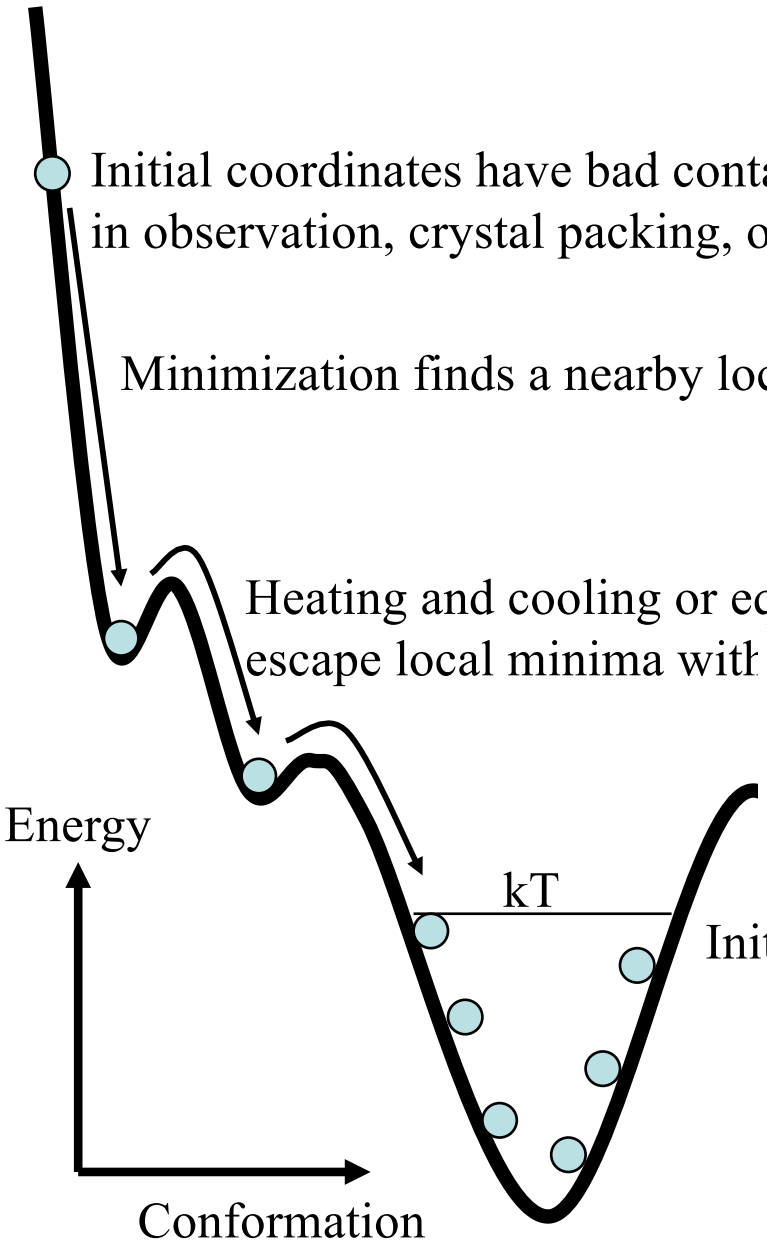
Molecular Dynamics to Sample Energy Landscape



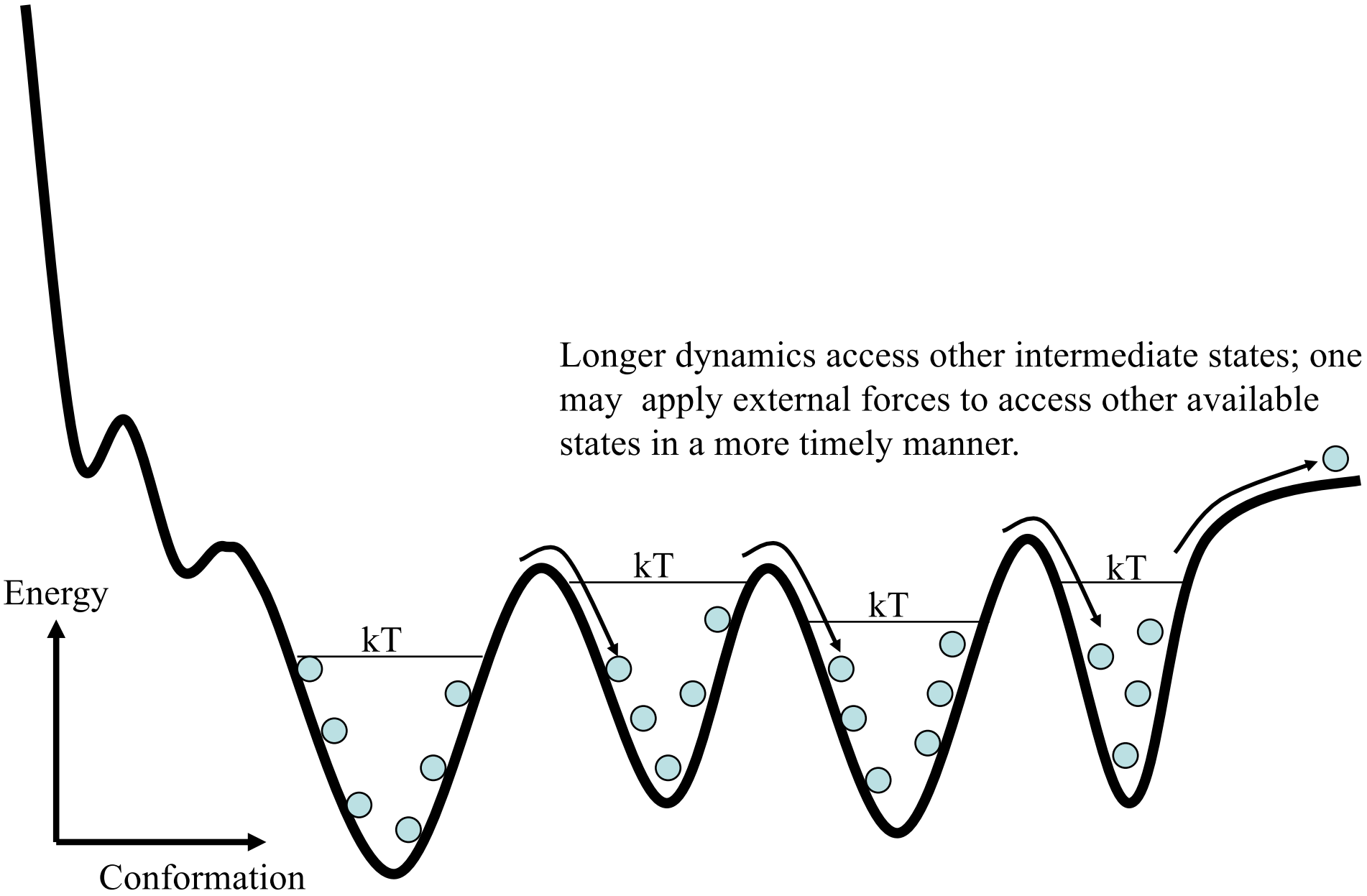
Initial coordinates have bad contacts, causing high energies and forces (due to averaging in observation, crystal packing, or due to difference between theoretical and actual forces)

Minimization finds a nearby local minimum.

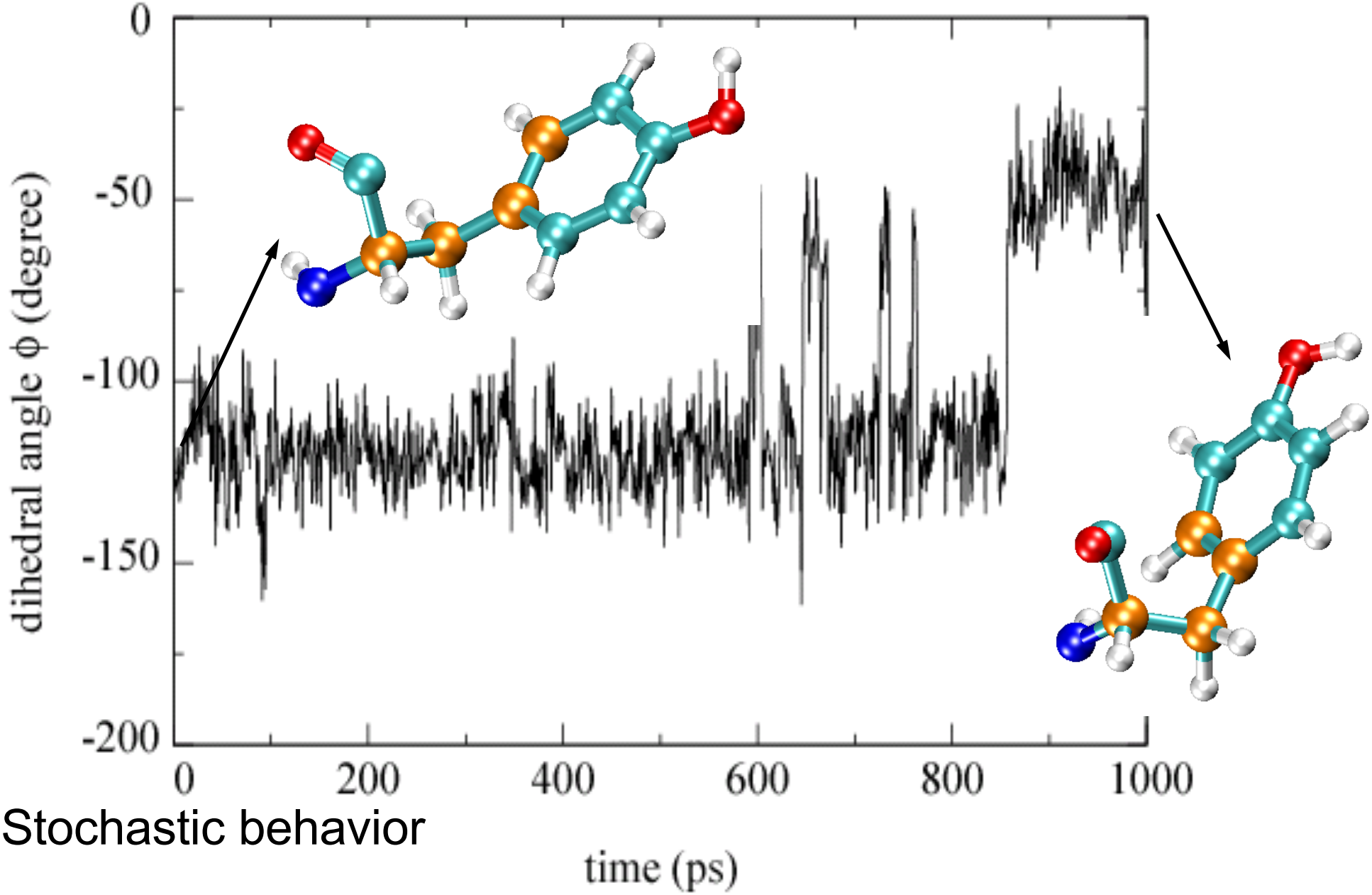
Heating and cooling or equilibration at fixed temperature permits biopolymer to escape local minima with

Initial dynamics samples thermally accessible states.

Molecular Dynamics to Sample Energy Landscape



Patience is required to observe Molecular Events



Stochastic behavior

Molecular Dynamics Ensembles

Constant energy, constant volume (NVE)

Constant temperature, constant volume (NVT)

Constant temperature, constant pressure (NPT)

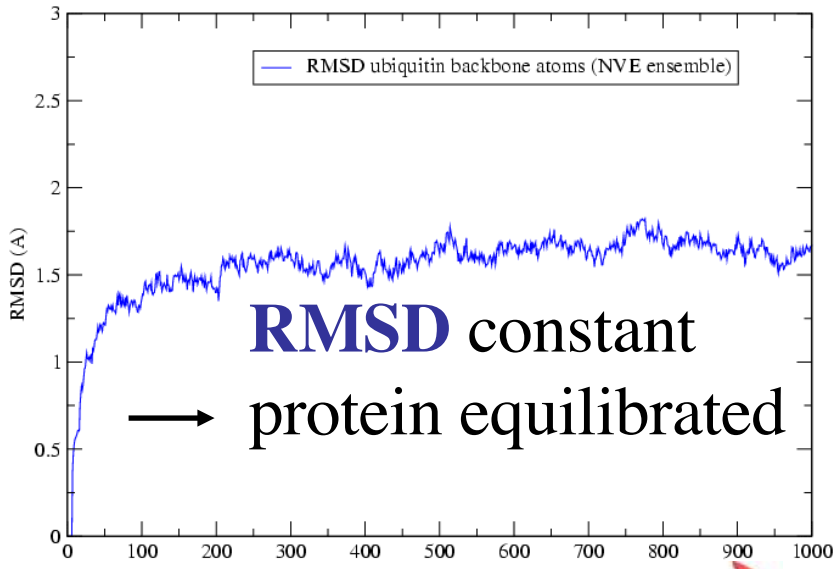
Choose the ensemble that best fits your system and start the simulations

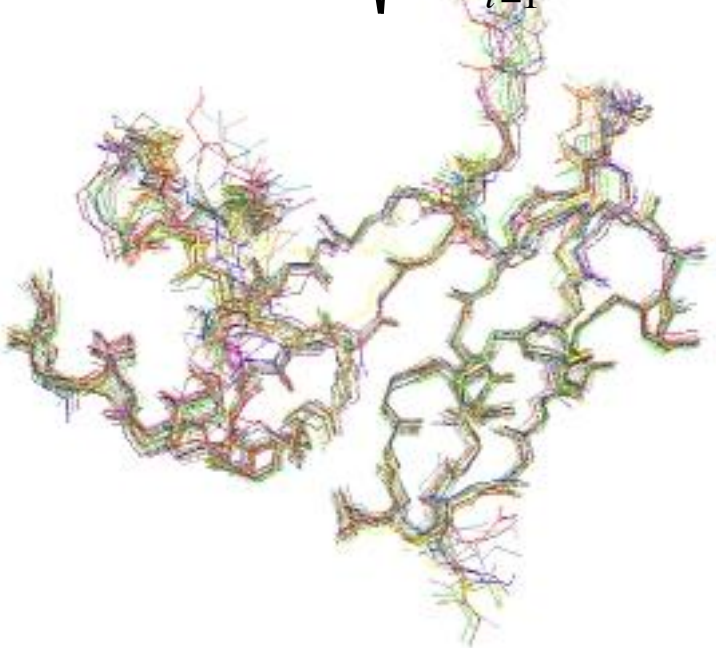
Equilibrium Properties of Proteins

Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

$$RMSD(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(R_i(t) - R_i(0)\right)^2}$$

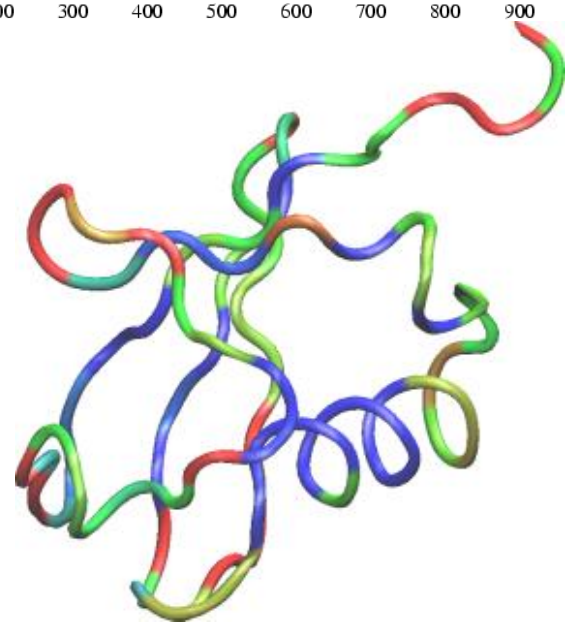




NMR structures

aligned together to see flexibility

Protein sequence exhibits characteristic permanent flexibility!

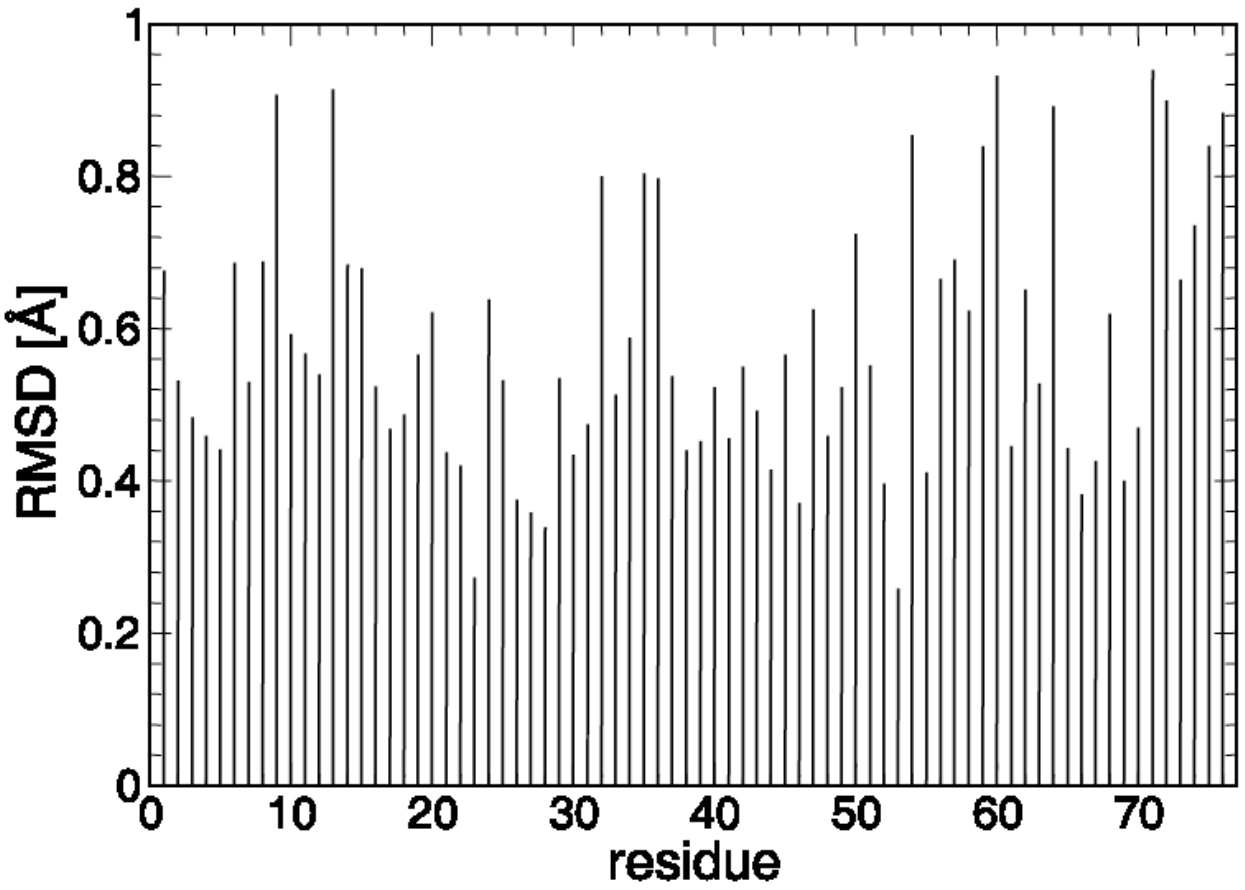


MD simulation

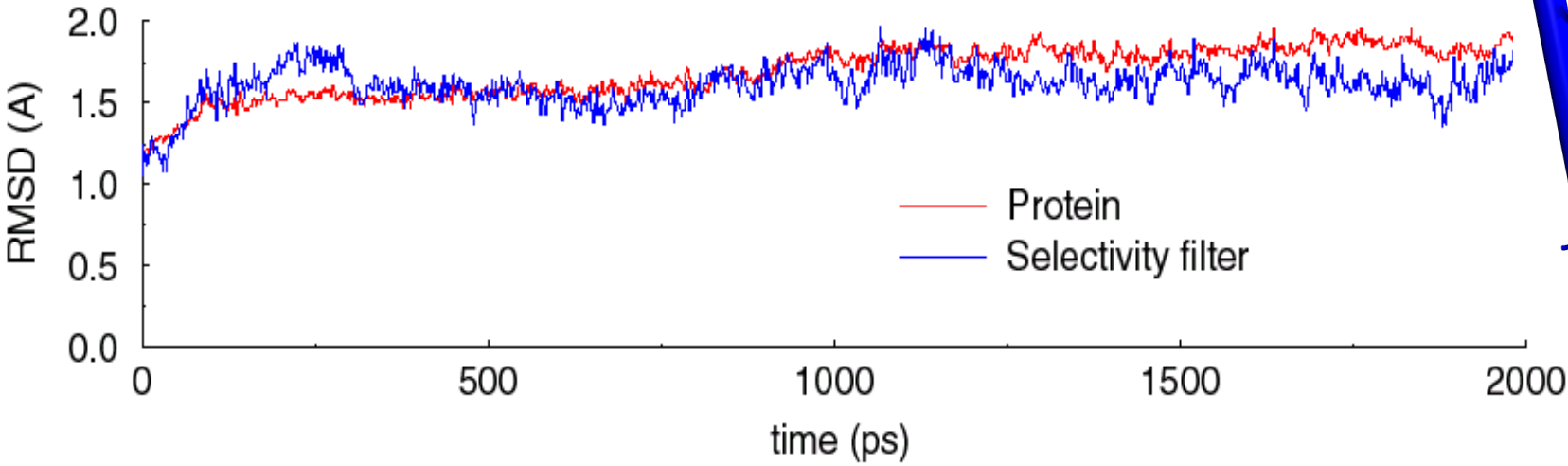
The color represents mobility of the protein through simulation (red = more flexible)

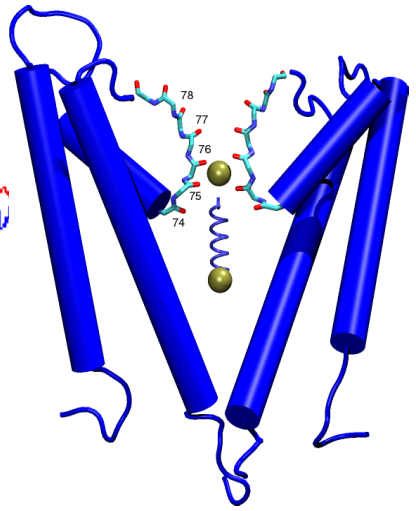
Thermal Motion of Ubiquitin from MD

RMSD values per residue

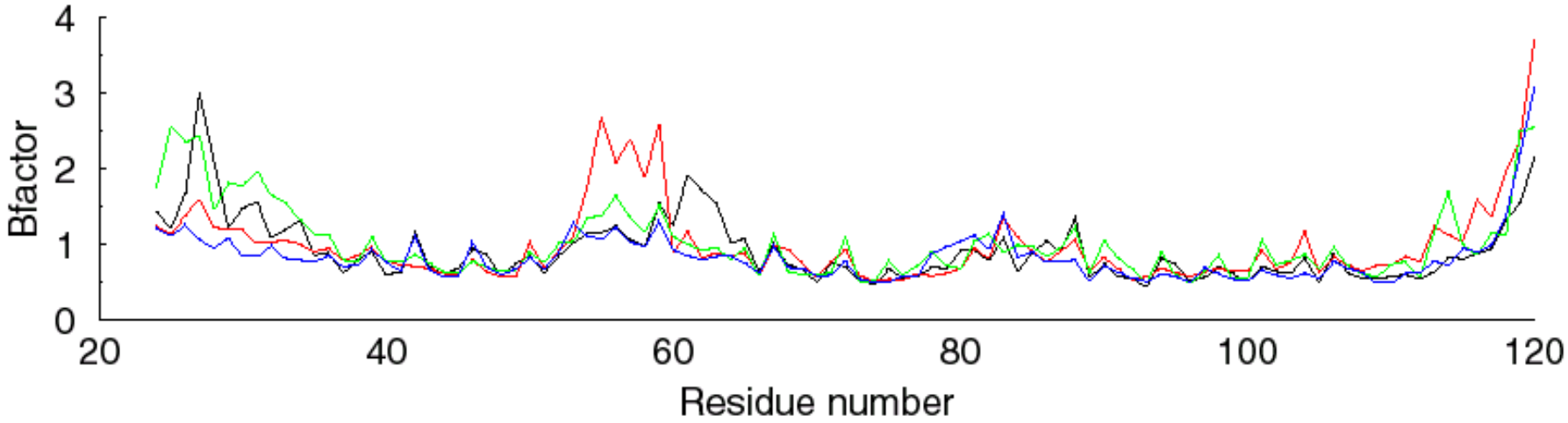


MD Results





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.



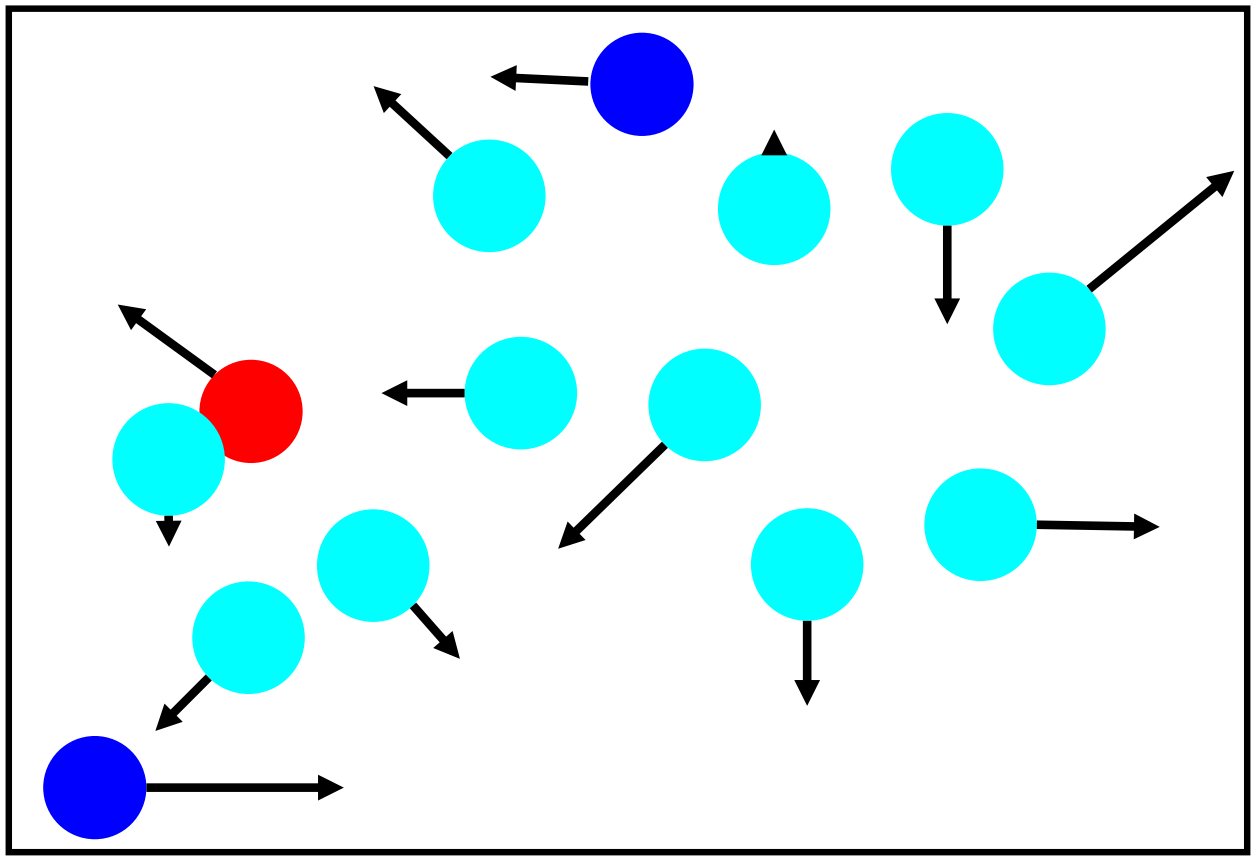
Temperature factors for individual residues in the four monomers of the KcsA channel protein indicate that the most flexible parts of the protein are the N and C terminal ends, residues 52-60 and residues 84-90. Residues 74-80 in the selectivity filter have low temperature factors and are very stable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

Improving the speed
cutoffs, PME, rigid bonds, multiple timesteps

- Nonbonded interactions require order N^2 computer time!
  - Truncating at R_{cutoff} reduces this to order $N\ R_{cutoff}^3$
  - Particle mesh Ewald (PME) method adds long range electrostatics at order N log N, only minor cost compared to cutoff calculation.

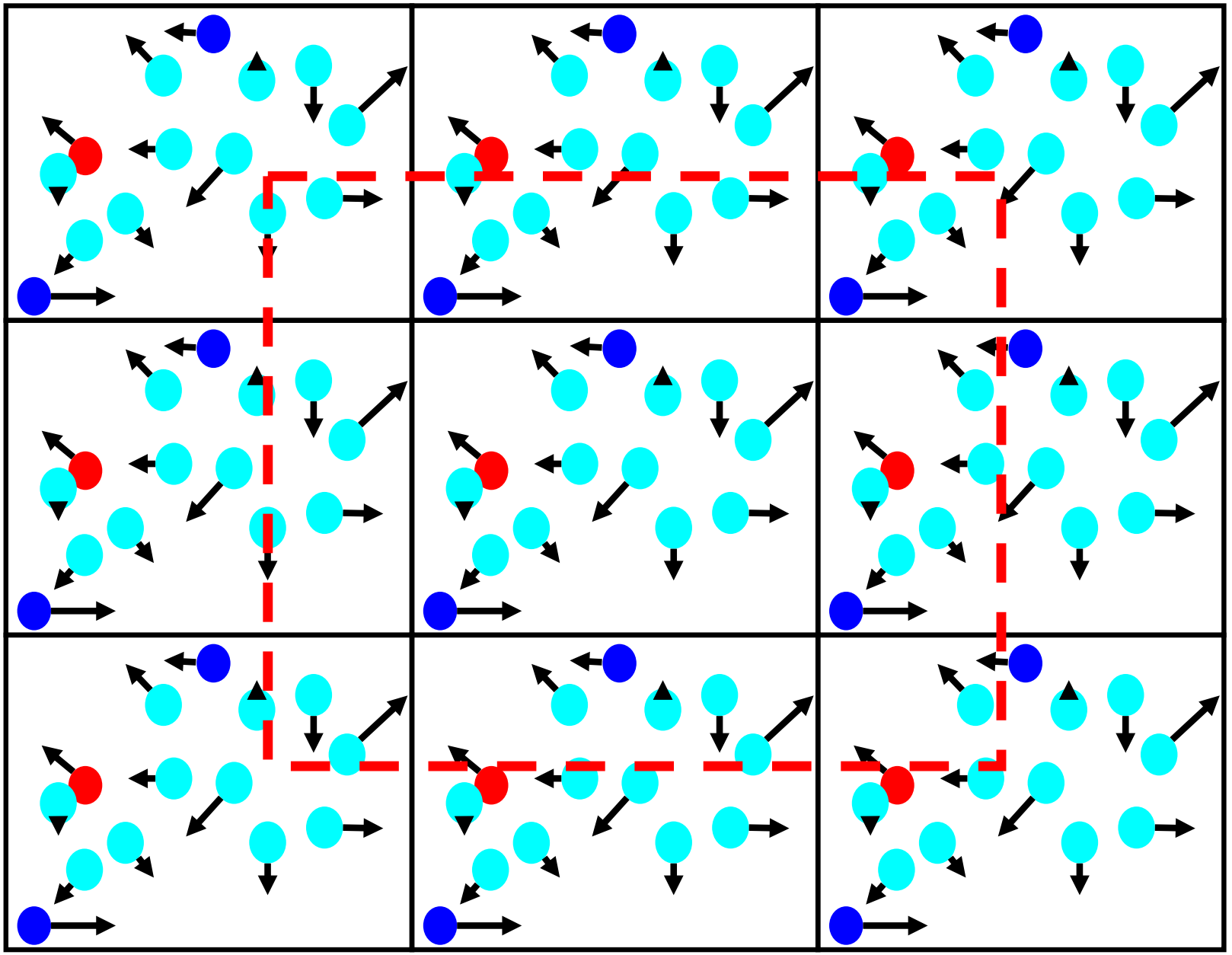
- Can we extend the timestep, and do this work fewer times?
  - Bonds to hydrogen atoms, which require a 1fs timestep, can be held at their equilibrium lengths, allowing 2fs steps.
  - Long range electrostatics forces vary slowly, and may be evaluated less often, such as on every second or third step.

Boundary Conditions?

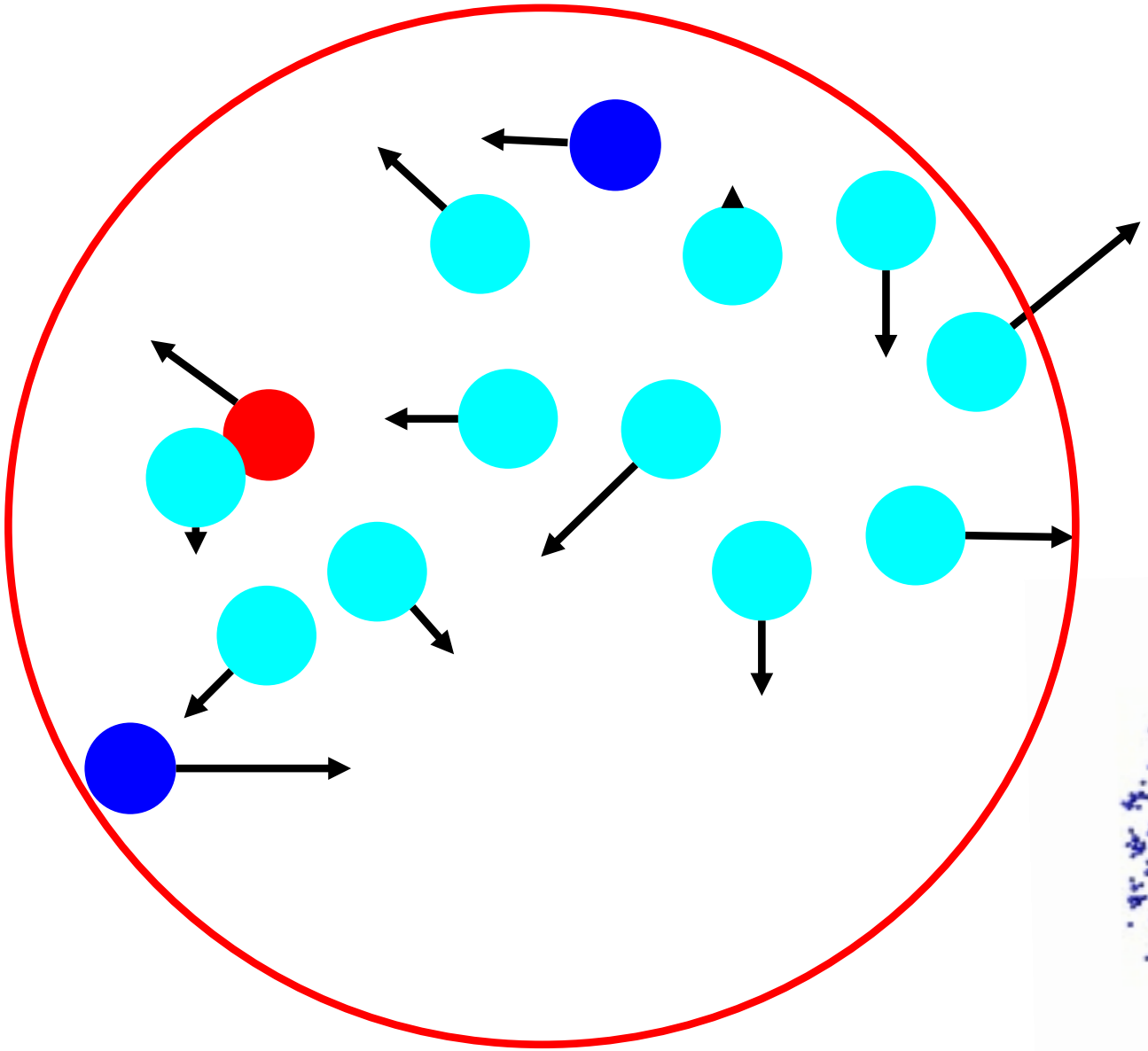


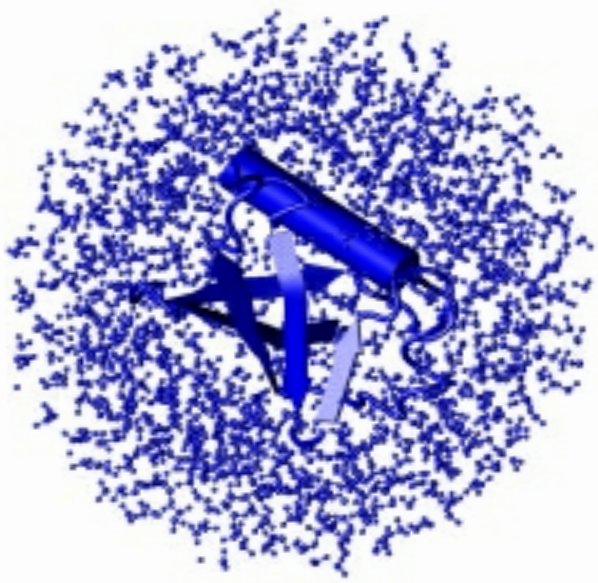
What happens if you put water under vacuum!?
Problems: Density, pressure, boundary effects, …
One solution: reflective boundaries, not quite good.

Periodic Boundary Conditions



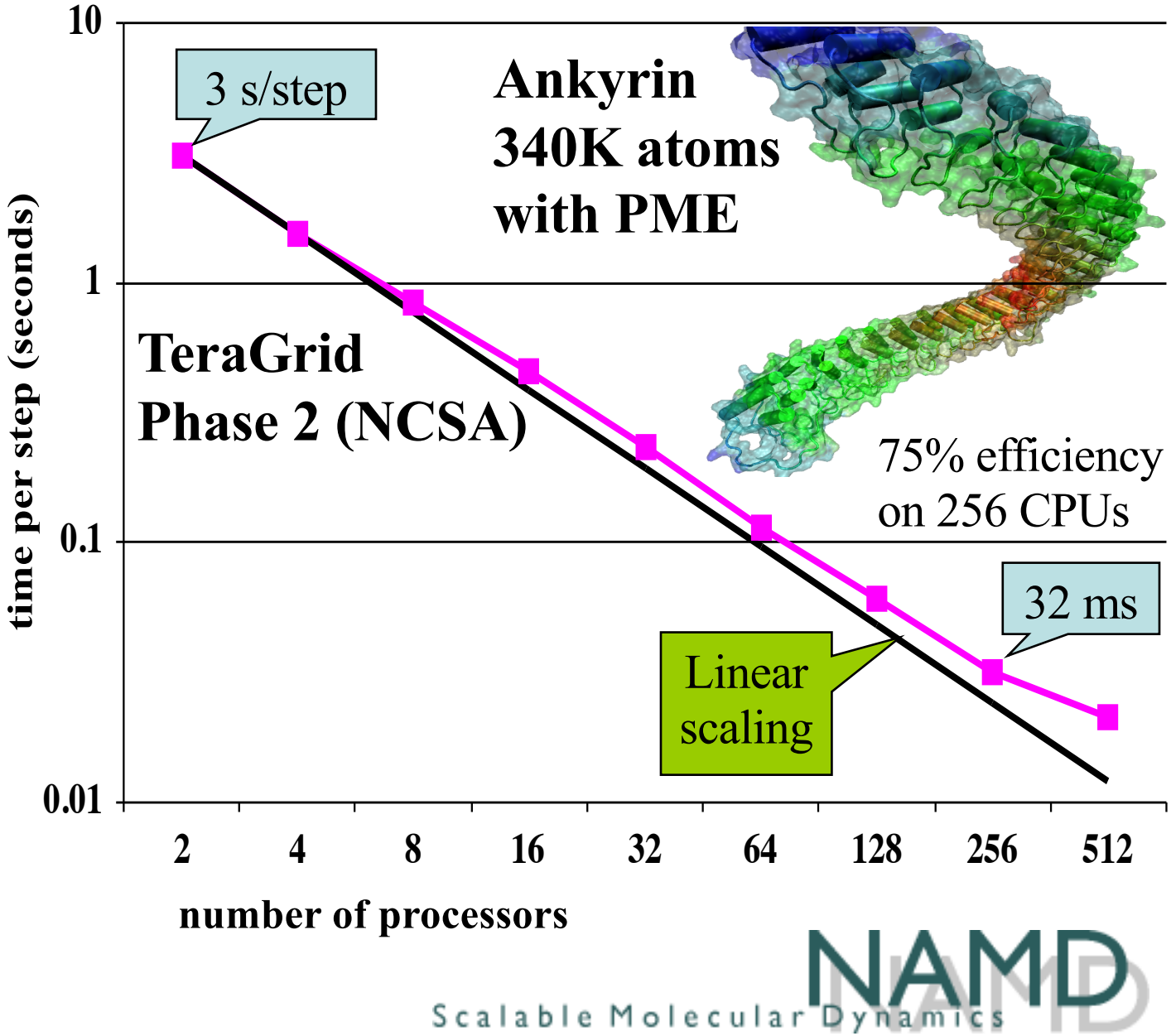
Spherical boundary conditions





NAMD: The Program we will Use

NAMD programmer J. Phillips Ph. D. UIUC Physics



Simulation of large biomolecular systems

2002 Gordon Bell Award for parallel scalability.

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

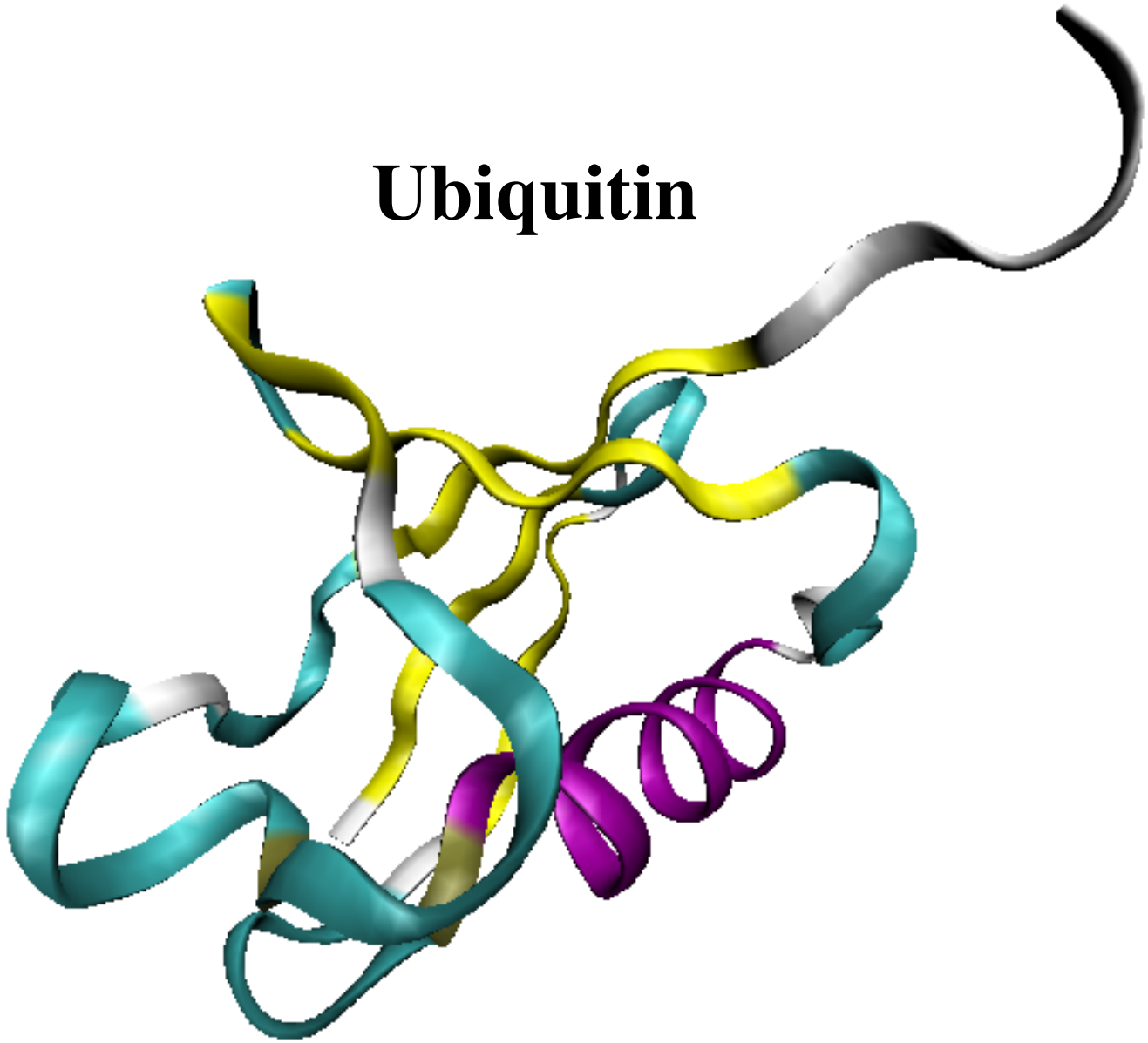
Available for **FREE** as precompiled binaries; includes source code.

10,000 registered users.

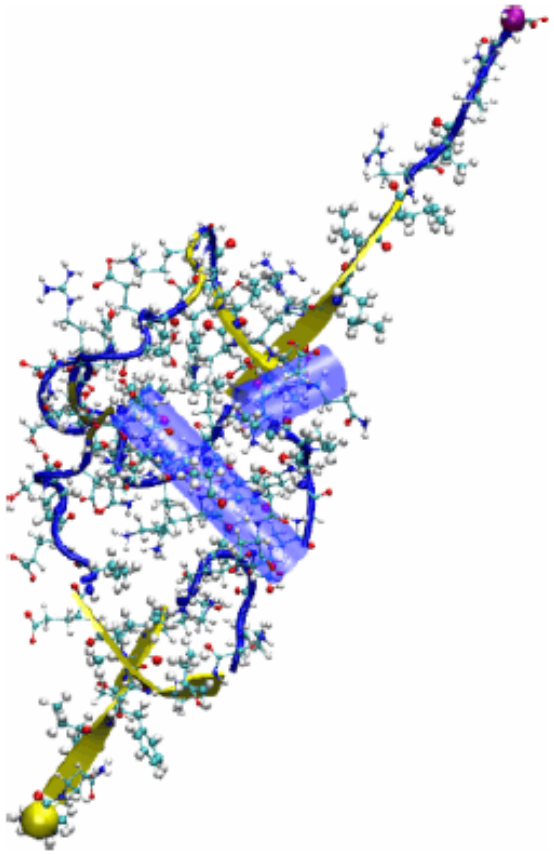
NAMD
Scalable Molecular Dynamics

Molecular Dynamics of Proteins

Ubiquitin



NAMD TUTORIAL



NAMD Developer: James Phillips

Timothy Isgro

James Phillips

Marcos Sotomayor

Elizabeth Villa

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

```
# 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

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

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.
```

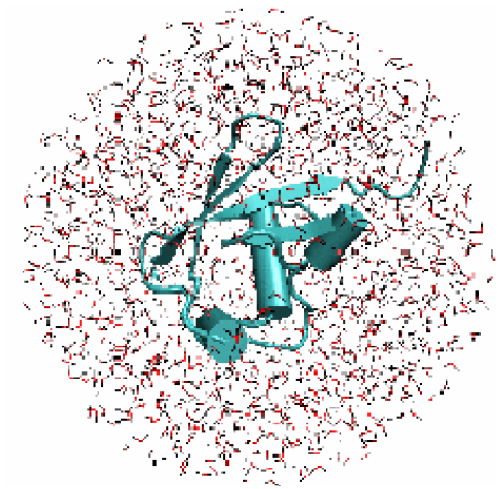
The NAMD Experience / 1

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

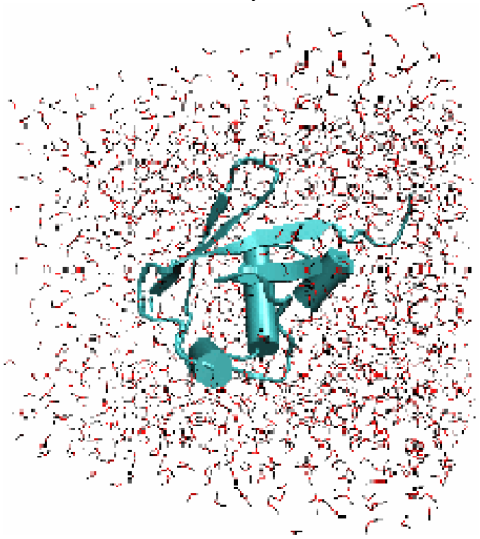
Generating a Protein Structure File (PSF)

- 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)



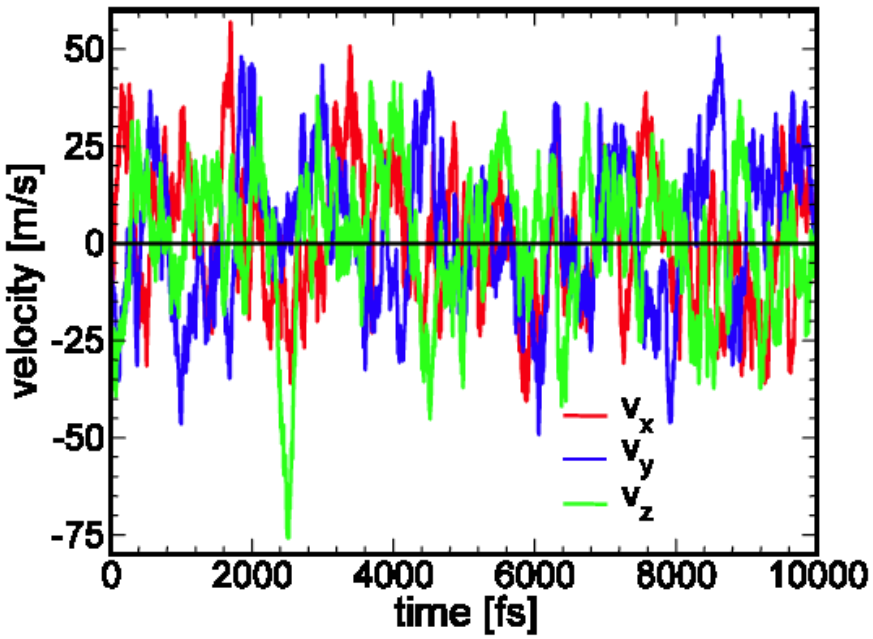
The NAMD Experience / 2

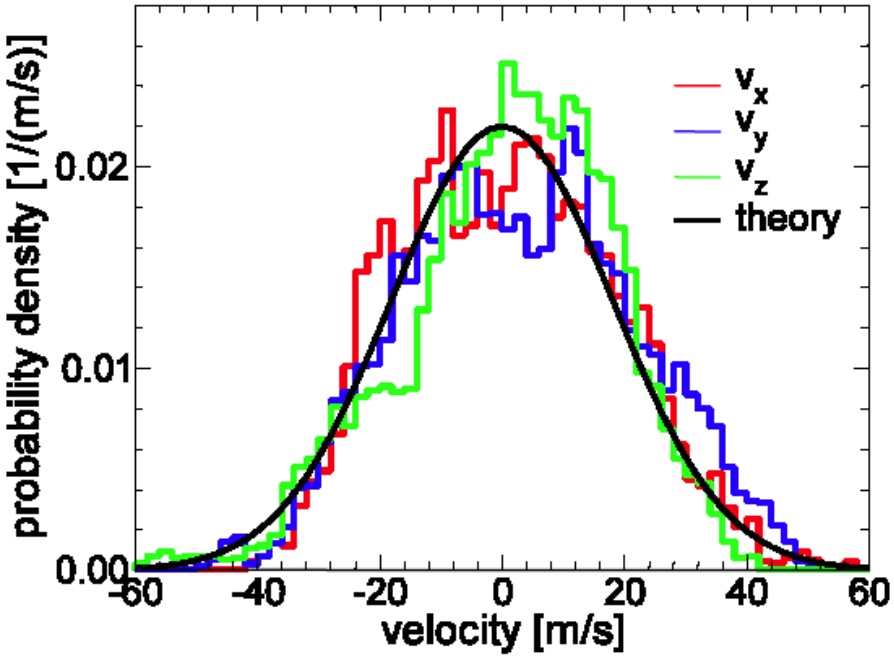
- RMSD value for equilibration
- Atomic RMSD values of equilibrated protein
- Velocity distribution
- Temperature distribution
- Specific heat
- Diffusion of whole protein
- Heat diffusion
- Temperature Echoes

NAMD Tutorial (Part 2)

- **2 Analysis**
  - 2.1 Equilibrium
    - 2.1.1 RMSD for individual residues
    - 2.1.2 Maxwell-Boltzmann Distribution
    - 2.1.3 Energies
    - 2.1.4 Temperature distribution
    - 2.1.5 Specific Heat

Maxwell Distribution of Atomic Velocities

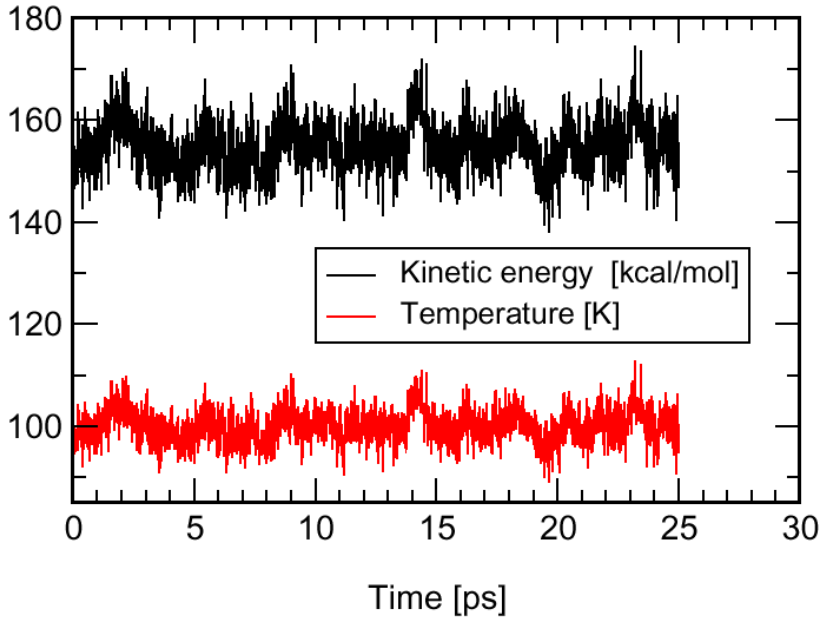


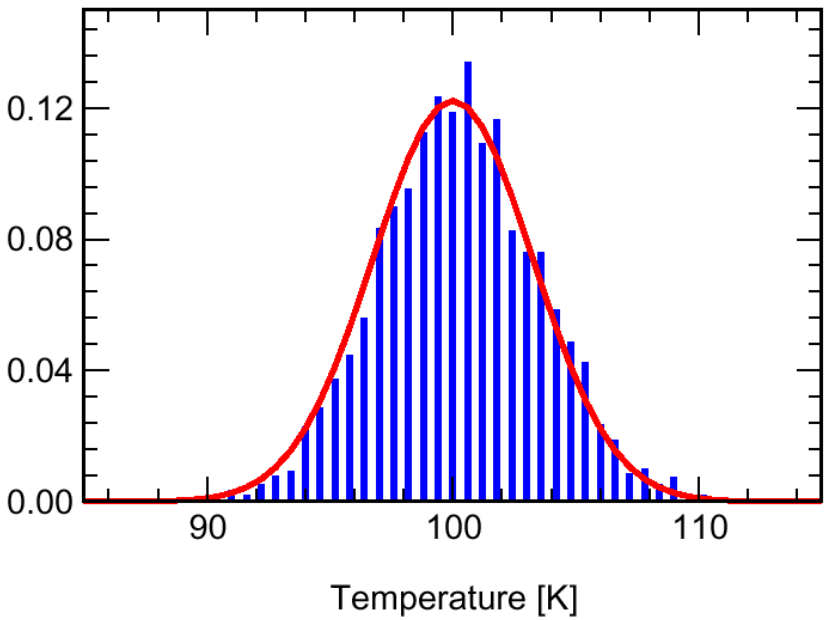


$$p(v_\sigma) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left[-\frac{m v_\sigma^2}{2 k_B T}\right]$$

$$\sigma = x,\ y,\ z$$

Analysis of E_{kin}, T (free dynamics)



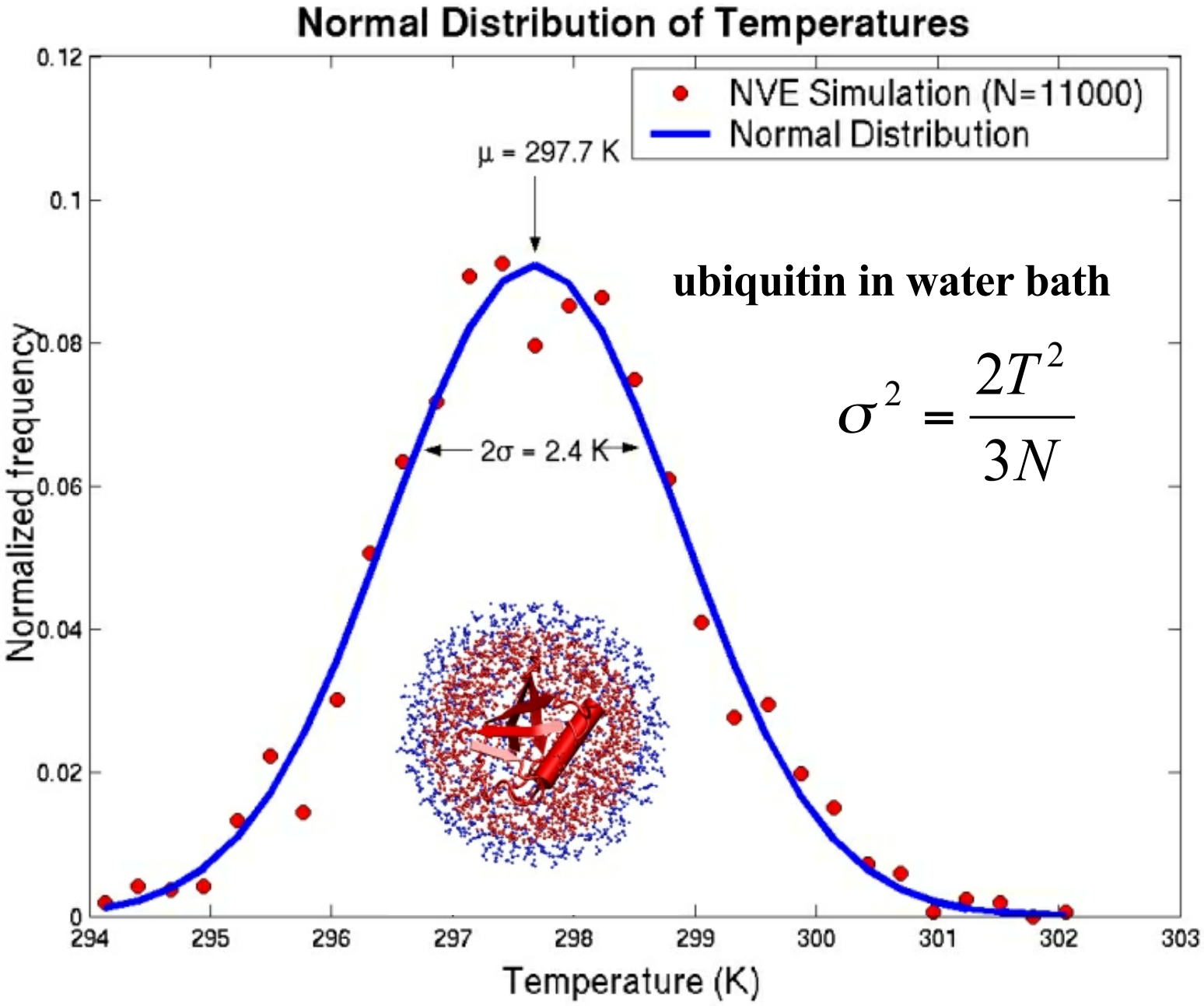


Definition of Temperature

$$\langle \sum_j \frac{1}{2} m_j v_j^2 \rangle = \frac{3}{2} N k_B T$$

$$T = \frac{2}{3N k_B} \langle \sum_j \frac{1}{2} m_j v_j^2 \rangle$$

The atomic velocities of a protein establish a thermometer.


Normal Distribution of Temperatures
NVE Simulation (N=11000)
Normal Distribution
μ = 297.7 K
2σ = 2.4 K
ubiquitin in water bath
$\sigma^2 = \frac{2T^2}{3N}$
Normalized frequency
Temperature (K)
0.12
0.1
0.08
0.06
0.04
0.02
0
294
295
296
297
298
299
300
301
302
303