1 NAMD  2.5b2ss03  Linux-i686-Clustermatic 47 umbriel jim
Info: Running on 47 processors.
```

The NAMD Output File / 2

Energies

```
ETITLE:      TS           BOND          ANGLE          DIHED          IMPRP
          ELECT            VDW       BOUNDARY           MISC       KINETIC
          TOTAL           TEMP         TOTAL2         TOTAL3       TEMPAVG
       PRESSURE     GPRESSURE         VOLUME      PRESSAVG     GPRESSAVG

ENERGY:    1000         0.0000         0.0000         0.0000         0.0000
    -97022.1848      9595.3175         0.0000         0.0000     14319.5268
    -73107.3405       300.2464    -73076.6148    -73084.1411       297.7598
      -626.5205      -636.6638    240716.1374      -616.5673      -616.6619
```

The NAMD Output File / 1

Writing out trajectories

⋮

```
OPENING COORDINATE DCD FILE
WRITING COORDINATES TO DCD FILE AT STEP 1000
```

⋮

Performance information

```
Info: Benchmark time: 47 CPUs 0.0475851 s/step 0.275377 days/ns 13540 kB memory

TIMING: 1000  CPU: 18.35, 0.01831/step  Wall: 50.1581, 0.0499508/step, 6.92374
hours remaining, 14244 kB of memory in use.
```

Warnings

```
Warning: Pairlistdist is too small for 1 patches during timestep 17.
Warning: Pairlists partially disabled; reduced performance likely.
Warning: 20 pairlist warnings since previous energy output.
```

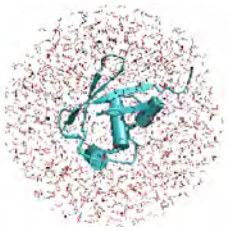
NAMD Example 1

You will first simulate ubiquitin in a water sphere and water box:

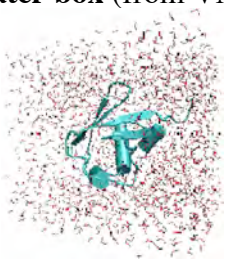
Generating a Protein Structure File (PSF)

- Go to 1-1-build directory
- Open VMD, choose extension TkCon
- Make from 1UBQ.pdb a structure without hydrogens, ubqp.pdb
- Create psf file for ubqp.pdb: ubq.pdb and ubq.psf
- Check if files exist

Solvate the protein in a water sphere (from VMD)



Solvate the protein in a water box (from VMD)



NAMD

- **RMSD value for equilibration**
- **Atomic RMSD values of equilibrated protein**
- **Velocity distribution**
- **Temperature distribution**

Inspection of the Results in VMD

- What to load?
- Some basic analysis tools
- Distances, angles, ….
- Salt bridges
- Fluctuations (RMSD) various types
- Using Tcl scripting to repeat the analysis for all frames loaded.

- Example of BPTI

Inspection of the Results in VMD

- Animating the results
- Smoothing the trajectory
- Making figures and movies
- Simultaneous display of multiple frames

Major Difficulties in Simulating Biological Systems

"Size" problem

Biological Systems are Complex

- Natural environment:
  Membrane, solvent, ions, …
- Many biosystems function in assemblies
  Photosynthetic apparatus
  Nuclear receptors, GPCRs, …

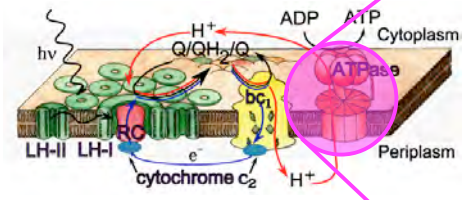
"Time scale" problem

Many biological events happen at μs–ms time scales

- Signaling and other regulatory mechanisms
- Protein folding

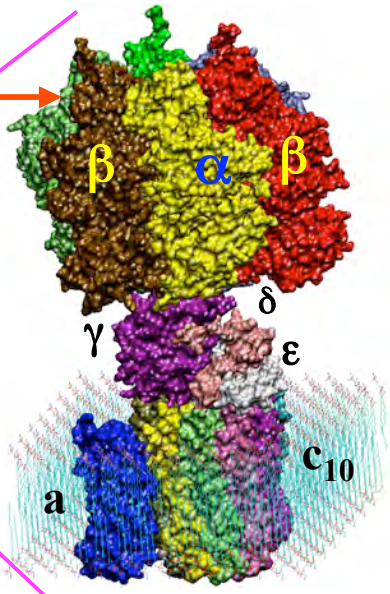
Large Systems

F_1 unit:
330,000 atoms



F_0 unit:
110,000 atoms

~ 10 millisecond timescale



ATP synthase

Size scale can be addressed effectively by larger clusters of computers

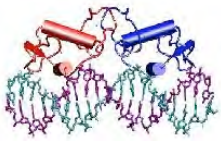
HP 735 cluster
12 processors
(1993)

SGI Origin 2000
128 processors (1997)

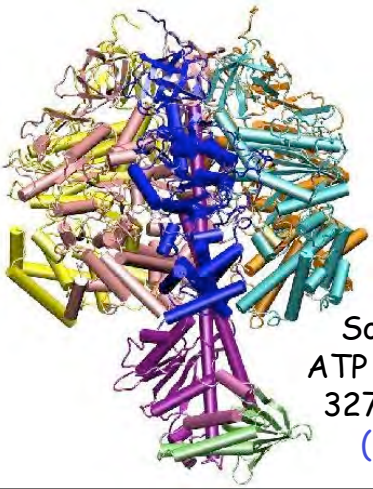
PSC LeMieux AlphaServer SC
3000 processors (2002)

Moving to Larger Molecules and Cellular Machines

BPTI
3K atoms
(1993)



Estrogen Receptor
36K atoms (1996)



Solvated
ATP Synthase
327K atoms
(2002)

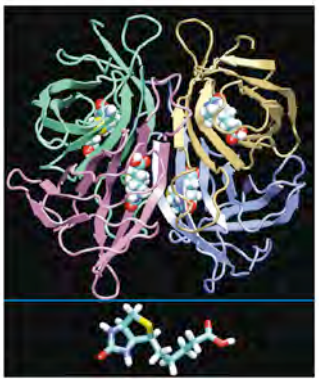
Time scales of several ns (100k) to μs (10k), but …

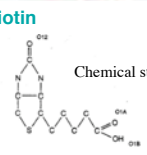
Steered Molecular Dynamics

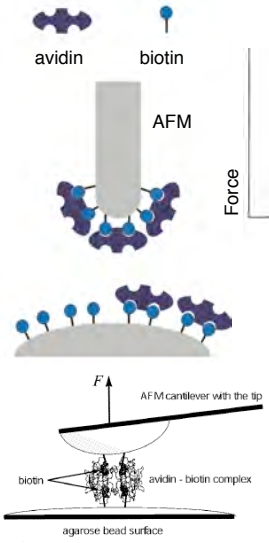
- Single Molecule Experiments: AFM
- Accelerating Events
  - Ligand Binding/Unbinding
  - Unfolding Experiments
  - Applying Surface Tension
  - Torque Application
  - Pressure induction
  - Inducing Large Domain Conformational Changes

- Constant-force: $f(t) = C$
- Constant-velocity: $f(t) = k\,[vt - (x_t - x_0)]$

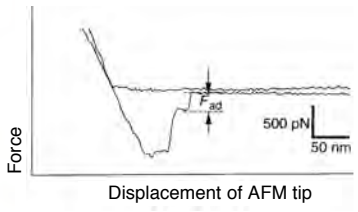
Atomic Force Microscopy Experiments of Ligand Unbinding

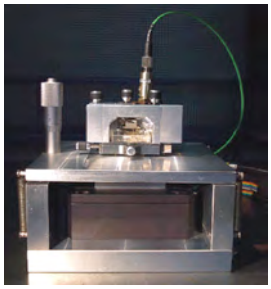






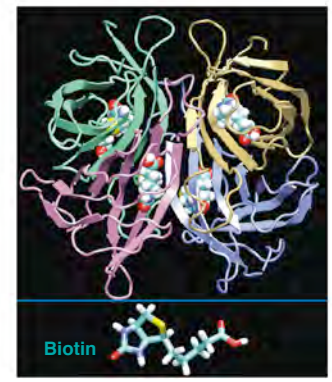
Florin et al., Science 264:415 (1994)

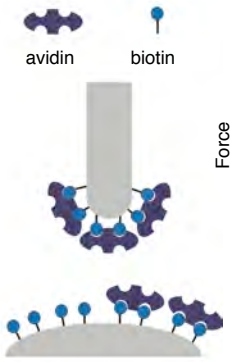


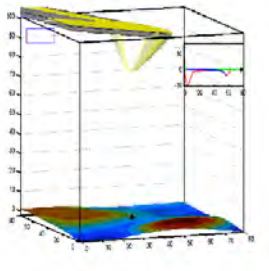


Atomic Force Microscopy Experiments of Ligand Unbinding

Florin et al., Science 264:415 (1994)

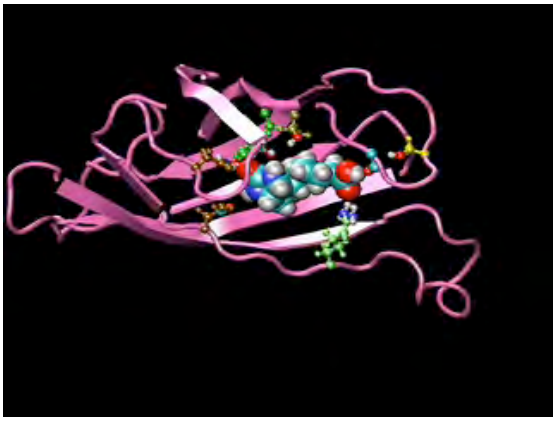






NIH Resource for Macromolecular Modeling and Bioinformatics
Theoretical Biophysics Group, Beckman Institute, UIUC

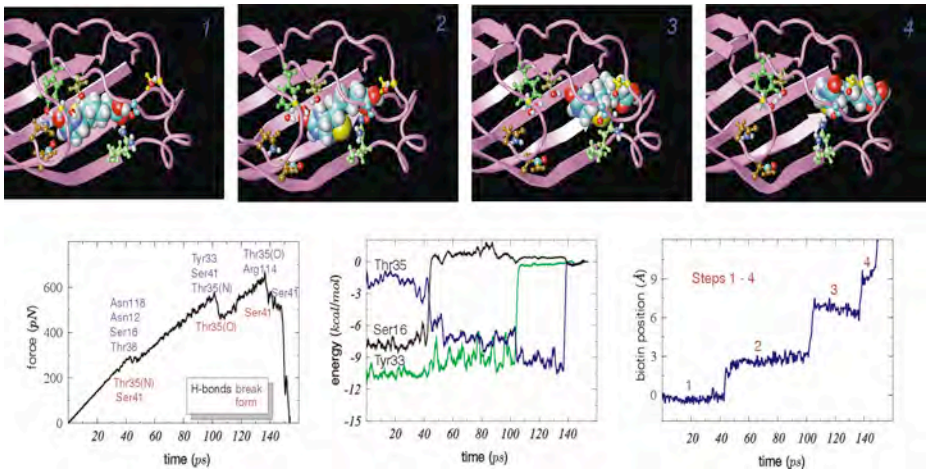
Pulling Biotin out of Avidin



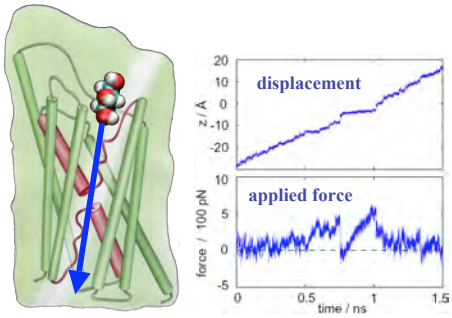
Molecular dynamics study of unbinding of the avidin-biotin complex. Sergei Izrailev, Sergey Stepaniants, Manel Balsera, Yoshi Oono, and Klaus Schulten. *Biophysical Journal*, 72:1568-1581, 1997.

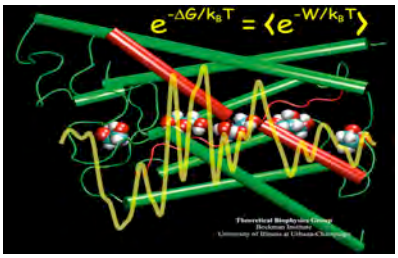
SMD of Biotin Unbinding: What Was Learnt

biotin slips out in steps, guided by amino acid side groups, water molecules act as lubricant, MD overestimates extrusion force



Quantitative Analysis of Substrate Permeation





Jensen et al, *PNAS* **99:** 6731-6736 (2002)

Calculation of the free energy profile of sugar transport from SMD simulations by Jarzynski's identity

Thermodynamics: $\Delta G \leq \langle W \rangle$

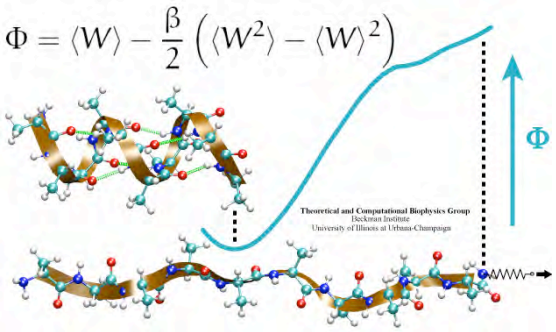
Is there any chance to discount the irreversible work? Yes!

Free Energy of Stretched Alpha-Helix

Thermodynamics: $\Delta G \le \langle W \rangle$

Jarzynski (1997): $e^{-\Delta G / k_B T} = \left\langle e^{-W / k_B T} \right\rangle$

$$\Phi = \langle W \rangle - \frac{\beta}{2}\left(\langle W^2 \rangle - \langle W \rangle^2\right)$$



Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. *Journal of Chemical Physics* , 119:3559-3566, 2003

Calculating potentials of mean force from steered molcular dynamics simulations. S. Park and K. Schulten. *Journal of Chemical Physics* , 120: 5946-5961, 2004