

# MD

## SASSIE CCP-SAS Workshop

January 23-25, 2017

ISIS Neutron and Muon Source

Rutherford Appleton Laboratory UK

# Today: “build, minimize, dynamics”

4:30 PM - 5:30 PM: **MD: NAMD MINIMIZATION to DYNAMICS & Analysis**

5:30 PM - 7:30 : Lab III: NAMD (w/ water, PBC, restart, equilibrium)

# Barriers: BUILD; EQUILIBRATE; PROPAGATE; ANALYZE

What software package(s) and force-fields do I use?

**Starting structure?**

**How do I clean up the structure?**

**How do I set up a trajectory (time or space)?**

How do I calculate scattering observables correctly?

# Overview

**PDB + PSF + INPUT FILE : run a NAMD trajectory**

Minimization (in vacuo)

Adding water & ions

PBC & PME

Fixed atom methods and control

Molecular dynamics parameters

IMD (link NAMD with VMD)

Restarting a simulation (continuation)

Output files

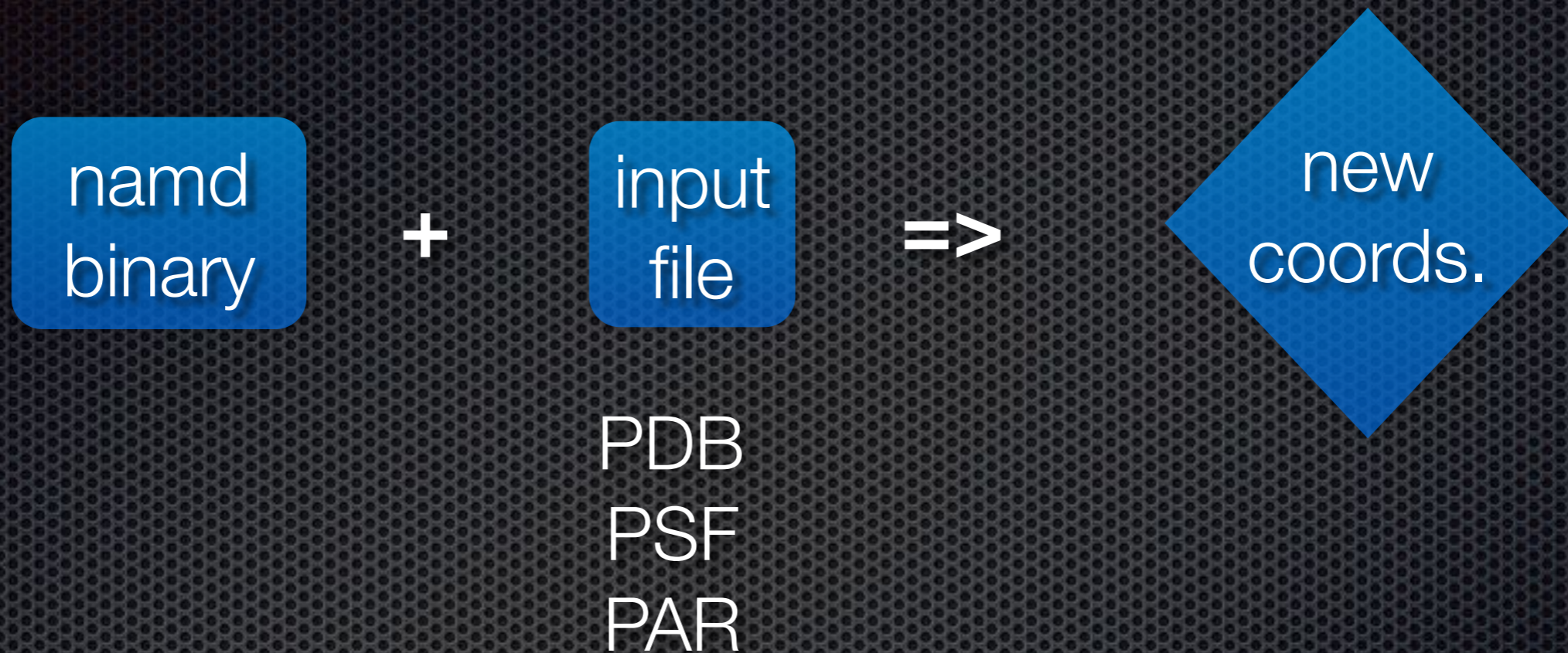
## **Analysis of MD & Connection to NS**

Testing for equilibration

Neutron scattering (dynamics)

Connection MD trajectory to  $S(q,w)$

# NAMD II



MD parameters

1. Minimization
2. MD (time)
3. Directed MD

# NAMD INPUT FILE

```
# sample NAMD configuration file for Minimization
```

```
# molecular system
```

```
coordinates      output/new_icln.pdb
structure        output/new_icln.psf
temperature      300
```

PDB

PSF

```
# force field
```

```
paratypecharmm  on
parameters      /home/mdschool/toppar/par_all27_prot_na.inp
```

PAR

```
# approximations
```

```
exclude         scaled1-4
1-4scaling      1.0
switching       on
switchdist      10
cutoff          12
```

MD

parameters

```
# output
```

```
outputname      output/min0
binaryoutput    no
```

```
# run control
```

```
minimize        1000
```

# RUNNING NAMD

## USAGE (unix prompt):

Start a run:

```
>namd2 min0 >& min0.out &
```

Check status:

```
>tail -90f min0.out
```

Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 min0 >& min0.out &
```

### Energy Minimization

---

run name	<input type="text" value="run_0"/>	
reference pdb	<input type="button" value="Browse..."/> dbd.pdb	OR <input type="button" value="Browse server"/> Local: dbd.pdb
input filename (dcd or pdb)	<input type="button" value="Browse..."/> dbd.pdb	OR <input type="button" value="Browse server"/> Local: dbd.pdb
PSF file name	<input type="button" value="Browse..."/> dbd.psf	OR <input type="button" value="Browse server"/> Local: dbd.psf
output file name (dcd)	<input type="text" value="min run 0.dcd"/>	
number of minimization steps	<input type="text" value="100"/>	
number of processors	<input type="text" value="1"/>	
keep run output files	<input type="button" value="no"/>	
DCD write frequency	<input type="text" value="20"/>	
run type	<input type="button" value="minimization"/>	

---

### Advanced Input

Check Box for Advanced Input

# NAMD OUTPUT (MIN.)

...

BRACKET: 0.00207889 7.0816 -1821.94 1646.18 8572.11  
TIMING: 1000 CPU: 24.2692, 0.0240947/step Wall: 24.3809, 0.0241311/step, 0  
hours remaining, 650704 kB of memory in use.

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

ENERGY:	1000	4561.4004	3396.5872	881.8814	15.2681
-148982.9533	16811.5791	0.0000	0.0000	0.0000	0.0000
-123316.2371	0.0000	-123316.2371	-123316.2371	0.0000	0.0000
-4162.6459	-4161.2347	366289.0000	-4162.6459	-4161.2347	

WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 1000  
WRITING COORDINATES TO OUTPUT FILE AT STEP 1000  
WRITING VELOCITIES TO OUTPUT FILE AT STEP 1000  
=====

WallClock: 302.991364 CPUtime: 309.270966 Memory: 690456 kB

Plotting  
(Lab II)

Check status:  
**>tail -90f min0.out | grep TOL**

findstr TOL

# MINIMIZATION NOTES

**Large negative energies (especially electrostatic and VDW)**

**Local minimum ...**

**“Minimized structures” often fail when you use min. coords for MD runs**

**Topologically unlikely structures can be minimized**

**Cycle of minimization and short MD (with constraints) may be necessary**

**Some convergence issues could indicate a poor structure (coor or PSF).**

# SOLVATE

Many tools (GUI & scripts) are available to automatically “dunk” your bio-molecule into a water-box. ---> VMD

The VMD plugin requires that you have a **PSF/PDB** of your system.

**The VMD plugin will generate NEW a PDB/PSF pair automatically.**

The screenshot shows the 'Solvate' GUI window with the following settings:

- Input:**
  - PSF: /Users/curt sj/mppp/new\_complex.psf (Browse)
  - PDB: /Users/curt sj/mppp/new\_complex.pdb (Browse)
  - Waterbox Only
  - Rotate to minimize volume
  - Rotation Increment (deg): 10
  - Selection for Rotation: all
- Output:**
  - solvate (Browse)
- Segmon: ID Prefix:** WT
- Boundary:** 2.4
- Box Size:**
  - Min: x: [ ] y: [ ] z: [ ]
  - Max: x: [ ] y: [ ] z: [ ]
  - Use Molecule Dimensions
- Box Padding:**
  - Min: x: 0 y: 0 z: 0
  - Max: x: 0 y: 0 z: 0
- Use nonstandard solvent
- Solvent: box PDB: [ ]
- Solvent: box PSF: [ ]
- Solvent: box topology: [ ]
- Solvent: box side length: [ ]
- Solvent: box key selection: [ ]

A 'Solvate' button is located at the bottom of the window.

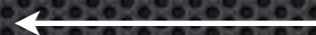
# SOLVATE II

The screenshot shows the Solvate application window with the following fields and options:

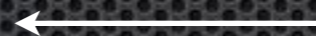
- Input:**
  - PSF:
  - PDB:
  - Waterbox Only
  - Rotate to minimize volume      Rotation increment: (deg):
  - Selection for Rotation:
- Output:**
  -
- Segment ID Prefix:**
- Boundary:**
- Box Size:**
  - Min: x:  y:  z:
  - Max: x:  y:  z:
  - Use Molecule Dimensions
- Box Padding:**
  - Min: x:  y:  z:
  - Max: x:  y:  z:
- Use nonstandard solvent
- Solvent box PDB:
- Solvent box PSF:
- Solvent box topology:
- Solvent box side length:
- Solvent box key selection:



path to topology file  
path to PDB file



name & path to OUTPUT FILES



YOU pick a segment name



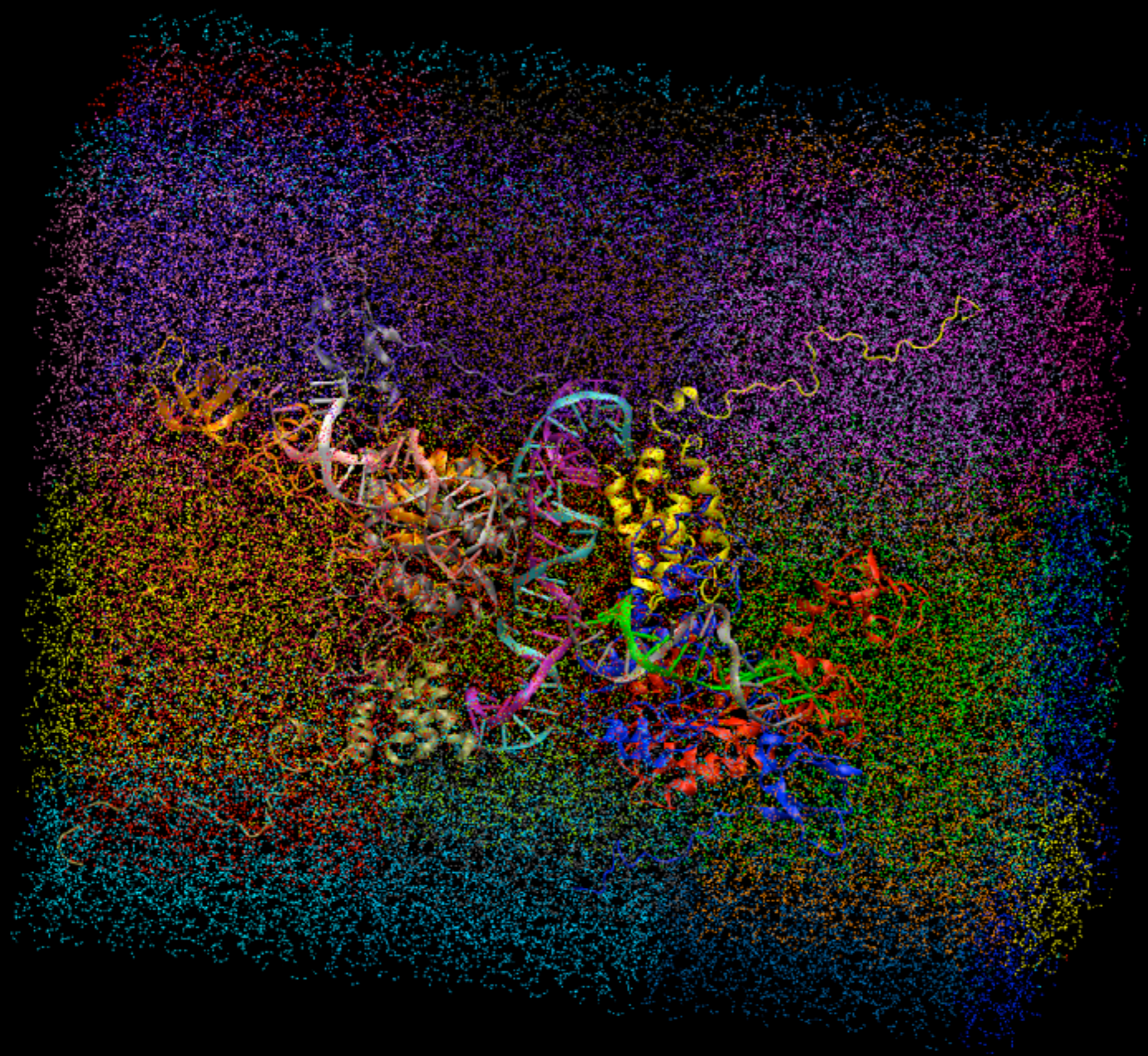
YOU pick MIN dimensions



YOU pick MAX dimensions

Write these down

Measure dimensions w/  
VMD and add ~12 Ang.  
pad (x,y,z +/- 12)





# SOLVATE II

Input  
PSF: /Users/curtsj/Desktop/summer\_school/ Browse  
PDB: /Users/curtsj/Desktop/summer\_school/ Browse  
 Waterbox Only  
 Rotate to minimize volume    Rotation increment: (deg): 10  
Selection for Rotation: all

Output  
solvated\_protein11 Browse

Segment ID Prefix: WT  
Boundary: 2.4

Box Size:  
Min: x: 30 y: 35 z: 32  
Max: x: 30 y: 35 z: 32  
 Use Molecule Dimensions

Box Padding:  
Min: x: 12 y: 12 z: 12  
Max: x: 14 y: 14 z: 14

Use nonstandard solvent  
Solvent box PDB:  
Solvent box PSF:  
Solvent box topology:  
Solvent box side length:  
Solvent box key selection:

Solvate

← path to topology file

← path to PDB file

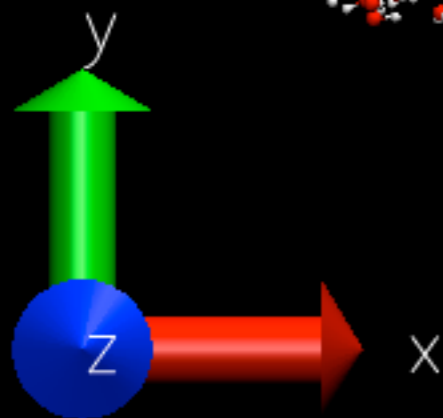
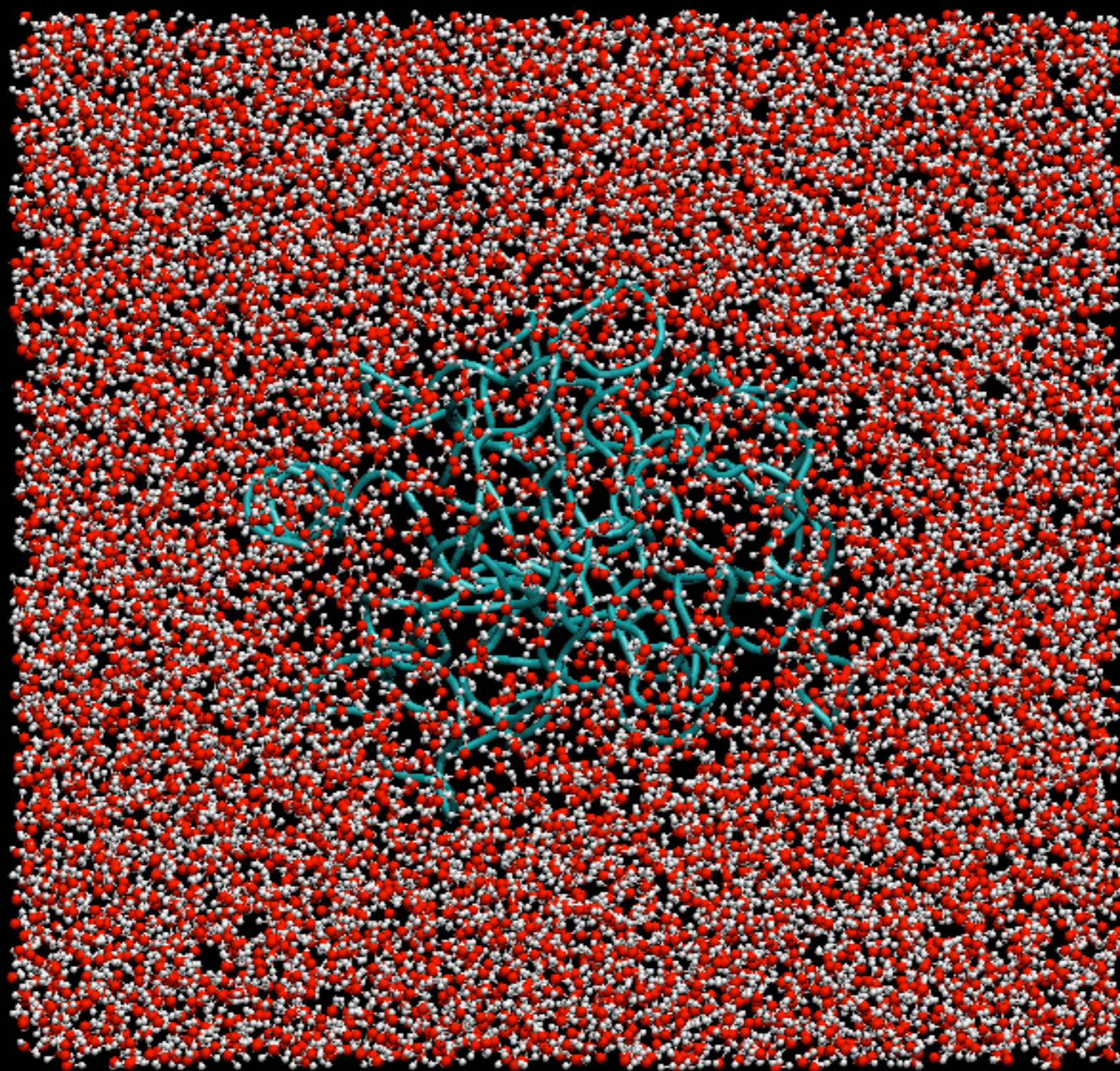
← name & path to OUTPUT FILES

← YOU pick a segment name

← YOU pick MIN padding

← YOU pick MAX padding

Write these down



Measure dimensions of final box ("2") graphically via VMD

OR

Use the VMD Console:

```
vmd > set everyone [atomselect top all]
```

```
atomselect32
```

```
vmd > measure minmax $everyone
```

```
{-9.96199989319 -10.5780000687 -14.7969999313}
```

```
{59.4099998474 55.422000885 60.5979995728}
```

```
dx ~ 70
```

```
dy ~ 66
```

```
dz ~ 76
```

# ADD IONS

Autoionize

Autoionize randomly places ions (NaCl/KCl) in a previously solvated system.

Input

PSF:

PDB:

Output prefix:

Ionic Concentration (defined as  $(\#Na + \#Cl)/V$ )

Concentration (mol/L):

User defined

Neutralize

# Na:

# Cl:

Min. distance from molecule (A):

Min. distance between ions (A):

Segment ID:

Switch to KCl instead of NaCl

← path to topology file

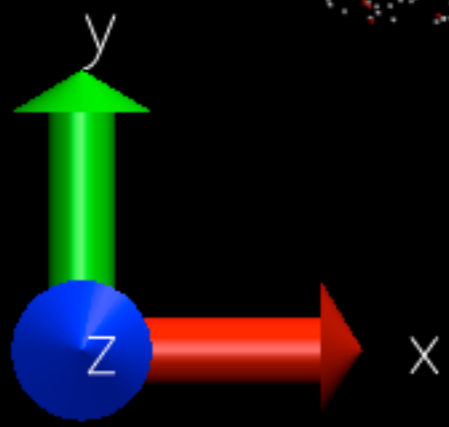
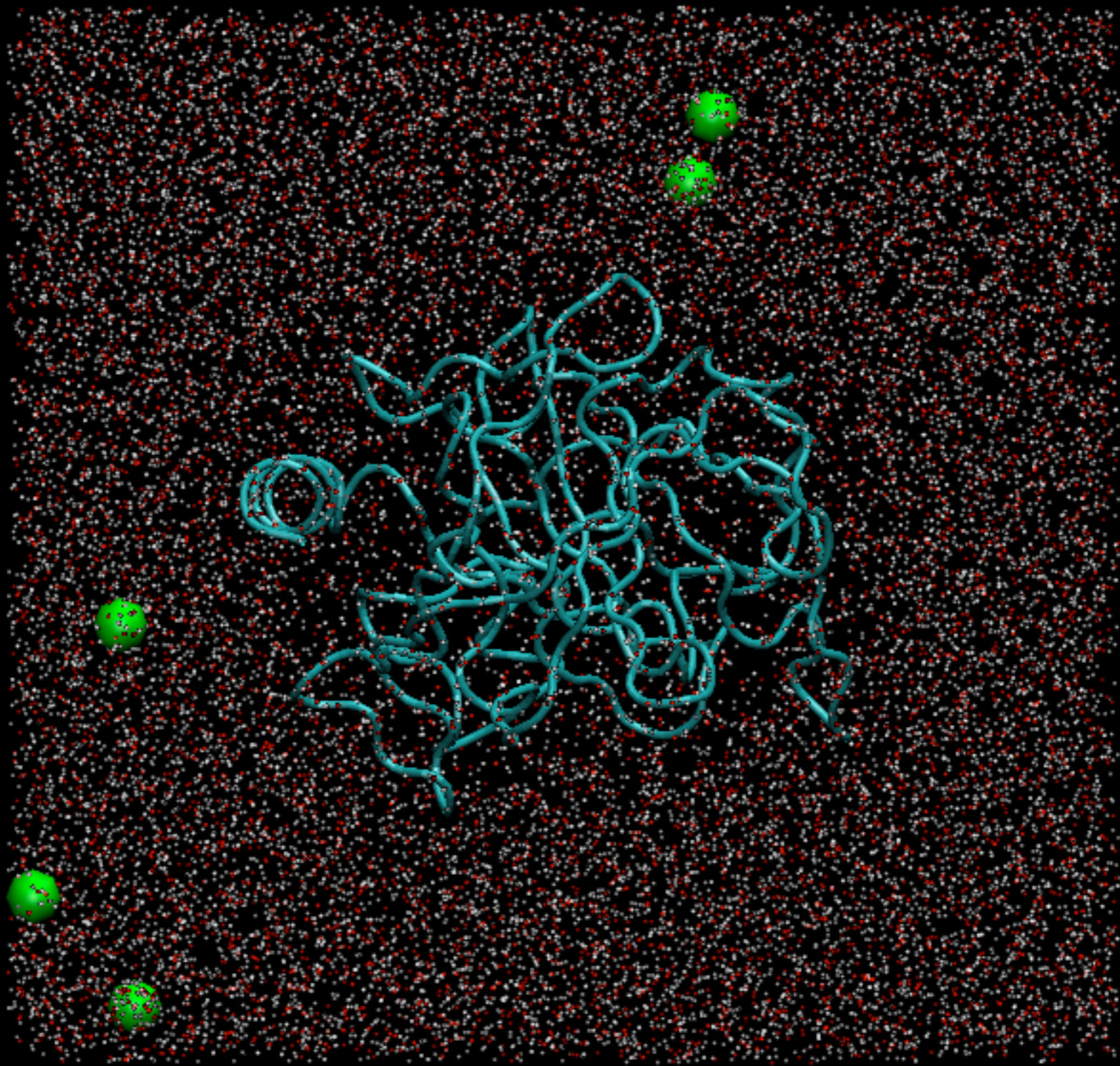
← path to PDB file

← name & path to OUTPUT FILES

← choose neutralize

Write this down

Other options available



# MINIMIZATION NOTES II

**Example steps for a solvated protein:**

- 0. Minimize vacuum structure (w/ restraints?) Then BUILD solvated system.**
- 1. Restrain protein atoms (let waters/ions relax)**
- 2. Restrain protein backbone (let side-chains and waters/ions relax)**
- 3. Restrain secondary structure elements (let loops, etc. relax)**
- 4. No restraints**
- 5. Start short MD run (10 ps or so) [[[ does it run? ]]]**

**Number of steps is something YOU need to monitor (1000 to 10000+ steps)**

# NAMD INPUT FILE

Now, we need to add to our NAMD input file to handle:

(1) Periodic boundary conditions

(2) Particle Mesh Ewald (PME) for PBC electrostatics

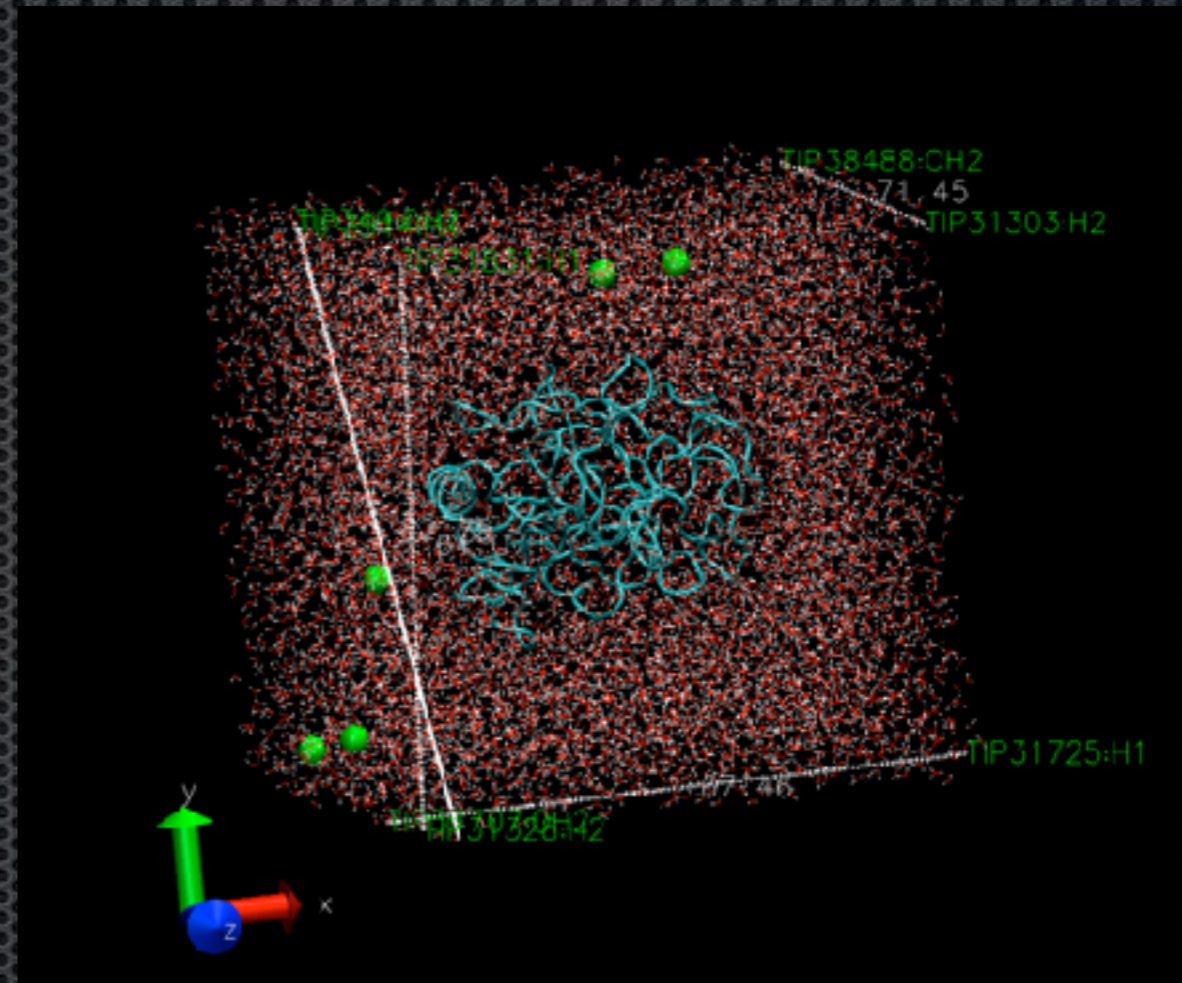
(3) Fixed atoms (for systematic “relaxation” and future needs)

--> TCL/PYTHON/PERL/MDTOOLS . . . “build utilities”

# PBC

Measure box (x,y,z)  
add values to the diagonal

```
. . .  
# pbc  
  
cellBasisVector1      71.0 0.0 0.0  
cellBasisVector2      0.0 67.0 0.0  
cellBasisVector3      0.0 0.0 77.0  
  
. . .  
# run control  
  
minimize              1000
```



It is generally prudent to add 1-2  
Angstroms in each dimension . . . NPT

# PME

Our test-case box (`ionized_solvated_proteinII`):

```
cellBasisVector1      68.0 0.0 0.0
cellBasisVector2      0.0 65.0 0.0
cellBasisVector3      0.0 0.0 72.0
```

```
. . .
# pme

PME                on
PMEGridSpacing     1.0

. . .

# run control

minimize           1000
```

Particle mesh is a 3D grid which the system charge is distributed. From this charge, F (and U) are calculated.

Grid spacing should be  $> 1/\text{Angstrom}$

# Fixed Atoms

```
. . .  
  
# restraints  
  
fixedAtoms          on  
fixedAtomsForces    on  
fixedAtomsFile       output/fixed_protein_atoms.pdb  
fixedAtomsCol        B  
  
. . .  
  
# run control  
  
minimize            1000
```

These specific commands:

(1) apply fixed atoms “on”

(2) instruct forces to be included FROM these atoms

(3) define the file that is used to tell NAMD which atoms are fixed

(4) define the COLUMN in the file that defines the fixed atoms:

1.00 == FIXED  
0.00 == NOT FIXED

# Fixed Atoms

. . .

```
# restraints
```

```
fixedAtoms on
```

```
fixed REMARK original generated coordinate pdb file
```

```
fixed ATOM File 1 OH2 TIP3 input/fixed_pdb_atoms.pdb 0.000 0.000 0.000 0.00 0.00 TIP3 O
```

```
fixed ATOM Col 2 H1 TIP3 0 1.000 0.000 0.000 0.00 0.00 TIP3 H
```

```
ATOM 3 H2 TIP3 0 -0.326 0.946 0.000 0.00 0.00 TIP3 H
```

```
END
```

. . .

```
# run control
```

```
minimize 1000
```

These specific commands:

(1) apply fixed atoms “on”

(2) instruct forces to be included FROM these atoms

OCC. BETA

(3) define the file that is used to tell NAMD which atoms are fixed

(4) define the COLUMN in the file that defines the fixed atoms:

1.00 == FIXED  
0.00 == NOT FIXED

# Fixed Atoms III

OCC. BETA

```
REMARK original generated coordinate pdb file
ATOM      1  OH2  TIP3      0      0.000  0.000  0.000  0.00  1.00  TIP3  O
ATOM      2  H1  TIP3      0      1.000  0.000  0.000  0.00  0.00  TIP3  H
ATOM      3  H2  TIP3      0     -0.326  0.946  0.000  0.00  0.00  TIP3  H
END
```

The PDB files that “we” have created so far have not specifically done anything to the “beta” (or “occupancy”) fields.

Ways to alter the field (all protein atoms, side-chains, backbone, etc.):

- (1) by hand or via clever unix command-line scripting (grep/sed/awk etc.)
- (2) using a interpretative language using stand-alone script (python, tcl, perl)
- (3) using other open source tools : ‘build utilities’
- (4) **tcl within VMD** (python with VMD is an option)

Useful in MANY ways!

# Tangent . . . tcl & vmd

Steps WE are going to take:

- (1) Open `ionized_solvated_proteinII.pdb` (hydrolyse in water box + 6 CLA)
- (2) Set all BETA values -> 0.00
- (3) Set **all protein atoms** BETA values -> 1.00 then save coords to a new PDB file (**fixed\_protein\_atoms.pdb**) ; reset all BETA -> 0.00
- (4) Set **all backbone protein atoms** BETA values -> 1.00 then save coords to a new PDB file (**fixed\_backbone\_atoms.pdb**) ; reset all BETA -> 0.00
- (5) Set **all alpha helix AND beta sheet and backbone protein atoms** BETA values -> 1.00 then save coords to a new PDB file (**fixed\_secondary\_structure\_atoms.pdb**)

THREE SPECIFIC PDB FILES FOR SUCCESSIVE MINIMIZATION RUNS

only BETA is used from  
these files

# New PDB I

With the **ionized\_solvated\_proteinll.pdb** file open in VMD . . . at the console type:

```
vmd > set everyone [atomselect top all]  
atomselect32
```

```
vmd > $everyone set beta 0.00
```

```
vmd > set sell [atomselect top "protein"]  
atomselect33
```

```
vmd > $sell set beta 1.00
```

```
vmd > $everyone writepdb fixed_protein_atoms.pdb
```

```
Info) Opened coordinate file fixed_protein_atoms.pdb for writing.
```

```
Info) Finished with coordinate file fixed_protein_atoms.pdb.
```

# New PDB I

With the **ionized\_solvated\_proteinll.pdb** file open in VMD . . . at the console type:

```
vmd > set everyone [atomselect top all]
atomselect32
```

```
vmd > $everyone set beta 0.00
```

```
vmd > set sell [atomselect top "protein"]
atomselect33
```

```
vmd > $sell set beta 1.00
```

```
vmd > $everyone writepdb fixed_protein_atoms.pdb
```

```
Info) Opened coordinate file fixed_protein_atoms.pdb for writing.
```

```
Info) Finished with coordinate file fixed_protein_atoms.pdb.
```

```
. . .
ATOM      3159  HD21  ASN  H   245           6.667   19.714   34.439   1.00   1.00      HYDR  H
ATOM      3160  HD22  ASN  H   245           7.401   21.295   34.329   1.00   1.00      HYDR  H
ATOM      3161   OH2  TIP3W    1          50.279   45.432   40.650   1.00   0.00      WT1  O
ATOM      3162   H1   TIP3W    1          51.097   45.825   40.462   1.00   0.00      WT1  H
. . .
```

# New PDB II

Now, lets only restrain the backbone atoms . . . continuing (not starting over) . . . at the console type:

```
vmd > $everyone set beta 0.00
```

```
vmd > set sel2 [atomselect top "protein and backbone"]  
atomselect34
```

```
vmd > $sel2 set beta 1.00
```

```
vmd > $everyone writepdb fixed_backbone_atoms.pdb
```

```
Info) Opened coordinate file fixed_backbone_atoms.pdb for writing.
```

```
Info) Finished with coordinate file fixed_backbone_atoms.pdb.
```

# New PDB II

Now, lets only restrain the backbone atoms . . . continuing (not starting over) . . . at the console type:

```
vmd > $everyone set beta 0.00
```

```
vmd > set sel2 [atomselect top "protein and backbone"]  
atomselect34
```

```
vmd > $sel2 set beta 1.00
```

```
vmd > $everyone writepdb fixed_backbone_atoms.pdb
```

```
Info) Opened coordinate file fixed_backbone_atoms.pdb for writing.
```

```
Info) Finished with coordinate file fixed_backbone_atoms.pdb.
```

```
. . .
```

ATOM	38	N	GLY	H	18	38.970	21.537	14.910	1.00	1.00	HYDR	N
ATOM	39	HN	GLY	H	18	39.749	21.183	15.426	1.00	0.00	HYDR	H
ATOM	40	CA	GLY	H	18	39.254	22.674	14.040	1.00	1.00	HYDR	C
ATOM	41	HA1	GLY	H	18	38.753	22.465	13.105	1.00	0.00	HYDR	H
ATOM	42	HA2	GLY	H	18	40.331	22.734	13.964	1.00	0.00	HYDR	H
ATOM	43	C	GLY	H	18	38.747	24.007	14.556	1.00	1.00	HYDR	C
ATOM	44	O	GLY	H	18	38.573	24.944	13.781	1.00	1.00	HYDR	O

```
. . .
```

# New PDB III

Now, lets only restrain the backbone atoms on conserved secondary structure elements. . . continuing (not starting over) . . . at the console type:

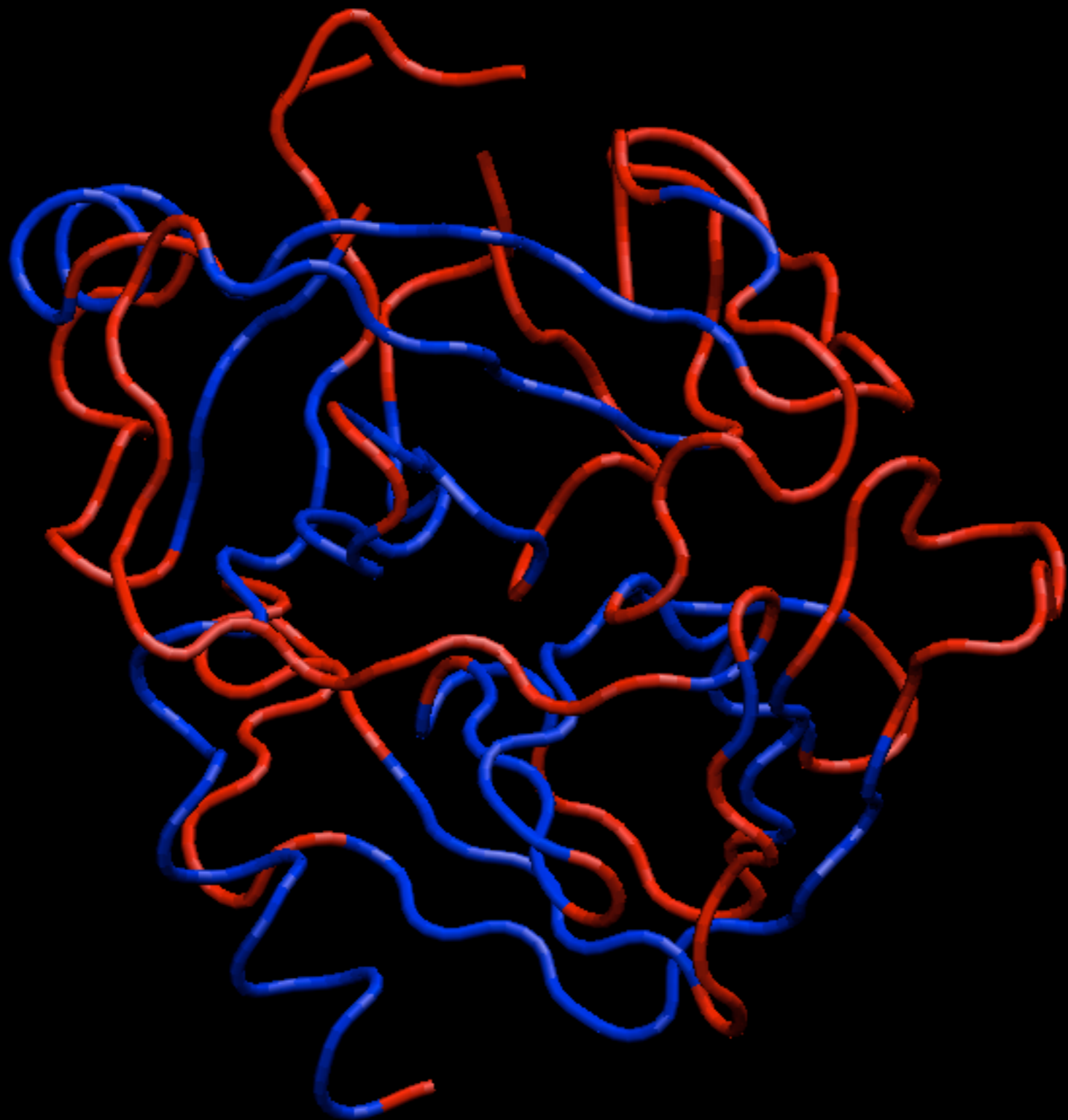
```
vmd > $everyone set beta 0.00
```

```
vmd > set sel3 [atomselect top "((protein and helix) or (protein and  
betasheet)) and backbone"]  
atomselect35
```

```
vmd > $sel3 set beta 1.00
```

```
vmd > $everyone writepdb fixed_secondary_structure_atoms.pdb
```

```
Opened coordinate file fixed_secondary_structure_atoms.pdb for writing.  
Info) Finished with coordinate file fixed_secondary_structure_atoms.pdb.
```



# MINIMIZATION in steps . . .

1. Restrain protein atoms (let waters/ions relax)
2. Restrain protein backbone (let side-chains and waters/ions relax)
3. Restrain secondary structure elements (let loops, etc. relax)
4. No restraints

Let's look at the successive NAMD input files (highlighting changes)

# NAMD INPUT FILE: smin0

```
# sample NAMD configuration file for Minimization w/ fixed atoms, PBC, PME

# molecular system
coordinates      output/ionized_solvated_proteinII.pdb
structure       output/ionized_solvated_proteinII.psf
temperature     300

# restraints

fixedAtoms      on
fixedAtomsForces on
fixedAtomsFile  output/fixed_protein_atoms.pdb
fixedAtomsCol   B

# force field
paratypecharmm on
parameters     /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude         scaled1-4
1-4scaling     1.0
switching      on
switchdist     10
cutoff         12

# pbc

cellBasisVector1 71.0 0.0 0.0
cellBasisVector2 0.0 67.0 0.0
cellBasisVector3 0.0 0.0 77.0
```



```
fixedAtomsFile      output/fixed_protein_atoms.pdb
fixedAtomsCol       B

# force field
paratypecharmm      on
parameters          /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude             scaled1-4
1-4scaling          1.0
switching           on
switchdist          10
cutoff              12

# pbc

cellBasisVector1    71.0 0.0 0.0
cellBasisVector2    0.0 67.0 0.0
cellBasisVector3    0.0 0.0 77.0

# electrostatics

PME                 on
PMEGridSpacing      1.0

# output

outputname          output/smin_fixed_protein
binaryoutput        no

# run control

minimize            1000
```



## USAGE (unix prompt):

### Start a run:

```
>namd2 smin0 >& smin0.out &
```

### Check status:

```
>tail -90f smin0.out
```

### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 smin0 >& smin0.out &
```



# NAMD INPUT FILE II: smin1

```
# sample NAMD configuration file for Minimization w/ fixed atoms, PBC, PME

# molecular system
coordinates      output/smin_fixed_protein.coor
structure        output/ionized_solvated_proteinII.psf
temperature      300

# restraints

fixedAtoms      on
fixedAtomsForces on
fixedAtomsFile output/fixed_backbone_atoms.pdb
fixedAtomsCol  B

# force field
paratypecharmm  on
parameters      /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude         scaled1-4
1-4scaling      1.0
switching       on
switchdist      10
cutoff          12

# pbc

cellBasisVector1 71.0 0.0 0.0
cellBasisVector2 0.0 67.0 0.0
cellBasisVector3 0.0 0.0 77.0
```



```
fixedAtomsFile      output/fixed_backbone_atoms.pdb
fixedAtomsCol       B

# force field
paratypecharmm      on
parameters          /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude             scaled1-4
1-4scaling          1.0
switching           on
switchdist          10
cutoff              12

# pbc

cellBasisVector1    71.0 0.0 0.0
cellBasisVector2    0.0 67.0 0.0
cellBasisVector3    0.0 0.0 77.0

# electrostatics

PME                 on
PMEGridSpacing      1.0

# output

outputname          output/smin_fixed_backbone
binaryoutput        no

# run control

minimize            1000
```



## USAGE (unix prompt):

### Start a run:

```
>namd2 smin1 >& smin1.out &
```

### Check status:

```
>tail -90f smin1.out
```

### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 smin1 >& smin1.out &
```



# NAMD INPUT FILE III: smin2

```
# sample NAMD configuration file for Minimization w/ fixed atoms, PBC, PME

# molecular system
coordinates      output/smin_fixed_backbone.coor
structure        output/ionized_solvated_proteinII.psf
temperature      300

# restraints

fixedAtoms      on
fixedAtomsForces on
fixedAtomsFile output/fixed_secondary_structure_atoms.pdb
fixedAtomsCol  B

# force field
paratypecharmm  on
parameters      /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude          scaled1-4
1-4scaling       1.0
switching        on
switchdist       10
cutoff           12

# pbc

cellBasisVector1 71.0 0.0 0.0
cellBasisVector2 0.0 67.0 0.0
cellBasisVector3 0.0 0.0 77.0
```

```
fixedAtomsFile      output/fixed_secondary_structure_atoms.pdb
fixedAtomsCol       B

# force field
paratypecharmm      on
parameters          /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude             scaled1-4
1-4scaling          1.0
switching           on
switchdist          10
cutoff              12

# pbc

cellBasisVector1    71.0 0.0 0.0
cellBasisVector2    0.0 67.0 0.0
cellBasisVector3    0.0 0.0 77.0

# electrostatics

PME                 on
PMEGridSpacing      1.0

# output

outputname          output/smin_fixed_secondary
binaryoutput        no

# run control

minimize            1000
```

## USAGE (unix prompt):

### Start a run:

```
>namd2 smin2 >& smin2.out &
```

### Check status:

```
>tail -90f smin2.out
```

### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 smin2 >& smin2.out &
```

# NAMD INPUT FILE IV: smin3

```
# sample NAMD configuration file for Minimization w/ fixed atoms, PBC, PME

# molecular system
coordinates      output/smin_fixed_secondary.coor
structure        output/ionized_solvated_proteinII.psf
temperature      300

# restraints

fixedAtoms      off
#fixedAtomsForces    on
#fixedAtomsFile     output/fixed_secondary_structure_atoms.pdb
#fixedAtomsCol      B

# force field
paratypecharmm  on
parameters      /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude          scaled1-4
1-4scaling       1.0
switching        on
switchdist       10
cutoff           12

# pbc

cellBasisVector1 71.0 0.0 0.0
cellBasisVector2 0.0 67.0 0.0
cellBasisVector3 0.0 0.0 77.0
```



```
#fixedAtomsFile      output/fixed_secondary_structure_atoms.pdb
#fixedAtomsCol       B

# force field
paratypecharmm       on
parameters           /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude              scaled1-4
1-4scaling           1.0
switching            on
switchdist           10
cutoff               12

# pbc

cellBasisVector1     71.0 0.0 0.0
cellBasisVector2     0.0 67.0 0.0
cellBasisVector3     0.0 0.0 77.0

# electrostatics

PME                  on
PMEGridSpacing       1.0

# output

outputname           output/smin3
binaryoutput         no

# run control

minimize             1000
```



## USAGE (unix prompt):

### Start a run:

```
>namd2 smin3 >& smin3.out &
```

### Check status:

```
>tail -90f smin3.out
```

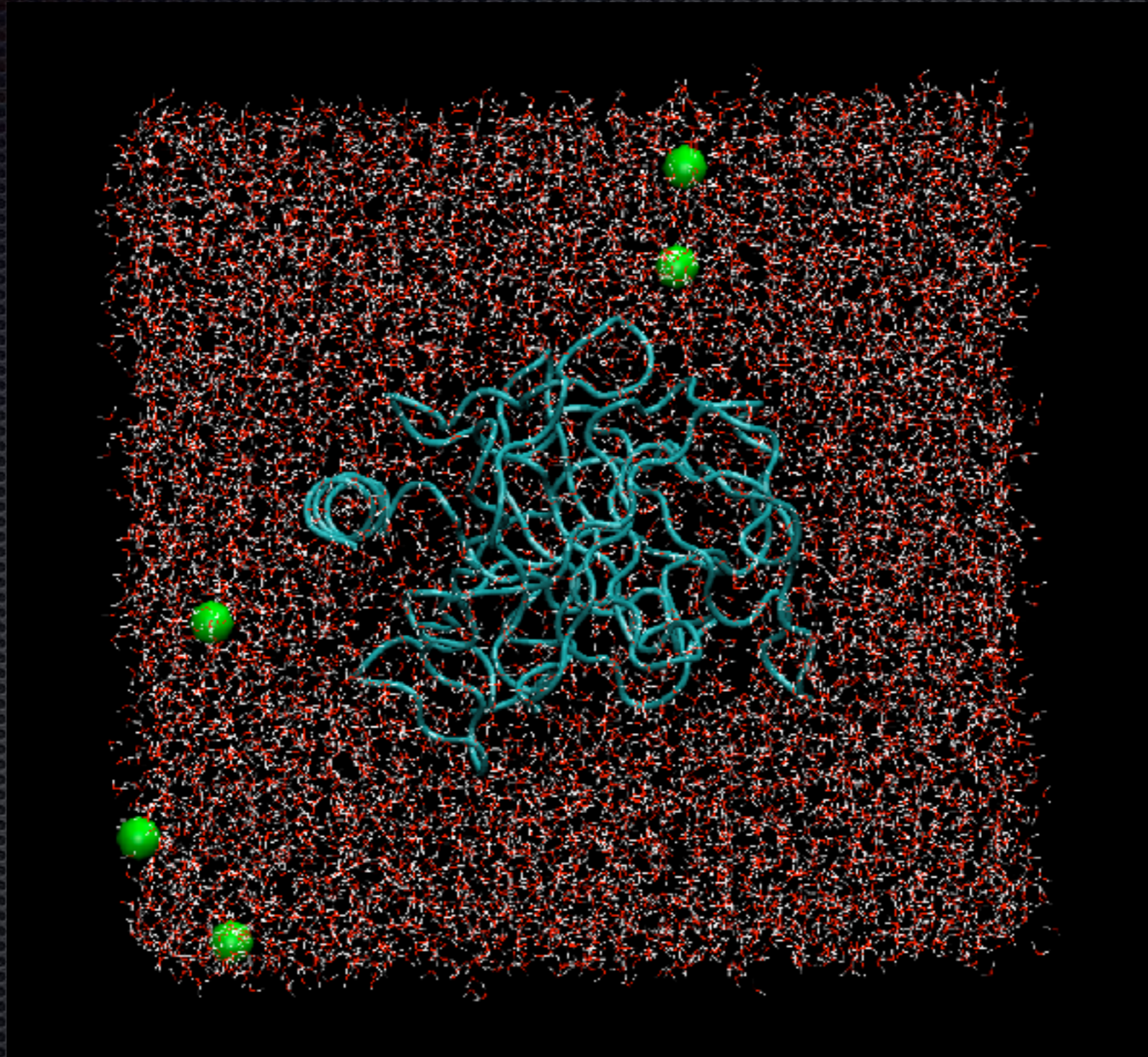
### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 smin3 >& smin3.out &
```



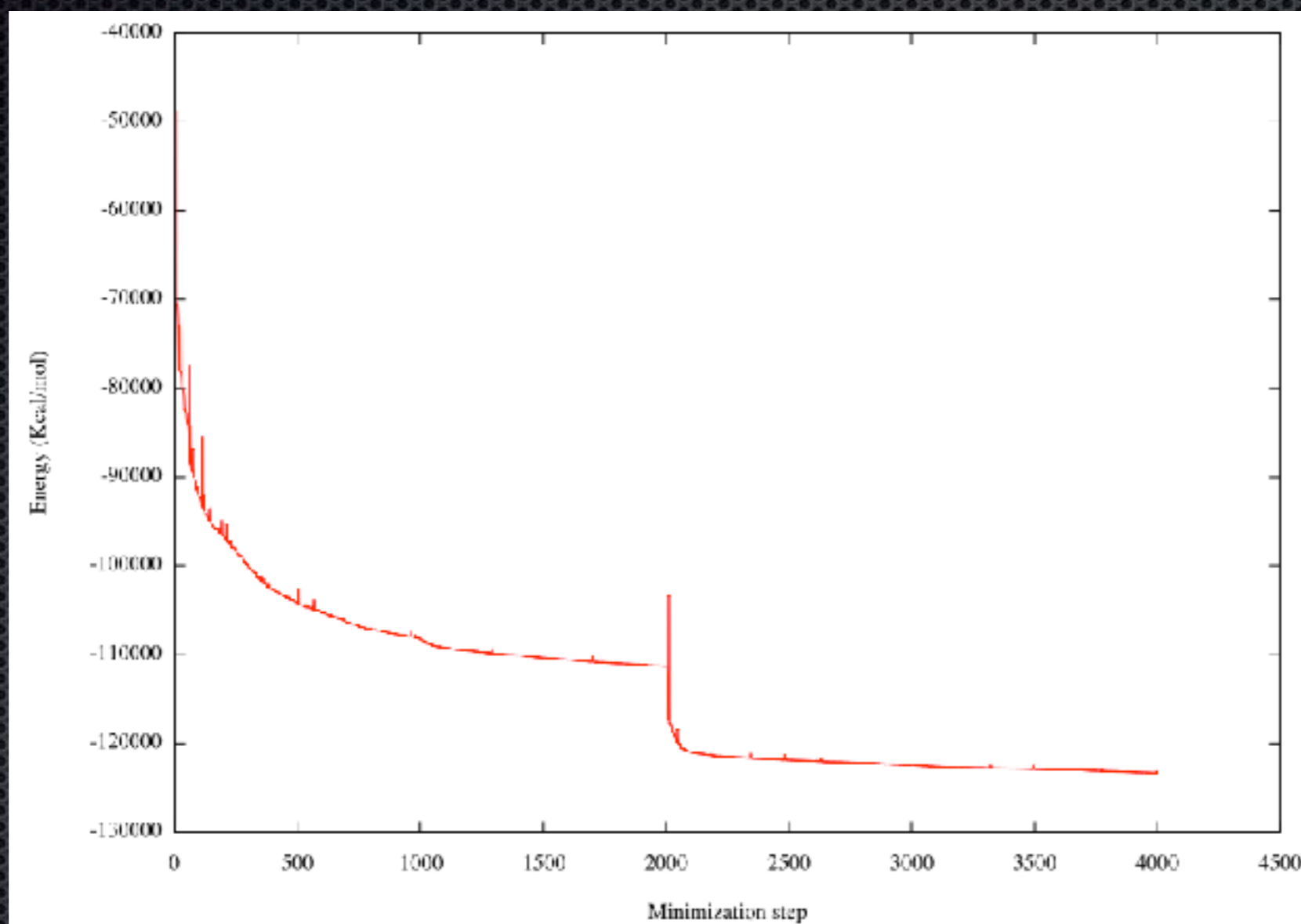
# MINIMIZATION: analysis

# MINIMIZATION: analysis



# MINIMIZATION: analysis

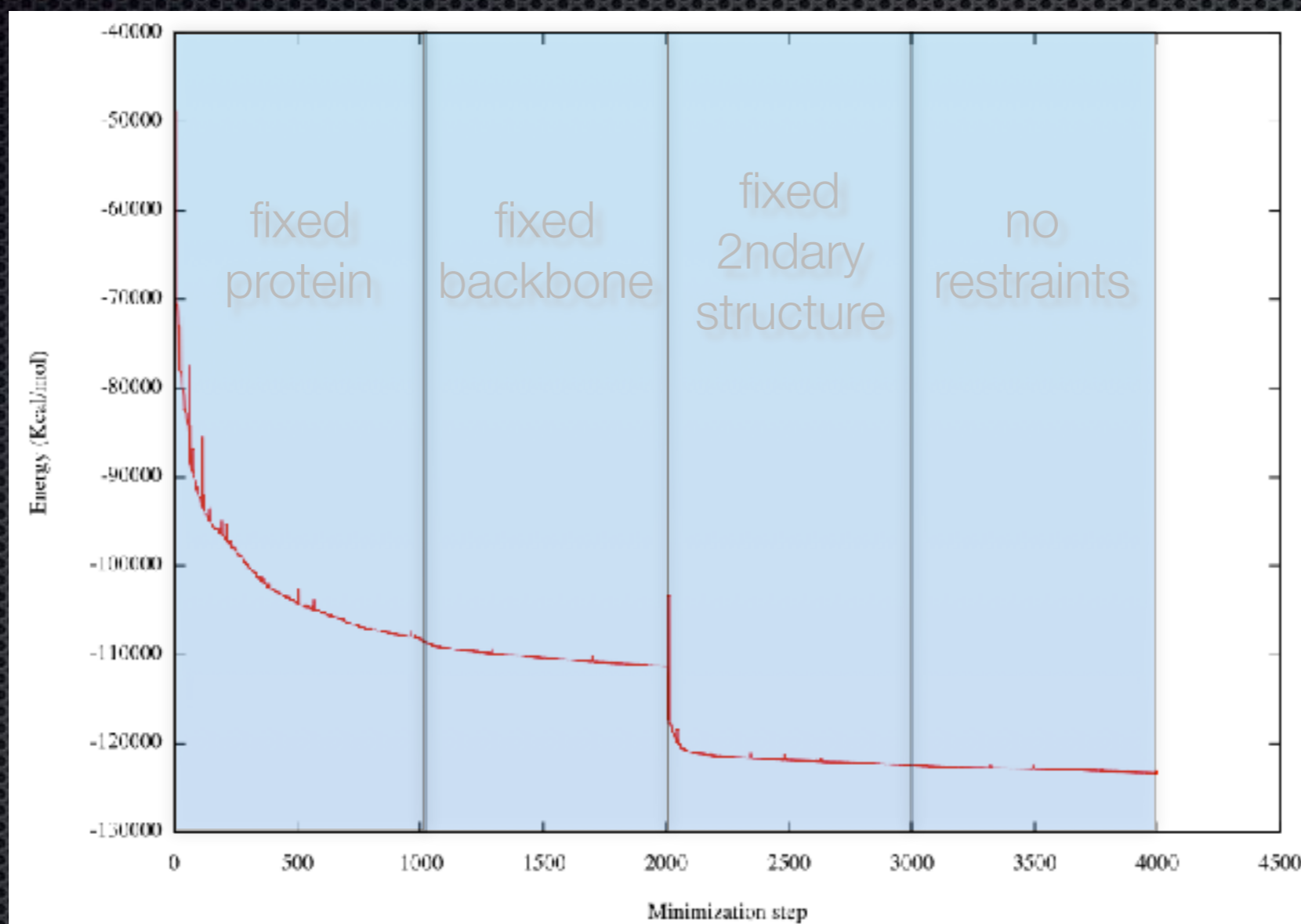
Plot created from the “output” files : (you will create these in your lab section later today)



**Cowboys: they sometimes SKIP the restrained minimizations  
... your mileage will vary.**

# MINIMIZATION: analysis

Plot created from the “output” files : (you will create these in your lab section later today)



**Cowboys: they sometimes SKIP the restrained minimizations  
... your mileage will vary.**

# DYNAMICS in steps . . .

Now, we have a new coordinate file (that NAMD “names”): `smin3.coor`

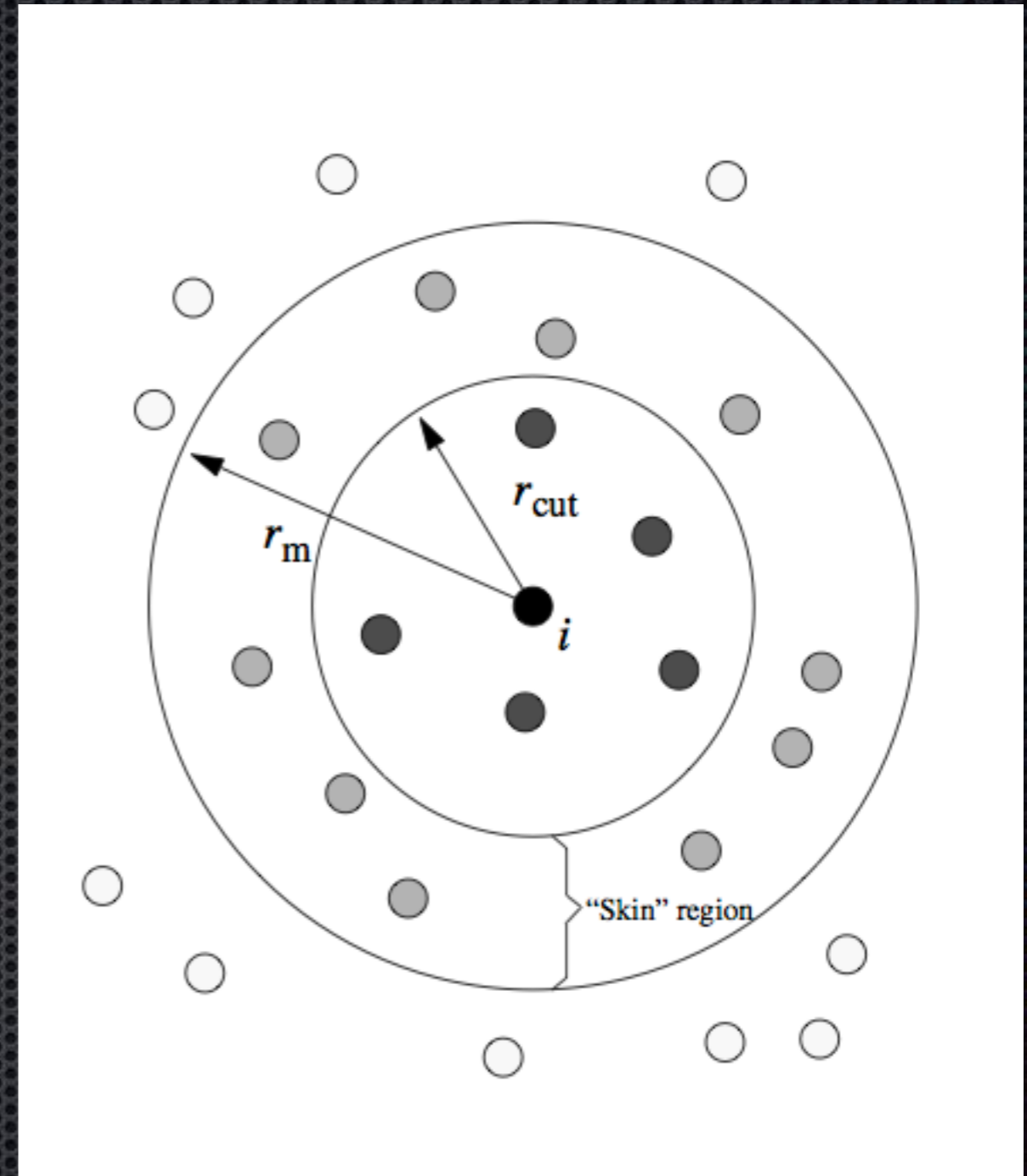
which has been minimized and is ready to go.

We are ready to start a molecular dynamics trajectory (time) so have to add some more commands to our input file.

- (1) Neighbor lists
- (2) Temperature control
- (3) Pressure control
- (4) Time steps & energy update cycle
- (5) Save restart & DCD files (coordinates, **velocities**, & cell dimensions)
- (6) IMD, screen printout stuff, & “run” command

# Neighbor Lists

```
. . .  
# neighbor lists  
  
pairlistdist      12  
margin            2  
  
. . .  
# run control  
  
run               10000
```



# Temperature Control

```
• • •  
# temperature control
```

```
langevin           on  
langevinDamping    1  
langevinTemp       300  
langevinHydrogen   no
```

← Temperature to maintain

```
• • •  
# run control
```

```
run                10000
```

$$m_i \frac{d^2 x_i(t)}{dt^2} = \mathbf{F}_i\{x_i(t)\} - \gamma_i \frac{dx_i(t)}{dt} m_i + \mathbf{R}_i(t)$$

Here, two additional terms on the right hand side accompany the ordinary force the particle experiences. The second term represents a frictional damping that is applied to the particle with frictional coefficient  $\gamma_i m_i$ . The third term represents random forces which act on the particle (as a result of solvent interaction). These two terms are used to maintain particle kinetic energy to keep system temperature, for instance, at a constant value. Because an unnecessarily high damping constant can significantly slow the system's dynamics, one should always find the minimum `langevinDamping` coefficient sufficient to maintain the temperature. **A value of 1.0 is often a good starting point.**

# Pressure Control

• • •  
**# pressure control**

**#useGroupPressure**           **yes**

**#useFlexibleCell**           **no**

**#LangevinPiston**           **on**

**#LangevinPistonTarget**   **1.01325**           ← **Pressure (bar) to maintain**

**#LangevinPistonPeriod**   **100**

**#LangevinPistonDecay**     **50**

**#LangevinPistonTemp**     **300**               ← **Temperature to maintain**

• • •

**# run control**

**run**                           **10000**

# Time, MTS, & Shake

...

```
timestep          1.0  
stepscycle       20  
  
fullElectFrequency 4  
nonbondedFreq     2
```

```
# shake
```

```
rigidBonds        water
```

...

```
# run control
```

```
run               10000
```

Update frequency of Nearest Neighbor list



# Time, MTS, & Shake

...

```
timestep      2.0  
stepscycle   10  
  
fullElectFrequency  2  
nonbondedFreq      1
```

← Update frequency of Nearest Neighbor list

```
# shake
```

```
rigidBonds    all
```

← Required for 2 fs time step

...

```
# run control
```

```
run          10000
```

# Binary Data & Restart Files

. . .

```
# output
```

```
outputname      output/dyn0
binaryoutput    no
```

```
DCDfile         output/dyn0.dcd
```

```
DCDfreq        1000
```

```
DCDUnitCell     yes
```

```
#velDCDfile     output/dyn0.vel.dcd
```

```
#velDCDfreq     1000
```

```
#XstFile        output/dyn0.xst
```

```
#XstFreq        1000
```

```
wrapAll        on
```

```
# restart files
```

```
binaryrestart   no
```

```
restartsave     no
```

```
restartname     output/dyn0.rest.pdb
```

```
restartfreq     10000
```

. . .

```
# run control
```

```
run            10000
```

← ATOM VELOCITIES: C(t) → Spectra

← UNIT CELL SIZE PER STEP (NPT)

← USED IF YOU EXPECT A SHUTDOWN OF SOME KIND BEFORE THE END OF A RUN . . . easy restart

# Interactive MD (IMD)

. . .

```
# interactive MD
```

```
IMDon      yes
IMDport    1085
IMDfreq    10
IMDwait    no
```

PORT



127.0.0.1

```
# screen output
```

```
outputenergies 10
outputtiming     100
```

or via internet  
with real IP

. . .

```
# run control
```

```
run          10000
```

ONE port per MD  
run you want to  
attach to . . .

slows computation!!!

# NAMD INPUT FILE V: dyn0

# NAMD configuration file for **INITIAL Dynamics** w/ fixed atoms, PBC, PME

**timestep** 1.0  
**stepspercycle** 20

**fullElectFrequency** 4  
**nonbondedFreq** 2

# molecular system  
coordinates output/**smin3.coor**  
structure output/ionized\_solvated\_proteinII.psf  
temperature 300

## # restraints

**fixedAtoms** off  
**fixedAtomsForces** on  
**fixedAtomsFile** output/**fixed\_secondary\_structure\_atoms.pdb**  
**fixedAtomsCol** B

# force field  
paratypecharmm on  
parameters /home/mdschool/toppar/par\_all27\_prot\_na.inp

# approximations  
exclude scaled1-4  
1-4scaling 1.0  
switching on  
switchdist 10  
cutoff 12



```
# NAMD configuration file for INITIAL Dynamics w/ fixed atoms, PBC, PME
```

```
timestep 1.0  
stepspercycle 20
```

```
fullElectFrequency 4  
nonbondedFreq 2
```

```
# molecular system  
coordinates output/smin3.coor  
structure output/ionized_solvated_proteinII.psf  
temperature 300
```

```
# restraints
```

```
fixedAtoms off  
fixedAtomsForces on  
fixedAtomsFile output/fixed_secondary_structure_atoms.pdb  
fixedAtomsCol B
```

```
# force field  
paratypecharmm on  
parameters /home/mdschool/toppar/par_all27_prot_na.inp
```

```
# approximations  
exclude scaled1-4  
1-4scaling 1.0  
switching on  
switchdist 10  
cutoff 12
```



```
PME on
PMEGridSpacing 1.0

# output

outputname output/dyn0
binaryoutput no

DCDfile output/dyn0.dcd
DCDfreq 1000
DCDUnitCell yes
#velDCDfile output/dyn0.vel.dcd
#velDCDfreq 1000
#XstFile output/dyn0.xst
#XstFreq 1000
wrapAll on

# restart files

binaryrestart no
restartsave no
restartname output/dyn0.rest.pdb
restartfreq 10000

# interactive MD

IMDon yes
IMDport 1085
IMDfreq 10
IMDwait no

# run control

run 10000
```



## USAGE (unix prompt):

### Start a run:

```
>namd2 dyn0 >& dyn0.out &
```

### Check status:

```
>tail -90f dyn0.out
```

### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 dyn0 >& dyn0.out &
```



# DYNAMICS “restarting”

Due to file sizes, queuing rules, chance of jobs stopping, . . . one needs a method to continue a simulation without a discontinuity in positions, velocities, boxsizes, etc.

Let's write a new input file to handle this.

- (1) Initial time step
- (2) Load final coordinates, velocities, box-dimensions from previous run
- (3) Comment out “Temperature” & PBC cell definition
- (4) Implement pressure control
- (5) Save restart & DCD files (coordinates, **velocities**, & **cell dimensions**)
- (6) IMD **OFF**, screen printout stuff, & “run” command (**1 ns**)

# NAMD INPUT FILE V: dyn1

# NAMD configuration file for **CONTINUING Dynamics** w/ fixed atoms, PBC, PME

```
timestep                1.0
stepspcycle             20
firsttimestep         10000

fullElectFrequency     4
nonbondedFreq          2

# molecular system
coordinates             output/dyn0.coor
velocities           output/dyn0.vel
extendedSystem       output/dyn0.xsc
structure              output/ionized_solvated_proteinII.psf
#temperature            300

# force field
paratypecharmm         on
parameters             /home/mdschool/toppar/par_all27_prot_na.inp

# approximations
exclude                scaled1-4
1-4scaling             1.0
switching              on
switchdist             10
cutoff                 12

# neighbor lists
listdist              12
```

```
# NAMD configuration file for CONTINUING Dynamics w/ fixed atoms, PBC, PME
```

```
timestep          1.0  
stepspcycle       20  
firsttimestep   10000
```

```
fullElectFrequency 4  
nonbondedFreq      2
```

```
# molecular system
```

```
coordinates        output/dyn0.coor  
velocities       output/dyn0.vel  
extendedSystem  output/dyn0.xsc  
structure          output/ionized_solvated_proteinII.psf  
#temperature    300
```

```
# force field
```

```
paratypecharmm    on  
parameters        /home/mdschool/toppar/par_all27_prot_na.inp
```

```
# approximations
```

```
exclude           scaled1-4  
1-4scaling        1.0  
switching         on  
switchdist       10  
cutoff            12
```

```
# neighbor lists
```

```
listdist          12
```

# restart files

binaryrestart no  
restartsave no  
restartname output/dyn1.rest.pdb  
restartfreq 10000

# interactive MD

IMDon no  
IMDport 1085  
IMDfreq 10  
IMDwait no

# screen output

outputenergies 10  
outputtiming 100

# run control

run 1000000

## USAGE (unix prompt):

### Start a run:

```
>namd2 dyn1 >& dyn1.out &
```

### Check status:

```
>tail -90f dyn1.out
```

### Multi-processor run:

```
>charmrun ++local ++p 2 /usr/local/bin/namd/namd2 dyn1 >& dyn1.out &
```

# NAMD OUTPUT (DYN.)

```
Charm++: scheduler running in netpoll mode.
Info: NAMD 2.6 for MacOSX-i686
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 Phillips et al., J. Comp. Chem. 26:1781-1802 (2005)
Info: in all publications reporting results obtained with NAMD.
Info:
Info: Based on Charm++/Converse 50900 for net-darwin-x86-smp
Info: Built Wed Aug 30 12:54:31 CDT 2006 by jim on juneau.ks.uiuc.edu
Info: 1 NAMD 2.6 MacOSX-i686 4 ocho.ncnr.nist.gov curtisj
Info: Running on 4 processors.
Info: 648656 kB of memory in use.
Info: Memory usage based on ps
Info: Configuration file is dyn0
TCL: Suspending until startup complete.
Warning: The following variables were set in the
Warning: configuration file but were not needed
Warning: fixedAtomsForces
Warning: fixedAtomsFile
Warning: fixedAtomsCol
Info: SIMULATION PARAMETERS:
Info: TIMESTEP 1
Info: NUMBER OF STEPS 0
Info: STEPS PER CYCLE 20
Info: PERIODIC CELL BASIS 1 71 0 0
Info: PERIODIC CELL BASIS 2 0 67 0
Info: PERIODIC CELL BASIS 3 0 0 77
Info: PERIODIC CELL CENTER 0 0 0
```

```
Info: MAX ITERATIONS : 500
Info: RIGID WATER USING SETTLE ALGORITHM
Info: NONBONDED FORCES EVALUATED EVERY 2 STEPS
Info: RANDOM NUMBER SEED      1308949783
Info: USE HYDROGEN BONDS?     NO
Info: COORDINATE PDB          output/smin3.coor
Info: STRUCTURE FILE         output/ionized_solvated_proteinII.psf
Info: PARAMETER file: CHARMM format! Info: PARAMETERS           /Users/
curtisj/research/toppar/par_all27_prot_na.inp
Info: USING ARITHMETIC MEAN TO COMBINE L-J SIGMA PARAMETERS
Info: SUMMARY OF PARAMETERS:
Info: 253 BONDS
Info: 628 ANGLES
Info: 1057 DIHEDRAL
Info: 73 IMPROPER
Info: 0 CROSSTERM
Info: 130 VDW
Info: 0 VDW_PAIRS
Info: *****
Info: STRUCTURE SUMMARY:
Info: 32067 ATOMS
Info: 22464 BONDS
Info: 15401 ANGLES
Info: 8444 DIHEDRALS
Info: 569 IMPROPER
Info: 0 CROSSTERMS
Info: 0 EXCLUSIONS
Info: 30463 RIGID BONDS
Info: 65738 DEGREES OF FREEDOM
Info: 11238 HYDROGEN GROUPS
Info: TOTAL MASS = 196616 amu
Info: TOTAL CHARGE = 1.78814e-06 e
```

# NAMD OUTPUT (DYN.) II

. . .

**TIMING: 10000 CPU: 1666.31, 0.166343/step Wall: 1607.39, 0.163882/step, 0 hours remaining, 711036 kB of memory in use.**

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

<b>ENERGY: 10000</b>	650.0702	1728.6069	1094.7965	104.4904	
-105158.6301	8258.5326	0.0000	0.0000	19724.1503	
-73597.9832	301.9758	-73561.6372	-73556.0246	302.8336	
-17532.4835	-1227.9700	366289.0000	-1103.7734	-1101.8505	

WRITING EXTENDED SYSTEM TO RESTART FILE AT STEP 10000  
WRITING COORDINATES TO DCD FILE AT STEP 10000  
WRITING COORDINATES TO RESTART FILE AT STEP 10000  
FINISHED WRITING RESTART COORDINATES  
WRITING VELOCITIES TO RESTART FILE AT STEP 10000  
FINISHED WRITING RESTART VELOCITIES  
WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 10000  
WRITING COORDINATES TO OUTPUT FILE AT STEP 10000  
CLOSING COORDINATE DCD FILE  
WRITING VELOCITIES TO OUTPUT FILE AT STEP 10000

=====

WallClock: 1619.131104 CPUtime: 1678.077148 Memory: 707260 kB

# SUMMARY

**WHEW!!! So many details, AGAIN!!!**

You observed how to build up a system from “minimized” vacuum coordinates to a solvated box running NPT MD.

### Energy Minimization

run name	<input type="text" value="run_0"/>		
reference pdb	<input type="button" value="Browse..."/> No file selected.	or <input type="button" value="Browse server"/> Server: no_project_specified/dyn0.pdb	
input filename (dcd or pdb)	<input type="button" value="Browse..."/> No file selected.	or <input type="button" value="Browse server"/> Server: no_project_specified/dyn0.pdb	
PSF file name	<input type="button" value="Browse..."/> No file selected.	or <input type="button" value="Browse server"/> Server: no_project_specified/ionized_solvated_proteinII.psf	
output file name (dcd)	<input type="text" value="min_run_0.dcd"/>		
number of processors	<input type="text" value="8"/>		
keep run output files	<input type="text" value="yes"/>		
run type	<input type="text" value="supply input file"/>		
namd input file	<input type="button" value="Browse..."/> dyn1	or <input type="button" value="Browse server"/> Local: dyn1	
check box to enter restart files	<input checked="" type="checkbox"/>		
velocity restart file	<input type="button" value="Browse..."/> dyn0.vel	or <input type="button" value="Browse server"/> Local: dyn0.vel	
extended system restart file	<input type="button" value="Browse..."/> dyn0.xsc	or <input type="button" value="Browse server"/> Local: dyn0.xsc	

DOCS FEEDBACK

# MD Analysis and NS

SASSIE CCP-SAS Workshop

January 23-25, 2017

ISIS Neutron and Muon Source

Rutherford Appleton Laboratory UK

# Barriers: BUILD; EQUILIBRATE; PROPAGATE; ANALYZE

What software package(s) and force-fields do I use?

**Starting structure?**

How do I clean up the structure?

How do I set up a trajectory (**time** or space)?

**How do I calculate scattering observables correctly?**

# Overview

## Analyze NAMD trajectory

Checking validity of equilibration run  
Standard analysis

## Neutron Scattering (dynamics)

What is measured

## $I(q,t)$ & $S(q,w)$ from MD

Connecting MD trajectory (time) to  $S(q,w)$   
Multiple Time Origins (MTO)  
Examples

## FURTHER READING

# OUTPUT FILES I

Picking up from where we left off yesterday using our `ionized_solvated_proteinII` system (hydrolase)

1. NVT 10 ps
2. NPT 1 ns

*Focussing on the 1ns equilibration run, several files were written “output/”:*

<code>dyn1.dcd</code>	←	<code>Coordinates sampled over run (BINARY)</code>
<code>dyn1.coor</code>	←	<code>Final coordinates (PDB) [[ VMD: namdbin ]]</code>
<code>dyn1.vel</code>	←	<code>Final velocities (PDB)</code>
<code>dyn1.xst</code>	←	<code>Final PBC box dimensions</code>
<code>dyn1.xsc</code>	←	<code>All PBC box dimensions sampled over run</code>
<code>dyn1.rest.pdb.coor</code> <code>dyn1.rest.pdb.vel</code> <code>dyn1.rest.pdb.xsc</code>	←	<code>Restart coords., velocities, and cell dimensions</code>

# namdstats.tcl

Run log files (e. g. dyn1.out) exist for nearly all MD/MMC packages. **PYTHON**

NAMD has a useful TCL script for parsing their output files: **namdstats.tcl**

# namdstats.tcl

Run log files (e. g. dyn1.out) exist for nearly all MD/MMC packages. **PYTHON**

NAMD has a useful TCL script for parsing their output files: **namdstats.tcl**

**USAGE (VMD prompt):**

Source the script:

```
vmd > source namdstats.tcl
```

run:

```
vmd > data_avg dyn1.out
```

chose a subset of the trajectory (say step 1000 to the end of the run):

```
vmd > data_avg dyn1.out 1000 last
```

# namdstats.tcl

```
vmd > source namdstats.tcl
vmd > data_avg dyn1.out
Calculating averages...
CALCULATING DATA FROM TIMESTEP 10000 TO 68830:
BOND: 1104.51939235
ANGLE: 1706.60535144
DIHED: 1090.54615195
IMPRP: 106.136829793
ELECT: -358288.083821
VDW: 8887.13411618
BOUNDARY: 0.0
MISC: 0.0
KINETIC: 20020.0703061
TOTAL: -325373.071694
TEMP: 299.396949898
TOTAL2: -325319.666106
TOTAL3: -325319.288662
TEMPAVG: 300.127975612
PRESSURE: -19167.892317
GPRESSURE: -0.444661811693
VOLUME: 317515.797541
PRESSAVG: -2.03941170972
GPRESSAVG: 0.706860605031
```

# THERMO. DISTRIBUTIONS

Select subset of output file to analyze thermodynamic variables

**USAGE (VMD prompt):**

Source the script:

```
vmd > source namdstats.tcl
```

run:

```
vmd > data_time TEMP dyn1.out
```

chose a subset of the trajectory (say step 1000 to the end of the run):

```
vmd > data_time TEMP dyn1.out 1000 last
```

**Saves data to a file called TEMP.dat (VDW.dat, DIHED.dat, etc.)**

Figure 0

10 ps NVT : dyn0.out

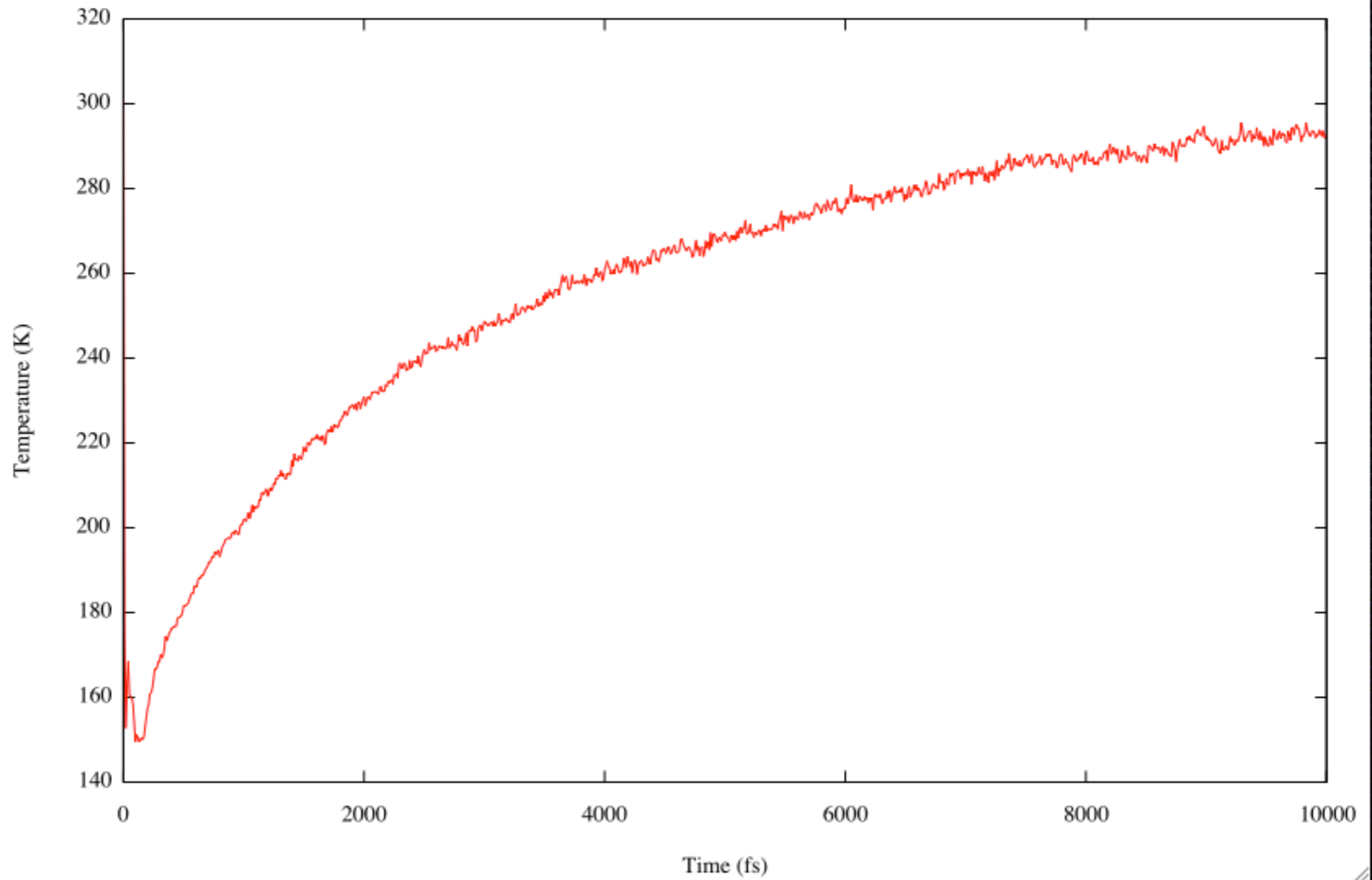
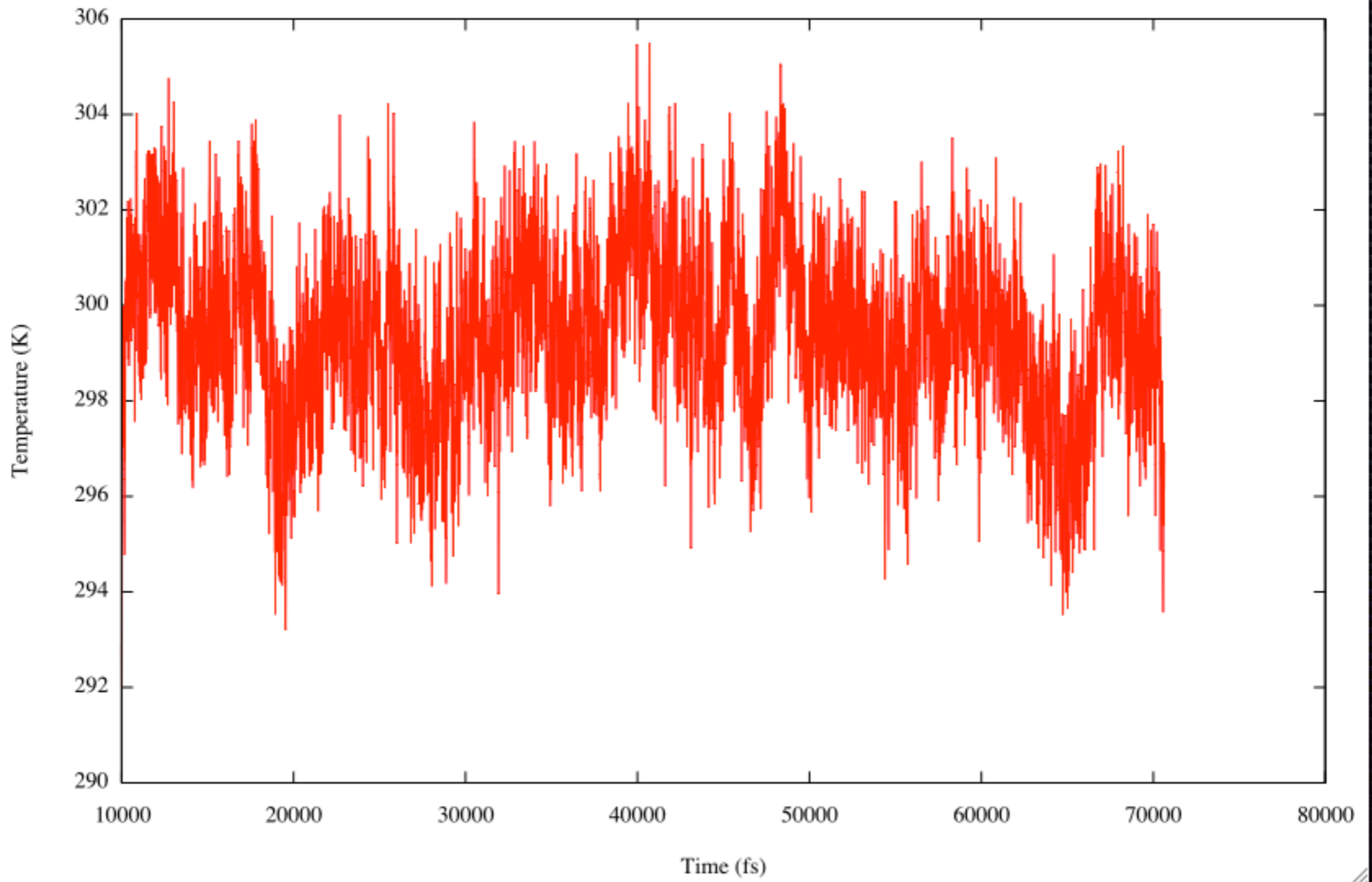
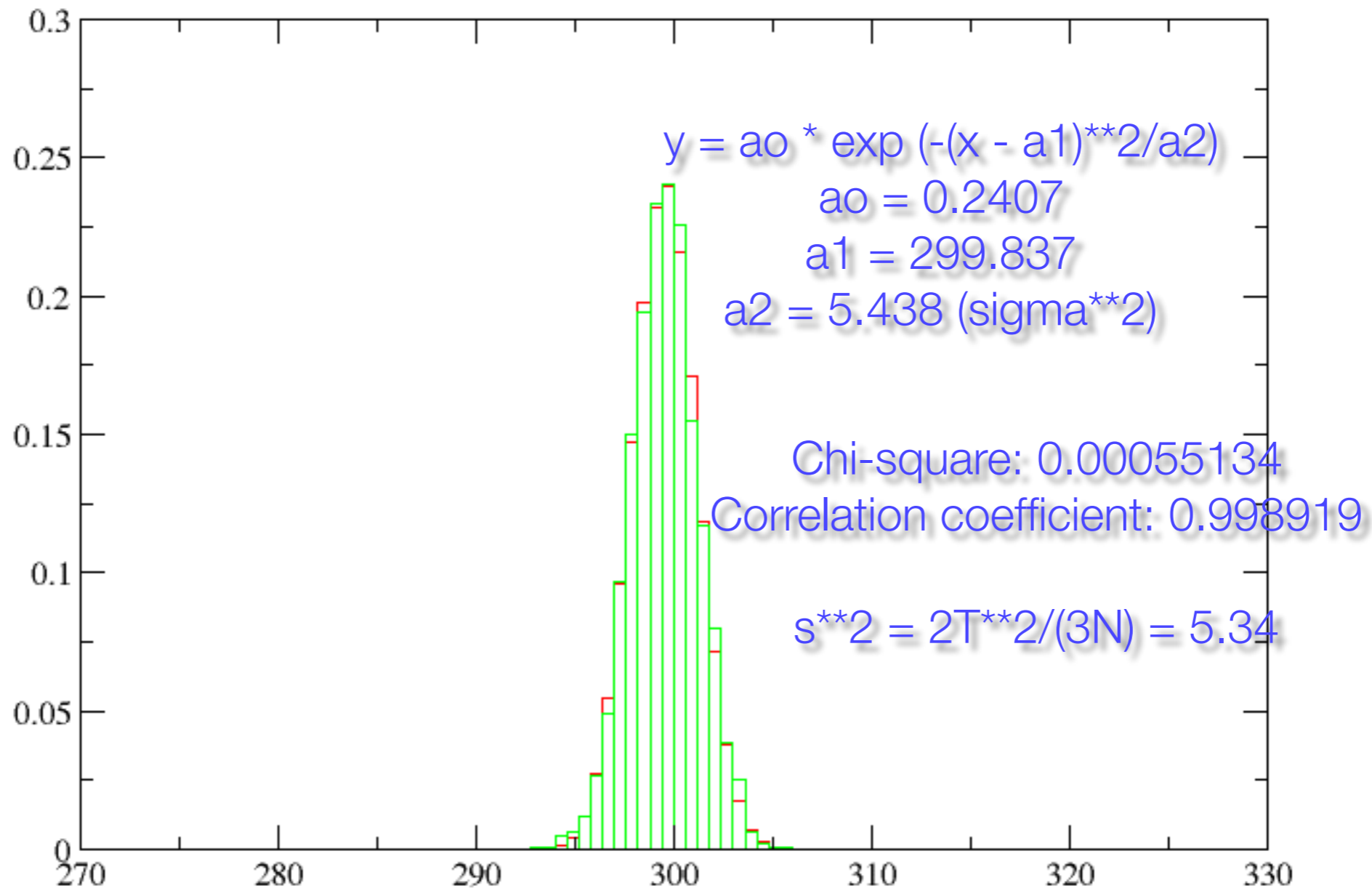


Figure 0

1 ns NPT: dyn1.out





# Protein conformational Eq.

Load `ionized_solvated_proteinII.psf` into VMD

Load `dyn1.dcd` into the same molecule

run the `rmsd.tcl` script within the vmd console

```
vmd > source rmsd.tcl
```

# Protein conformational Eq.

Load `ionized_solvated_protein.tps` into VMD

Load `dyn1.dcd` into the same molecule

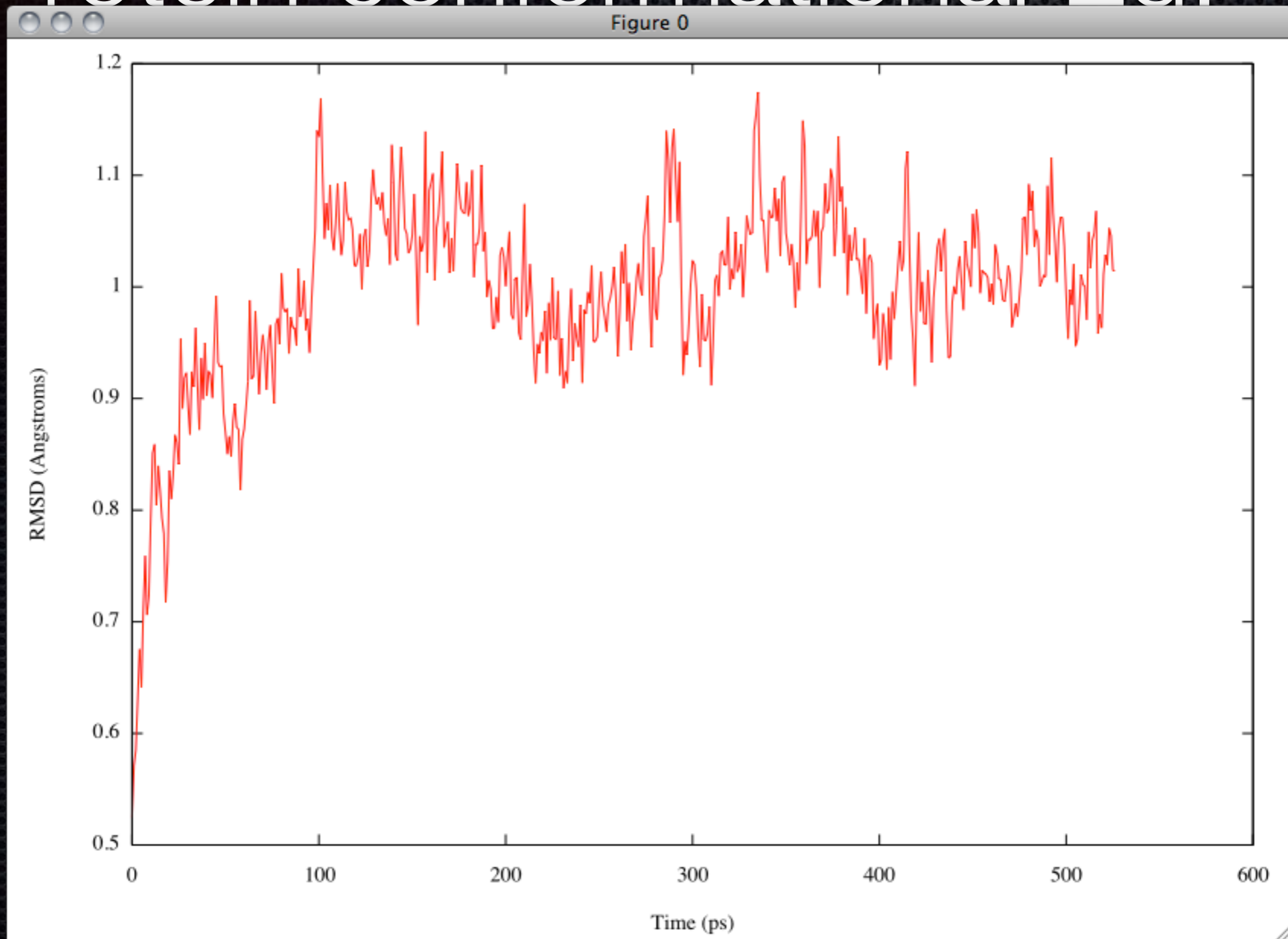
run the `rmsd.tcl` script within the vmd console

```
set outfile [open rmsd.dat w]
set nf [molinfo top get numframes]
set frame0 [atomselect top "protein and backbone and noh" frame 0]
set sel [atomselect top "protein and backbone and noh"]

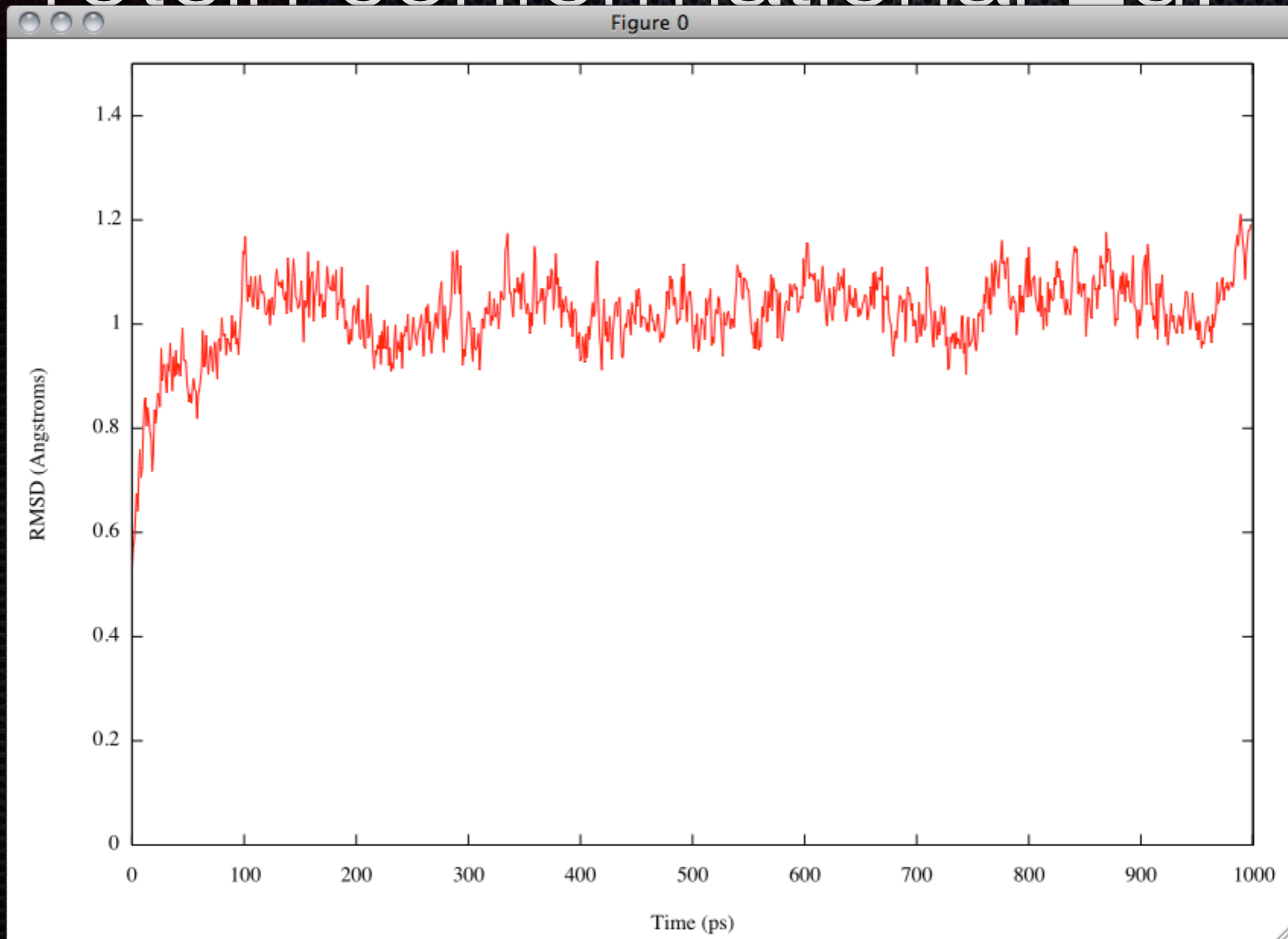
for { set i 1 } { $i <= $nf } { incr i } {
  $sel frame $i
  $sel move [measure fit $sel $frame0]
  puts $outfile "[measure rmsd $sel $frame0]"
}
close $outfile
```

`vmd > source rmsd.tcl`

# Protein conformational Eq.



# Protein conformational Eq.



# Protein Conformational Eq.

Load `ionized_solvated_proteinll.psf` into VMD

Load `dyn1.dcd` into the same molecule

Run the `residue_rmsd.tcl` script within the vmd console

```
vmd > source residue_rmsd.tcl  
vmd > set sel_resid [[atomselect top "protein and alpha"] get resid]  
vmd > rmsd_residue_over_time top $sel_resid
```

which will write a file “`residue_rmsd.dat`” to your directory AND add the residue rmsd values to a “User” field so you can view on the molecule

Hot loops, colder core . . .



# Equilibration notes

**Thermodynamic variables will fluctuate about a mean value.**

**Biological molecules will fluctuate about a mean structure (1-3 Å)**

**Once the thermodynamic variables and structure are fluctuating normally, you are ready to continue the trajectory to collect a “production run”.**

**These production runs vary in time (1 ns to 1 μs) depending on what observable you wish to predict.**

**One should do a similar analysis of the complete trajectory to validate the system prior to further analysis.**

# Points to Remember

## Wrapping -->

If you are **never** going to be interested in motions of the solvent then it is okay to wrap coordinates prior to saving to file (\*.coor, \*.dcd).

You can “wrap” the coordinates afterwards via boxsize information from DCD header or \*.xsc files.

## Energy loss -->

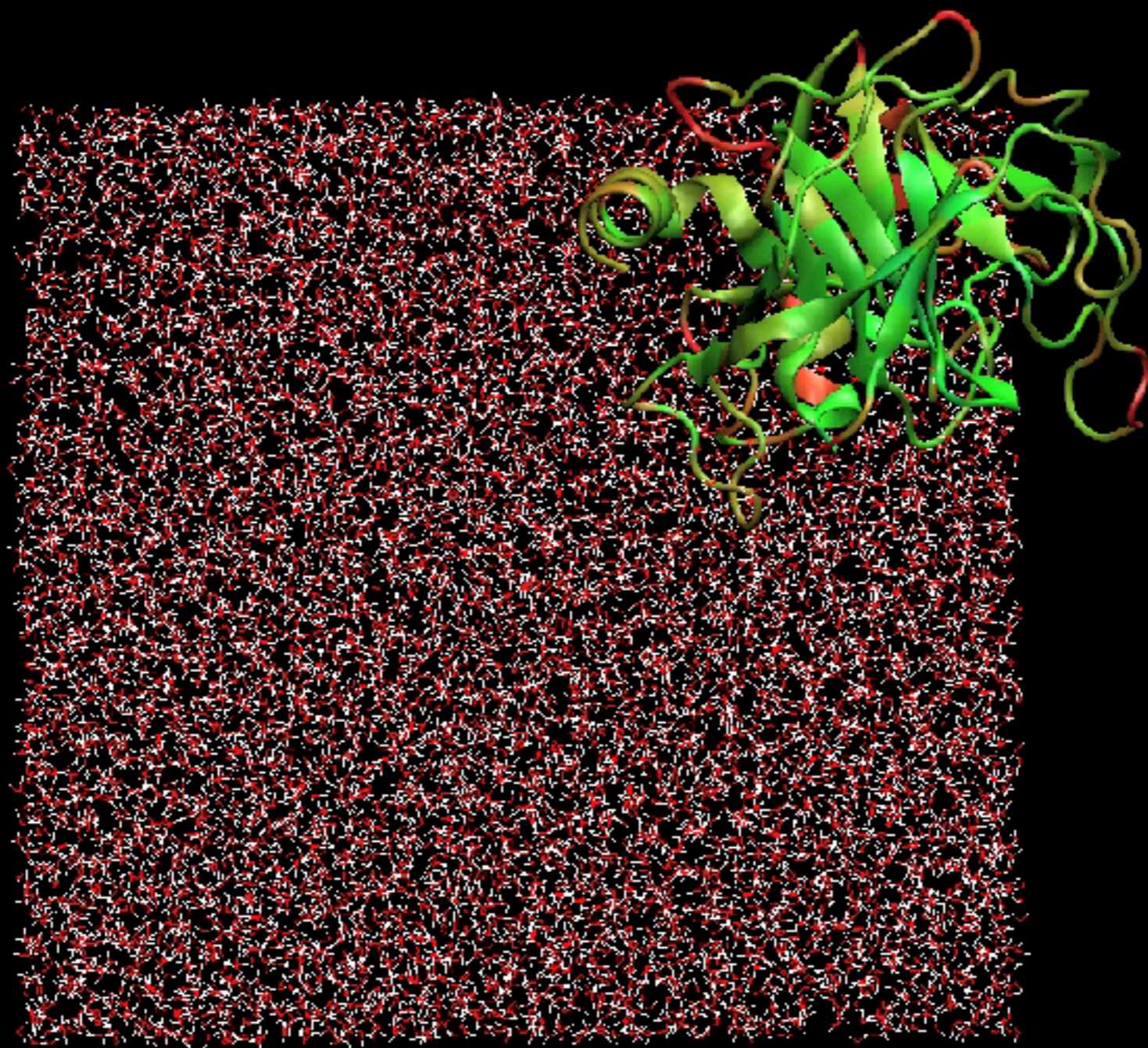
MD integrators / numerical precision / etc. do not conserve linear or angular momentum: need to “center” entire system and “overlap” bio-molecule prior to calculating dynamical quantities.

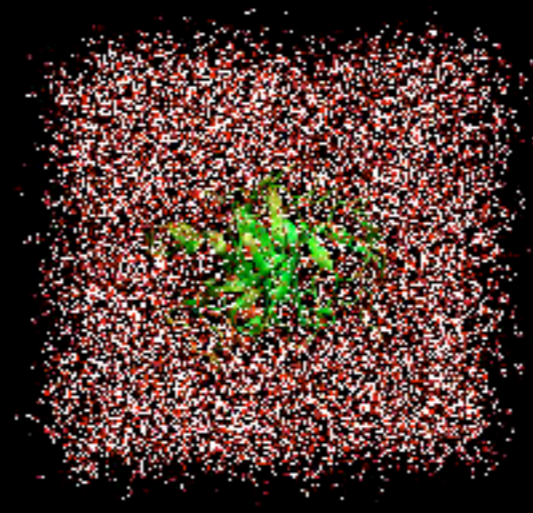
## Velocities -->

One should save these if you are interested in calculating time-autocorrelation functions: “classical spectra” = FT ( $\langle v(t) \cdot v(0) \rangle$ )

# Points to Remember

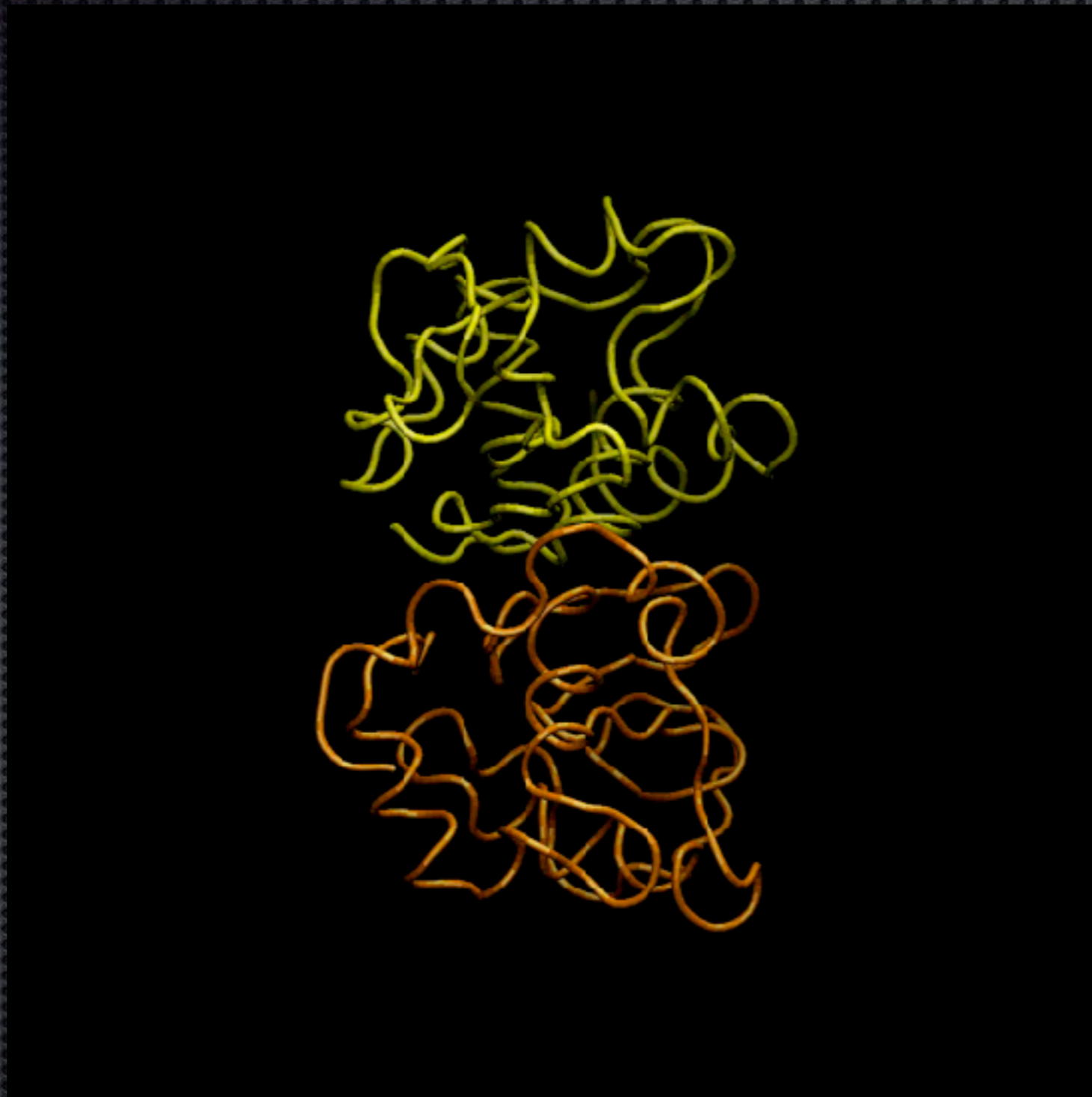
**Wrap or not to wrap:**





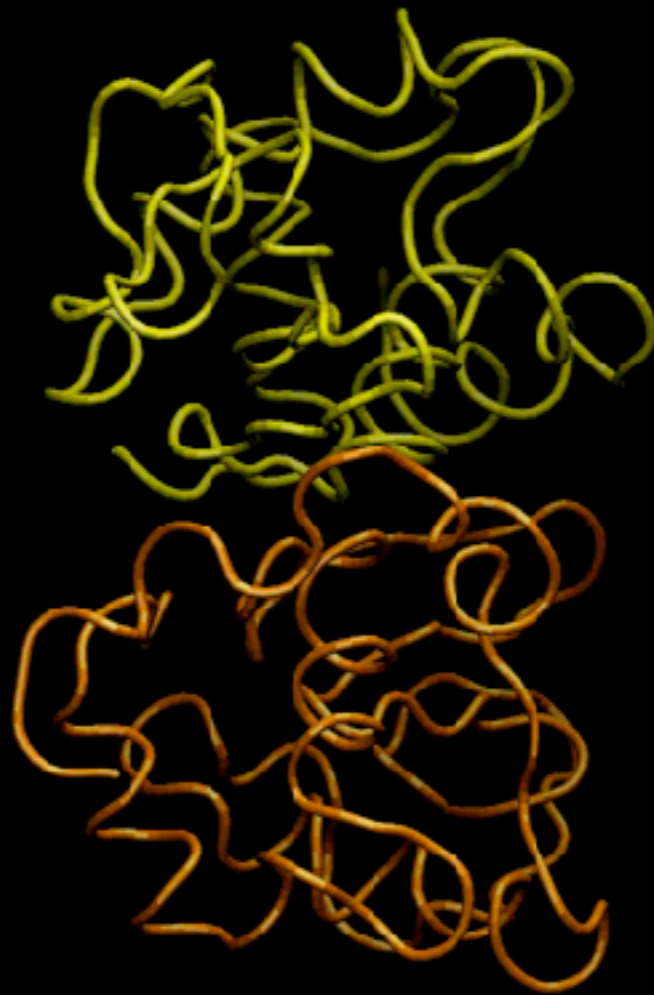
# Points to Remember

**Energy loss:** Lysozyme (2) 6 ns trajectory : dt for movie = 100 ps (1 fs in sim.)



# Points to Remember

**Energy loss:** Lysozyme (2) 6 ns trajectory : dt for movie = 100 ps (1 fs in sim.)



Kabsch W (1976). "A solution for the best rotation to relate two sets of vectors". *Acta Crystallographica* **32** (5): 922–923

# So many options . . .

Bio-simulation (All-atom MD packages):

Amber\*  
CHARMM\*  
NAMD\*  
GROMACS\*  
GROMOS  
LAMPPS\*  
PINYMD\*  
HIPPO  
GPIUTMD  
DL\_POLY\*  
ESPReso  
MacroModel\*  
MACSIMUS  
MOLDY  
MOSCITO  
ProtoMol  
TINKER\*  
MDGrape  
Materials Studio (InsightII)\*

Classical force fields:

Amber\*  
CHARMM\*  
CVFF  
COSMOS-NMR  
GROMACS\*  
GROMOS  
OPLS\*  
ENZMIX  
ECP/2  
QCFF/PI  
CFF\*  
MMFF  
MM2, MM3, MM4\*  
XPOL  
SIBFA  
AMOEB  
VALBOND  
DRF90  
CG MD\*

Methods & details\*:

Thermostats  
Barostats  
Electrostatics (PBC)  
Polarizability  
Implicit Solvent  
Langevin Dynamics  
Replica Exchange  
Parallel Tempering  
Steered MD  
Free-Energy Calculations  
Umbrella Sampling  
Normal Mode Analysis  
VMD  
Pymol  
Chimera, O  
APBS  
Gaussian, Gamess, CPMD  
Beowulf, GPU Clusters  
Rosetta  
PHYRE  
Folding @ Home  
TIP3, TIP3P, TIP4P, SPC, ST2

“HOW TO” . . .  
Start you on the path

# Further Reading

## NAMD

<http://www.ks.uiuc.edu/Training/Tutorials/>

## CHARMM & AMBER

<http://www.charmm.org/>

<http://ambermd.org/>

## MD

Frenkel & Smit : Understanding Molecular Simulation, ISBN: 0122673514

Allen & Tildesley : Computer Simulation of Liquids, ISBN: 9780198556459

Leach : Molecular Modelling: Principles and Applications, ISBN: 0582382106

Schlick : Molecular Modeling and Simulation, ISBN: 038795404X

## Key Papers

Electrostatics calculations: latest methodological advances. P. Koehl. *Current Opinion in Structural Biology*, 16:142-151, 2006.

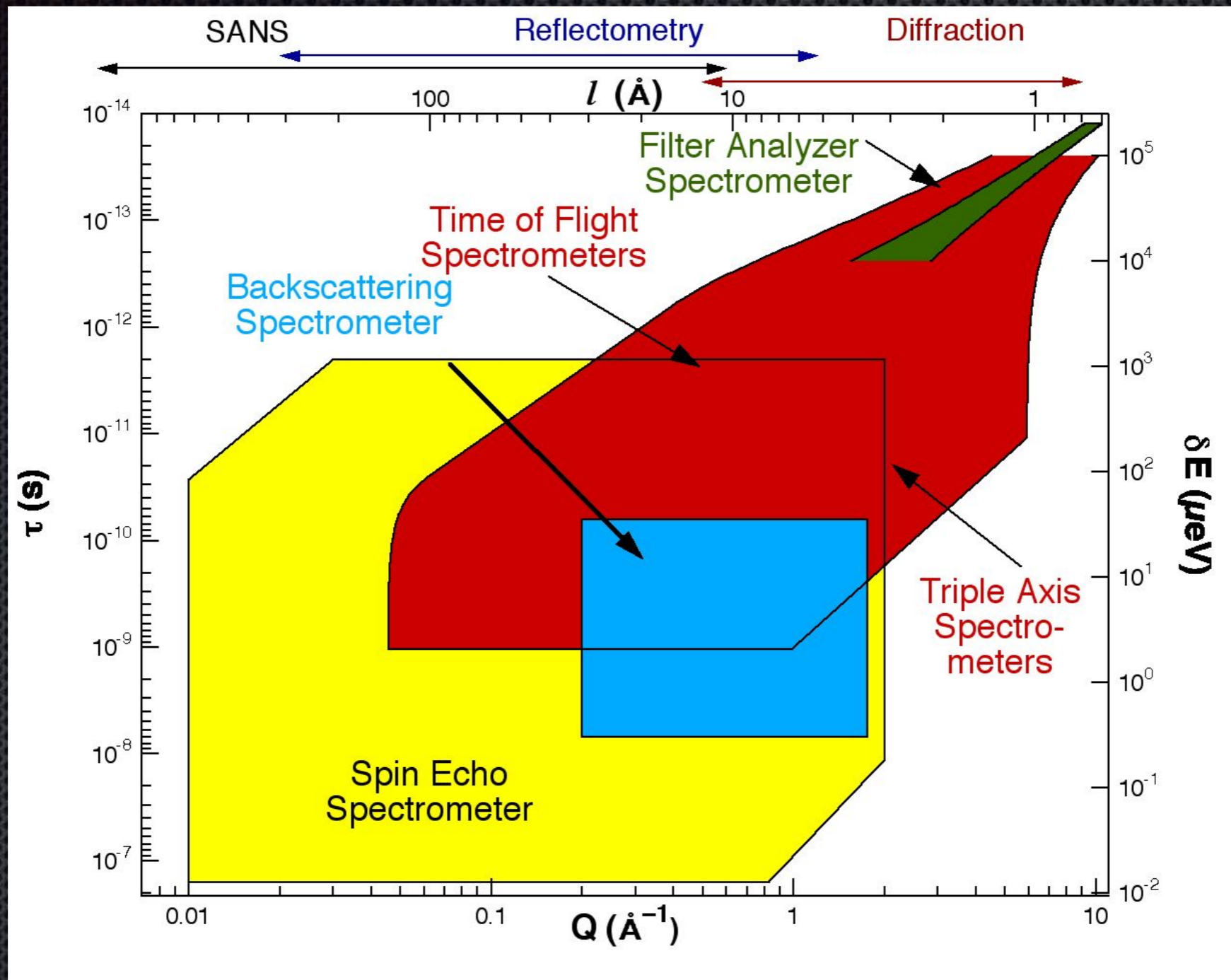
Empirical force fields for biological macromolecules: Overview and issues. A. D. MacKerell. *Journal of Computational Chemistry*, 25:1584-1604, 2004.

Molecular dynamics simulation of nucleic acids: successes, limitations, and promise. T. E. Cheatham and M. A. Young. *Biopolymers*, 56:232-256, 2000.

Molecular dynamics simulations of lipid bilayers. S. E. Feller. *Current Opinion in Colloid & Interface Science*, 5:217-223, 2000.

The MARTINI forcefield: coarse grained model for biomolecular simulations. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *Journal of Physical Chemistry B*, 111:7812-7824, 2007.

# NCNR Dynamics Instruments



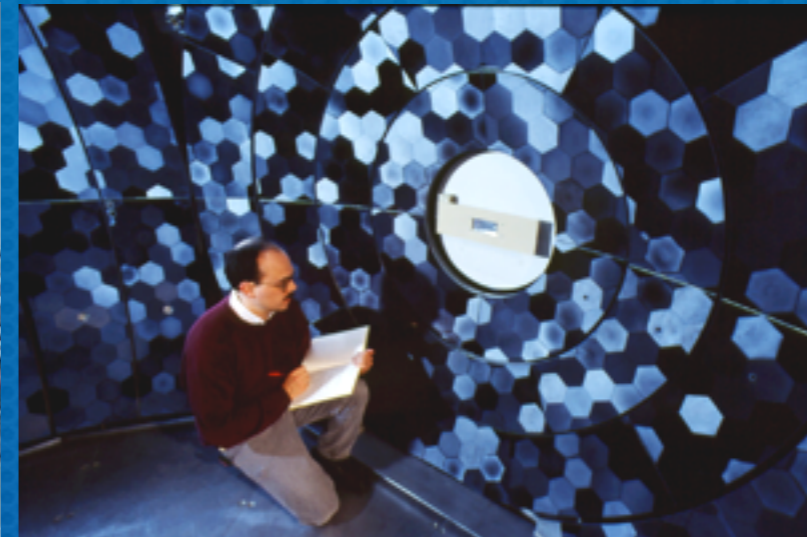
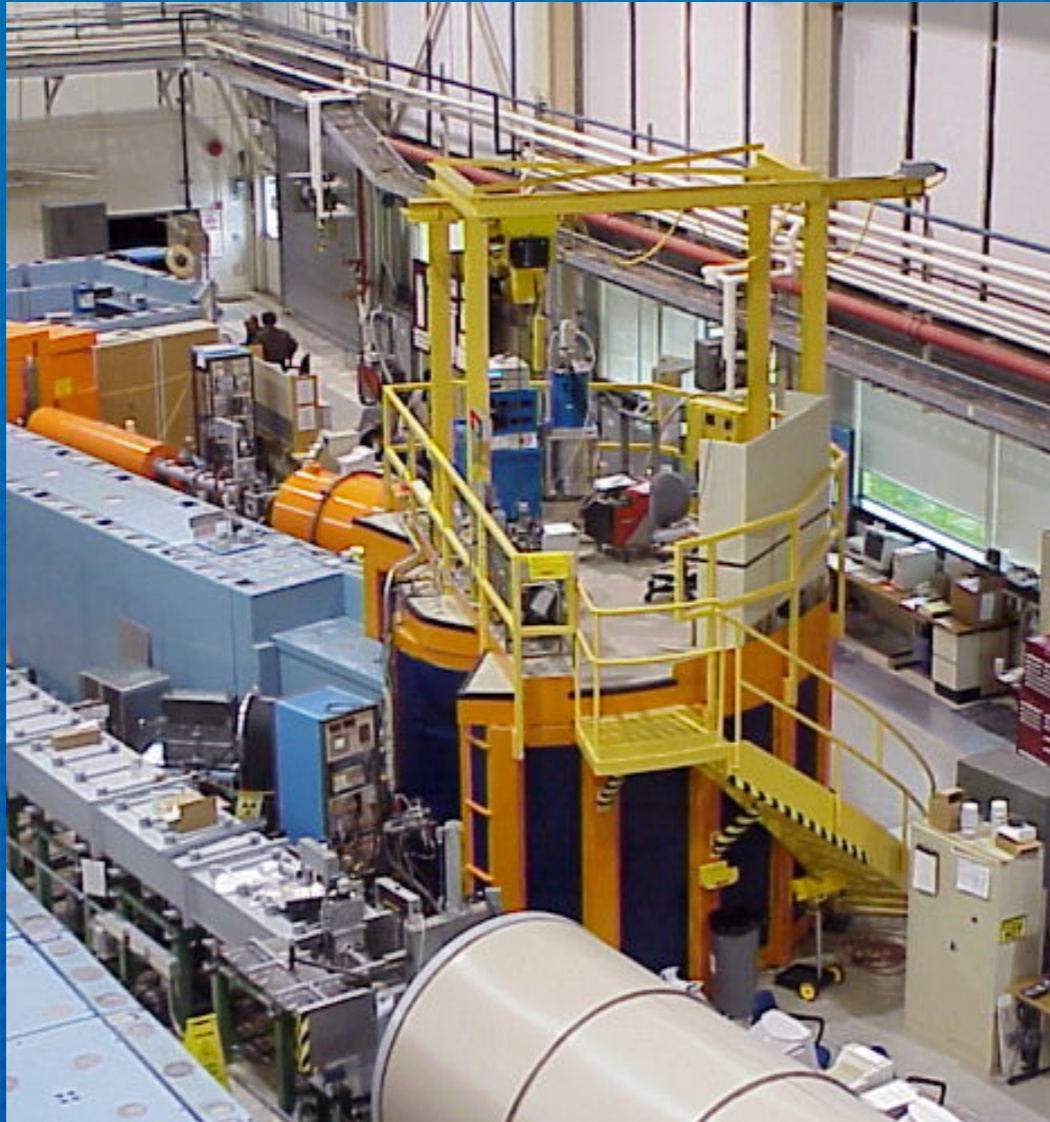
# NS Instruments (dynamics)

## Disk Chopper Spectrometer at the NCNR



The DCS is an extremely versatile instrument. Useful incident wavelengths range from  $< 2\text{\AA}$  to at least  $9\text{\AA}$ ; correspondingly the elastic energy resolution (FWHM) varies from  $\sim 1500$  to  $\sim 15 \mu\text{eV}$ : **picosecond dynamics**

# NS Instruments (dynamics) II



$\sim \text{ns} : \text{MSD} (S(q,0))$

The instrumental energy resolution is a Gaussian-like lineshape, with a full-width-half-maximum (FWHM) about  $1 \mu\text{eV}$ : **picosecond to  $\sim 5$  nanoseconds**

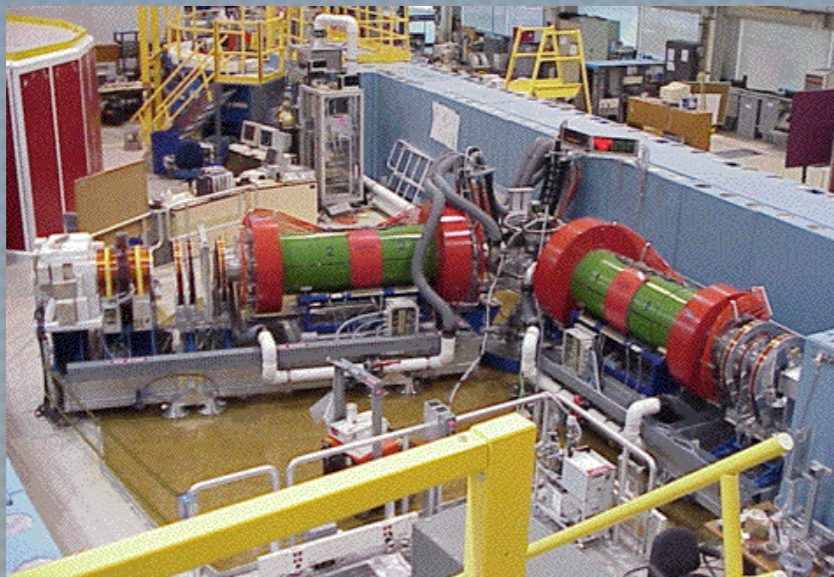
# NS Instruments (dynamics) III

Directly measure the energy transfer => Spin Echo

Advantage - high “energy” resolution using a broad energy band

**Actually measures  $I(Q,t)/I(Q,0)$**

Easier to measure Coherent Scattering, Particularly good for small Q



Cold neutron instruments

=> “Energy” Resolution 0.00001 to 0.001 meV

Examples – IN11, IN15, MESS, G1bis, SPAN, FRJ2, NSE

**Dynamic range:**

from 5 ps to 15 ns at  $\lambda = 6$

from 10 ps to 40 ns at  $\lambda = 8$

from 35 ps to 100 ns at  $\lambda = 12$

# Example “Bio” Publications

## DCS:

J.M. Borreguero, J. He, F. Meilleur, K. Weiss, C.M. Brown, D.A. Myles, K.W. Herwig, and P.K. Agarwal, "Redox-Promoting Protein Motions in Rubredoxin", J. Phys. Chem. B, Article ASAP.

D. Russo, J. Teixeira, L. Kneller, J.R.D. Copley, J. Ollivier, S. Perticaroli, E. Pellegrini, and M.A. Gonzalez, "Vibrational density of states of hydration water at biomolecular sites: hydrophobicity promotes low density amorphous ice behaviour", J. Am. Chem. Soc. **133**, 4882 (2011).

M.E. Johnson, C. Malardier-Jugroot, and T. Head-Gordon, "Effects of co-solvents on peptide hydration water structure and dynamics", Phys. Chem. Chem. Phys. **12**, 393 (2010).

C. Malardier-Jugroot, D.T. Bowron, A.K. Soper, M.E. Johnson and T. Head-Gordon, "Structure and water dynamics of aqueous peptide solutions in the presence of co-solvents", Phys. Chem. Chem. Phys. **12**, 382 (2010).

## HFBS:

B. Wang, M. Cicerone, Y. Aso, and M.J. Pikal, "The impact of thermal treatment on the stability of freeze-dried amorphous pharmaceuticals: II. aggregation in an IgG1 fusion protein", Journal of Pharmaceutical Sciences, **99**, 683, (2010).

X.-q. Chu, M. Lagi, E. Mamontov, E. Fratini, P. Baglioni, and S.-H. Chen, "Experimental evidence of logarithmic relaxation in single-particle dynamics of hydrated protein molecules", Soft Matter, **6**, 2623–2627, (2010).

S. Khodadadi, J. H. Roh, A. Kisliuk, E. Mamontov, M. Tyagi, S. A. Woodson, R. M. Briber, A. P. Sokolov, "Dynamics of Biological Macromolecules: Not a Simple "Slaving" by Hydration Water", Biophys. J., **98**, 1321–1326, (2010).

M.E. Johnson, C. Malardier-Jugroot, R.K. Murarka, and T. Head-Gordon, "Hydration Water Dynamics near Biological Interfaces", J. Phys. Chem. B, **113**, 4082, (2009).

## NSE:

J.-H. Lee, S.-M. Choi, C. Doe, A. Faraone, P.A. Pincus, and S.R. Kline, "Thermal Fluctuation and Elasticity of Lipid Vesicles Interacting with Pore-Forming Peptides" Phys. Rev. Lett., **105**, 038101 (2010).

M.B. Boggara, A. Faraone, and R. Krishnamoorti, "Effect of pH and Ibuprofen on the Phospholipid Bilayer Bending Modulus" J. Phys. Chem. B, **114**, 8061 (2010).

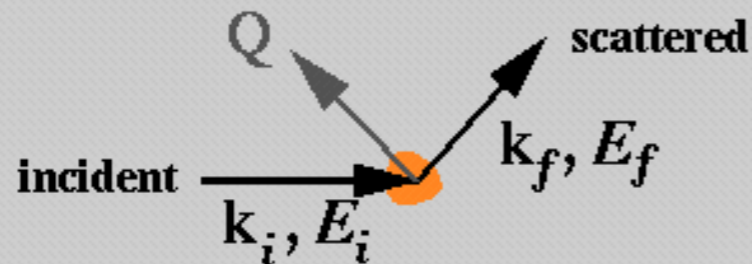
L. Porcar, P. Falus, W.-R. Chen, A. Faraone, E. Fratini, K. Hong, P. Baglioni, and Y. Liu, "Formation of the Dynamic Clusters in Concentrated Lysozyme Protein Solutions" J. Phys. Chem. Lett., **1**, 126 (2010).

# Observables from MD

Experiment	Corresponding time correlation function
Dynamical quantity	
Neutron scattering $r$ position of atom	$\left\langle \sum e^{-iQ \cdot r_m(0)} e^{iQ \cdot r_n(t)} \right\rangle$
Raman scattering $u$ transition dipole	$\langle P_2[u(0) \cdot u(t)] \rangle$
Infrared absorption $u$ transition dipole	$\langle u(0) \cdot u(t) \rangle$
Dielectric relaxation $u$ permanent dipole	$\sum \langle u_i(0) \cdot u_j(t) \rangle$
NMR lineshape $M_x$ magnetization	$\langle M_x(0) \cdot M_x(t) \rangle$
Self-diffusion $v$ velocity	$\langle v(0) \cdot v(t) \rangle$

# Neutron Scattering

Neutrons scatter from the nuclei in a sample:



$$Q = k_f - k_i = \text{momentum transfer}$$

$$\Delta E = E_f - E_i = \hbar\omega = \text{energy transfer}$$

Measure double differential cross-section:

$$\frac{\partial^2 \sigma}{\partial \Omega \partial \omega} = \text{probability that neutron scatters within solid angle } d\Omega \text{ of direction } \Omega \text{ and } \omega \text{ within } d\omega \text{ of } E_i / \hbar$$
$$\propto \frac{k_f}{k_i} [\sigma^{\text{coh}} S^{\text{coh}}(Q, \omega) + \sigma^{\text{inc}} S^{\text{inc}}(Q, \omega)]$$

where  $S(Q, \omega)$ , the dynamical structure factors, contain information on the structure and dynamics of the sample

coherent  $\Rightarrow$  structure (diffraction), collective motions

incoherent  $\Rightarrow$  single-particle motions (e.g. vibrations, diffusion)

Each nucleus has unique cross-sections,  $\sigma^{\text{coh}}$  and  $\sigma^{\text{inc}}$

# Neutron Scattering II

The composition of the system determines whether the scattering is primarily coherent, incoherent, or both

Because  $\sigma_{\text{H}}^{\text{inc}} \gg$  all other  $\sigma$ , in organic molecules the scattering is primarily incoherent

$\Rightarrow$  probe single particle motions of H atoms

Connection to the time-domain through the Van Hove Theory for the structure factor (1954):

$$S(Q, \omega) = (1/2\pi) \int_{-\infty}^{\infty} I(Q, t) e^{-i\omega t} dt$$

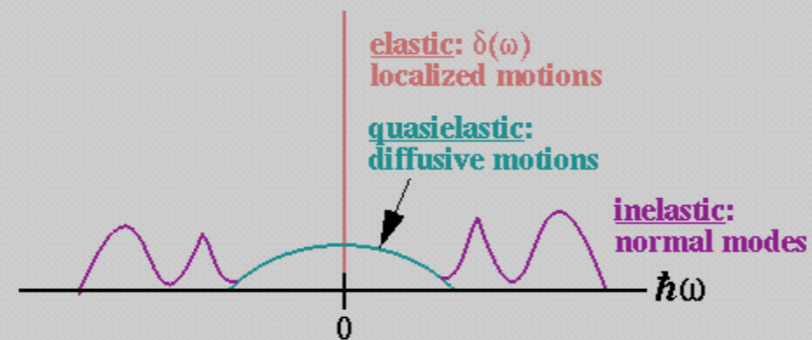
$$I(Q, t) = (1/N) \sum_j \langle e^{iQ \cdot r_j(t)} e^{-iQ \cdot r_j(0)} \rangle$$

is the incoherent intermediate scattering function which we can calculate from simulation and FT to obtain the structure factor

Typical instrumental and MD time and length scales overlap very well

# Neutron Scattering III

Typical spectrum,  $S(Q,\omega)$ , has three parts:

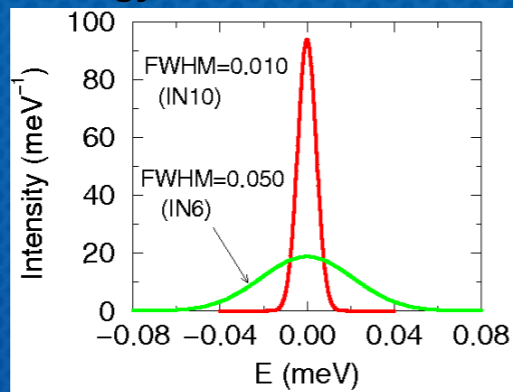


In practice, spectrum is broadened by finite energy resolution, which sets a limit on time scale probed by the experiment

$$S_{measured}(Q,\omega) = S(Q,\omega) \otimes R(\omega)$$

$R(\omega)$  determines time scale probed by measurement:

Energy resolution functions

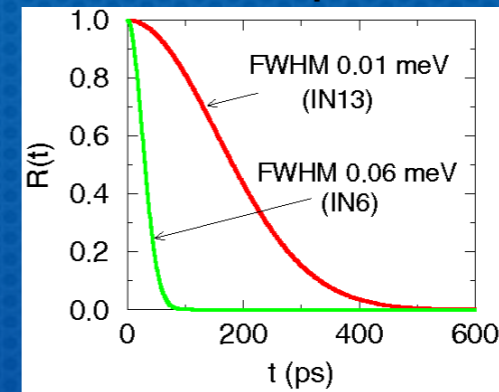


Fourier transform



Convolution with energy resolution function = multiplication by its FT in time domain

Time domain equivalents

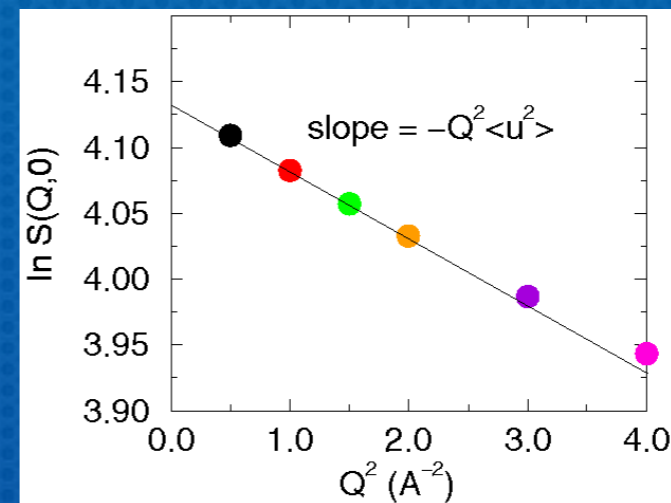
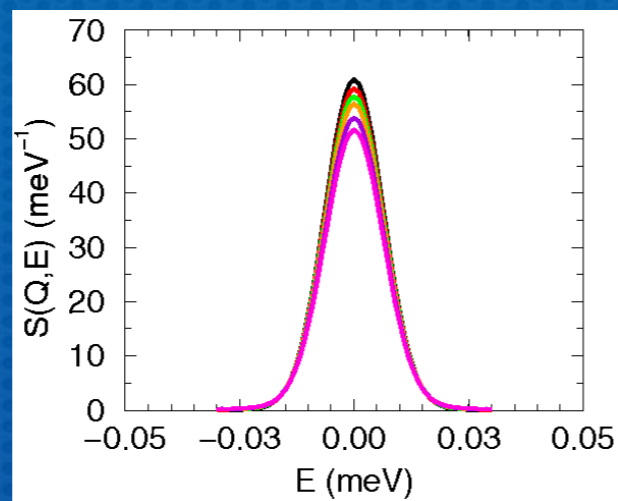


# Neutron Scattering IV

Most neutron scattering studies of protein dynamics to date have been based on measuring the temperature dependence of mean-squared displacements of H atoms ( $\langle u^2 \rangle$ ) via elastic scattering

Estimate mean-squared displacement from elastic intensity via Debye-Waller factor:  $S(Q, E=0) = \exp(-Q^2 \langle u^2 \rangle)$

In practice, the number of counts within the instrumental resolution ( $\Delta E$ ) of the elastic line is used, i.e.  $S(Q, 0) \sim S(Q, \pm \Delta E)$

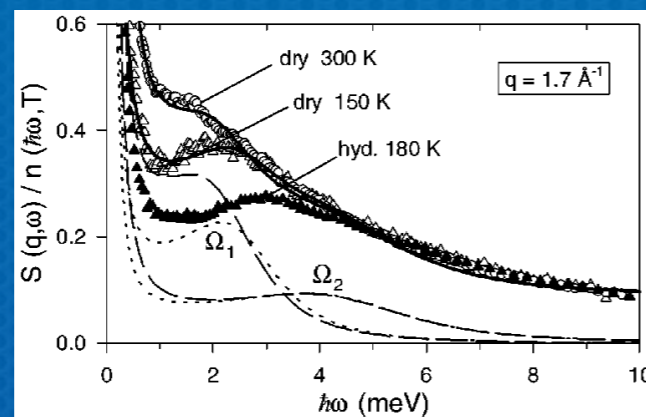


$$\langle u^2 \rangle = -\text{slope} / Q^2$$

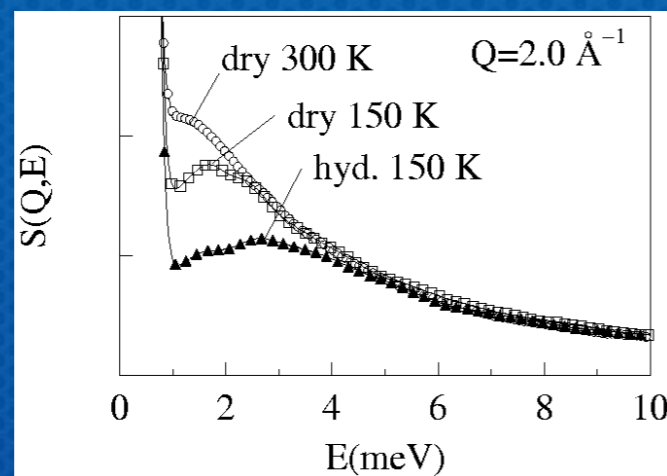
# Neutron Scattering V

## “Boson peak” in low frequency inelastic spectrum

Incoherent neutron scattering on Mb powders  
[Leyser et al., Phys. Rev. Lett. (1999)]

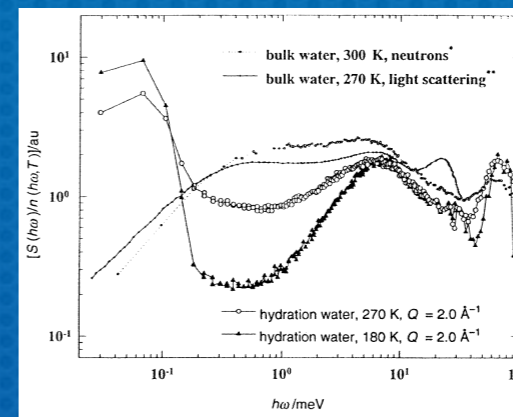


MD simulations of RNase powders  
[Tarek & Tobias, J. Chem. Phys. (2000)]

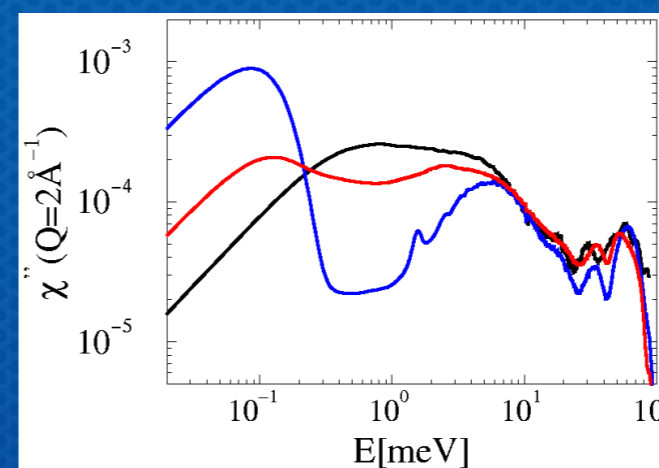


## Dynamical susceptibility

Dynamical structure factors:  
bulk vs. protein hydration water  
[Settle & Doster, Faraday Disc. 103, 269 (1996)]



MD simulations of RNase powders  
(protein signal subtracted)  
[Tarek & Tobias, Biophys. J. 79, 3244 (2000)]



# Steps to Calculate $I(q,t)$

$$I(Q,t) = \frac{1}{N} \sum_j \left\langle e^{-iQ \cdot r_j(0)} e^{iQ \cdot r_j(t)} \right\rangle$$

Correct of COM & rotation of system

Create subset trajectory with only non-exchangeable hydrogens

Need a method to sample over “Q”

Write code (Python & C) to extract positions and calculate  $I(q,t)$  using MTO

# Steps to Calculate $I(q,t)$ II

$$I(Q,t) = \frac{1}{N} \sum_j \left\langle e^{-iQ \cdot r_j(0)} e^{iQ \cdot r_j(t)} \right\rangle$$



```
def random_vector_sphere():  
    """  
    von Neumann method  
    """  
    looking=True  
    while(looking):  
        r1=random.random()  
        r2=random.random()  
        r3=random.random()  
        z1=1-2*r1 ; z2=1-2*r2 ; z3=1-2*r3  
        sumz2=z1*z1+z2*z2+z3*z3  
        if(sumz2 < 1):  
            z=math.sqrt(sumz2)  
            vector=[z1/z,z2/z,z3/z]  
            looking=False  
    return vector
```

**Need a method to sample over “Q”**

**Code to generate a random vector on a unit sphere**

**Call this snippet to create an array of q-vectors to loop over (8 - 32 for convergence)**

**The magnitude of q,  $q = |q|$ \*unit vector**

# Steps to Calculate $I(q,t)$ II

$$I(Q,t) = \frac{1}{N} \sum_j \left\langle e^{-iQ \cdot r_j(0)} e^{iQ \cdot r_j(t)} \right\rangle$$



```
argo=tqx*xo+tqy*yo+tqz*zo ;  
argt=tqx*xt+tqy*yt+tqz*zt ;  
re=cos(argo)*cos(argt)+sin(argo)*sin(argt) ;  
im=sin(argo)*cos(argt)-cos(argo)*sin(argt) ;
```

Create array of q vectors

Read in coordinates

Loop over atoms

Loop over q-vectors

Calculate  $q \cdot r$

$I(q,t) = I(q,t) + \cos(\text{old}) * \cos(\text{new}) \dots$

Normalize

Print results

**Poor convergence with a finite trajectory.** Need a method to get more out of the data ---> Multiple Time Origins (MTO)

# MTO

A correlation function of some function,  $A(t)$ , can be written as:

$$C(\tau) = \langle A(\tau)A(0) \rangle = \frac{1}{\tau_{max}} \sum_{\tau_0=1}^{\tau_{max}} A(\tau_0)A(\tau_0 + \tau)$$

Thus, we average over  $t_{max}$  time origins the product of  $A(t_0)$  and  $A(t_0+dt)$  later.

Thus, each successive data point is used as an origin ( $t_0$ ).

In practice, use every 5 or 10th point to avoid correlations.

Must not exceed data length, so each “window” uses a different number of points. --> need to normalize correctly.

```
def corfun(trun, tcor, a, acf, norm):
```

```
    for t in xrange(tcor):
```

```
        acf(t) = 0.0
```

```
        norm(t) = 0.0
```

```
    for t0 in range(1, trun):
```

```
        a0 = a(t0)
```

```
        tt0max = min(trun, t0 + tcor)
```

```
        for tt0 in xrange(t0, tt0max):
```

```
            t = tt0 - t0
```

```
            acf(t) = acf(t) + a0 * a(tt0)
```

```
            norm(t) = norm(t) + 1.0
```

```
    for t in xrange(tcor):
```

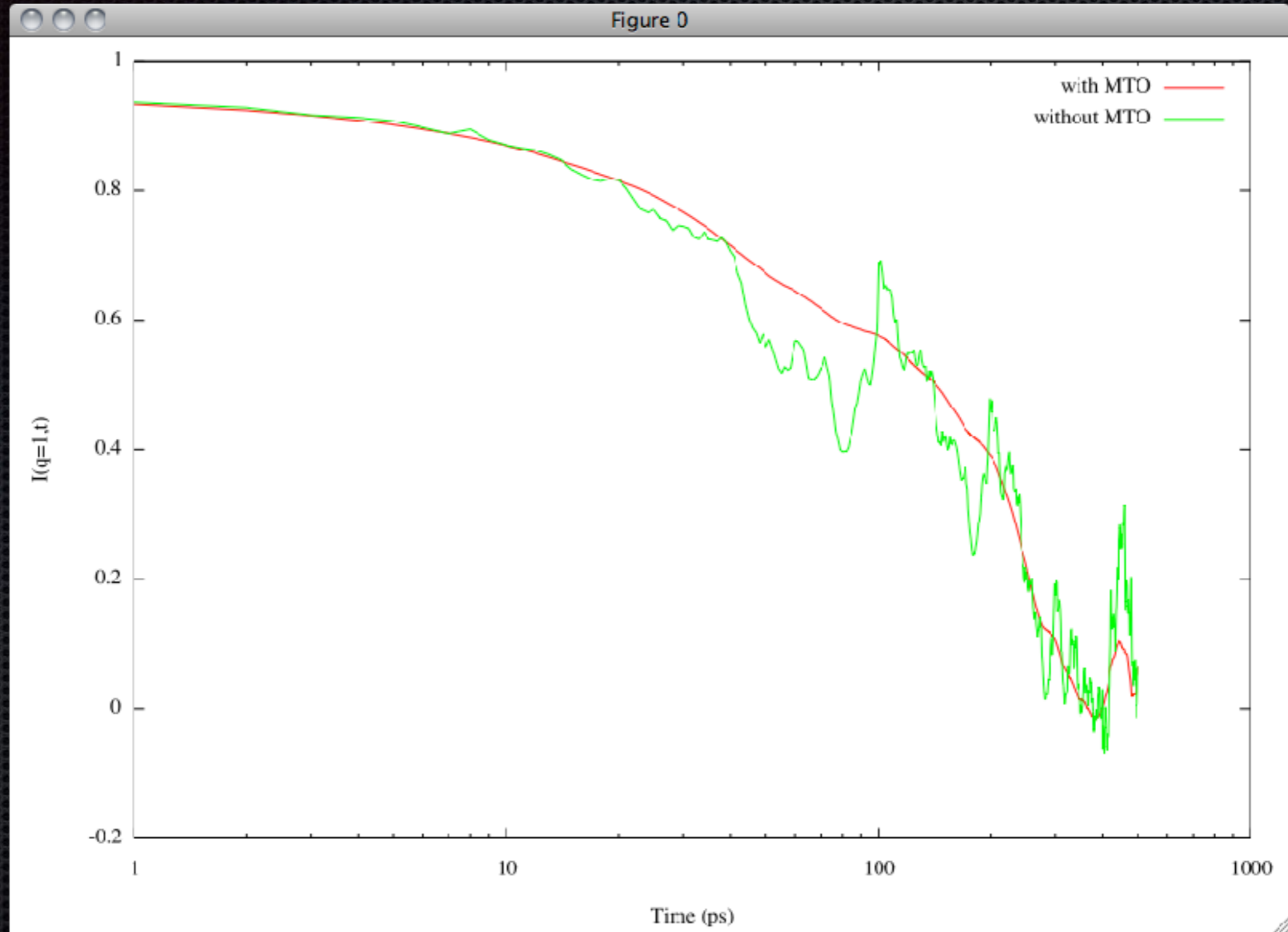
```
        acf(t) = acf(t) / norm(t)
```

```
    return acf
```

See: Allen & Tildesley “Computer Simulation of Liquids”, Oxford Univ. Press, NY, 1997, page 185-188.

PLUS FFT methods ...

# MTO Example



# Steps to Calculate $S(q,w)$ , $X''(q,w)$ & $\langle u^2 \rangle$

$$S(Q,\omega) = (1/2\pi) \int_{-\infty}^{\infty} I(Q,t) e^{-i\omega t} dt$$

$$S_{measured}(Q,\omega) = S(Q,\omega) \otimes R(\omega)$$

**GOAL:** to compare MD  $S(q,w)$  to measured  $S(q,w)$

Multiply  $I(q,t)$  by  $R(t)$  to incorporate instrumental resolution function

Take the discrete Fourier Transform of  $I(q,t) \times R(t) = S(q,w)$

**BONUS:**  $X''(q,w) \sim w^* S(q,w)$

Estimate  $\langle u^2 \rangle$  by plotting  $\ln(S(q,0))$  -vs-  $q^2$

$$\langle u^2 \rangle = -\text{slope}/q^2$$

$\langle u^2 \rangle$  is usually studied as a function of temperature  $\langle u^2(T) \rangle$

DWF

$$S(q,w=0) = \exp(-q^2 \langle u^2 \rangle)$$

# Examples

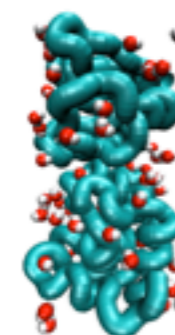
## Lysozyme Powder Under Pressure

INS of hydrated lysozyme ( $h=0.3$ ) at two pressures and two length scales

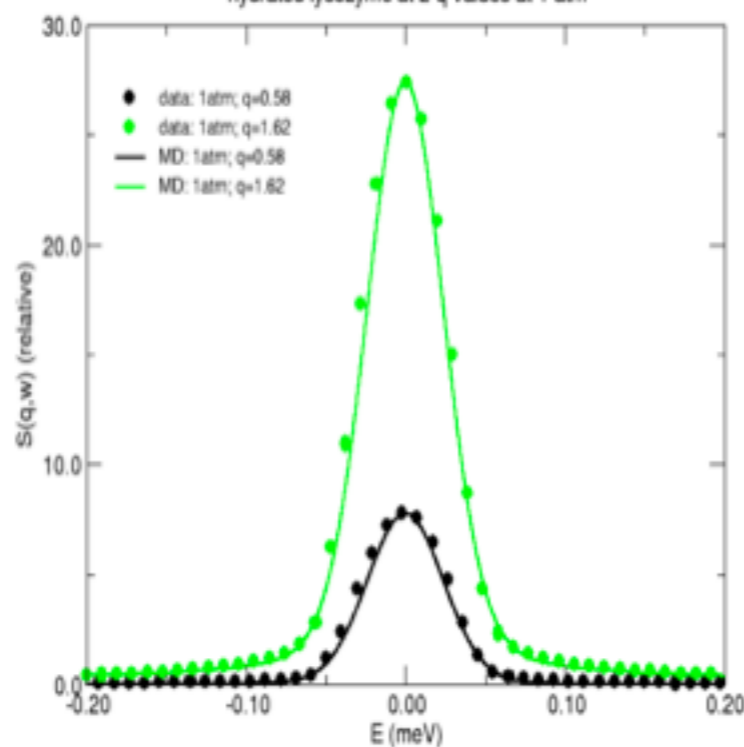
Agreement with simulation is robust and allows for further atomic scale analysis

### RMSD ( $\text{\AA}^2$ )

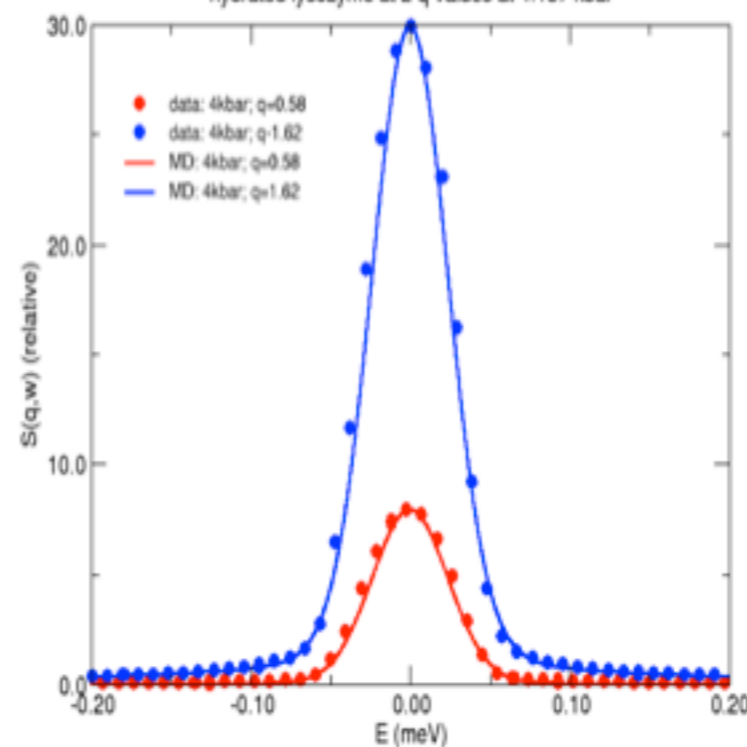
Ref.	1 atm	4kbar
xtal	2.01	2.22
Self	0.62	0.53
1 atm	-	2.36



$S(q,w)$  Comp Plot: Expt and MD sim  
hydrated lysozyme at 2 q-values at 1 atm

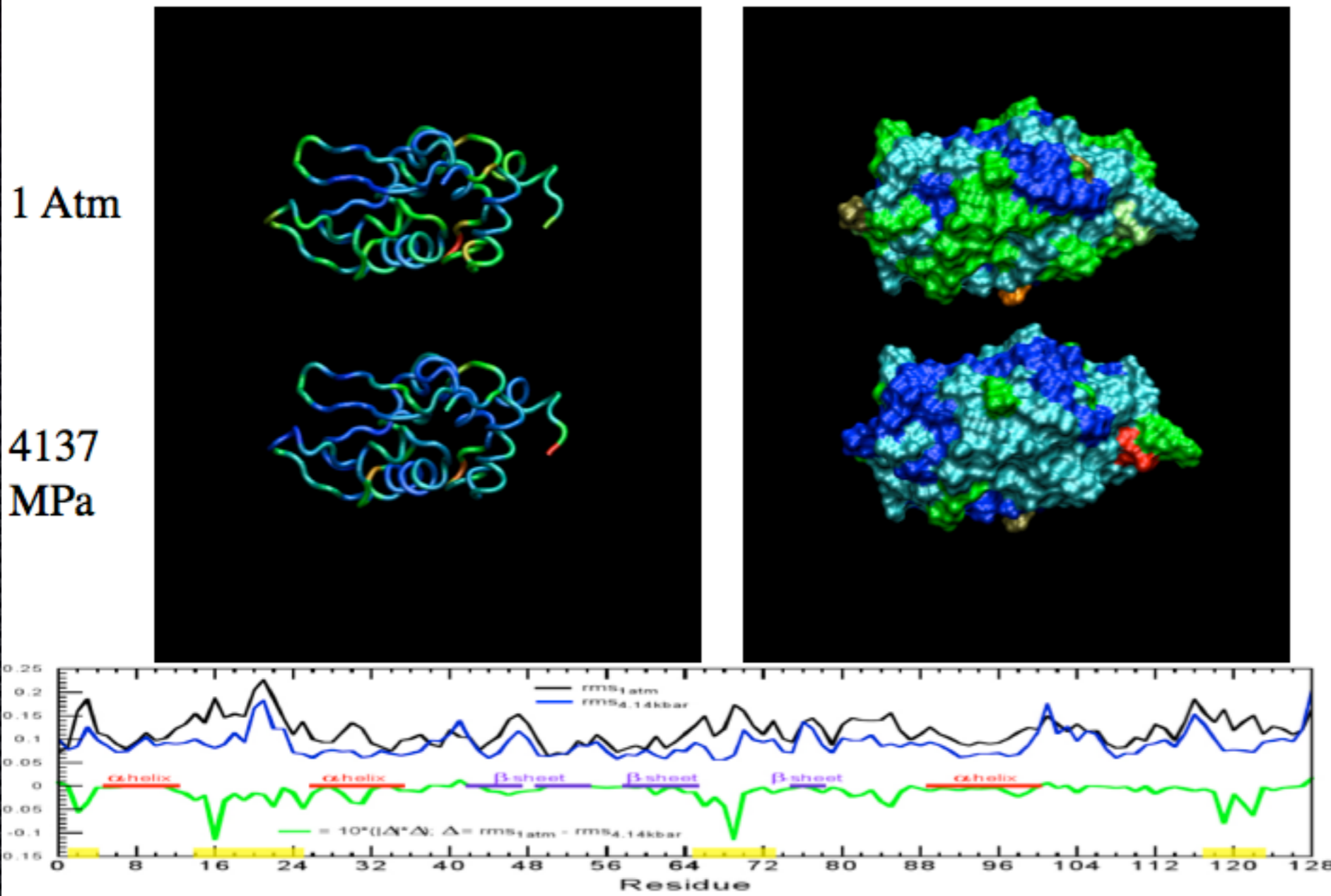


$S(q,w)$  Comp Plot: Expt and MD sim  
hydrated lysozyme at 2 q-values at 4.137 kbar

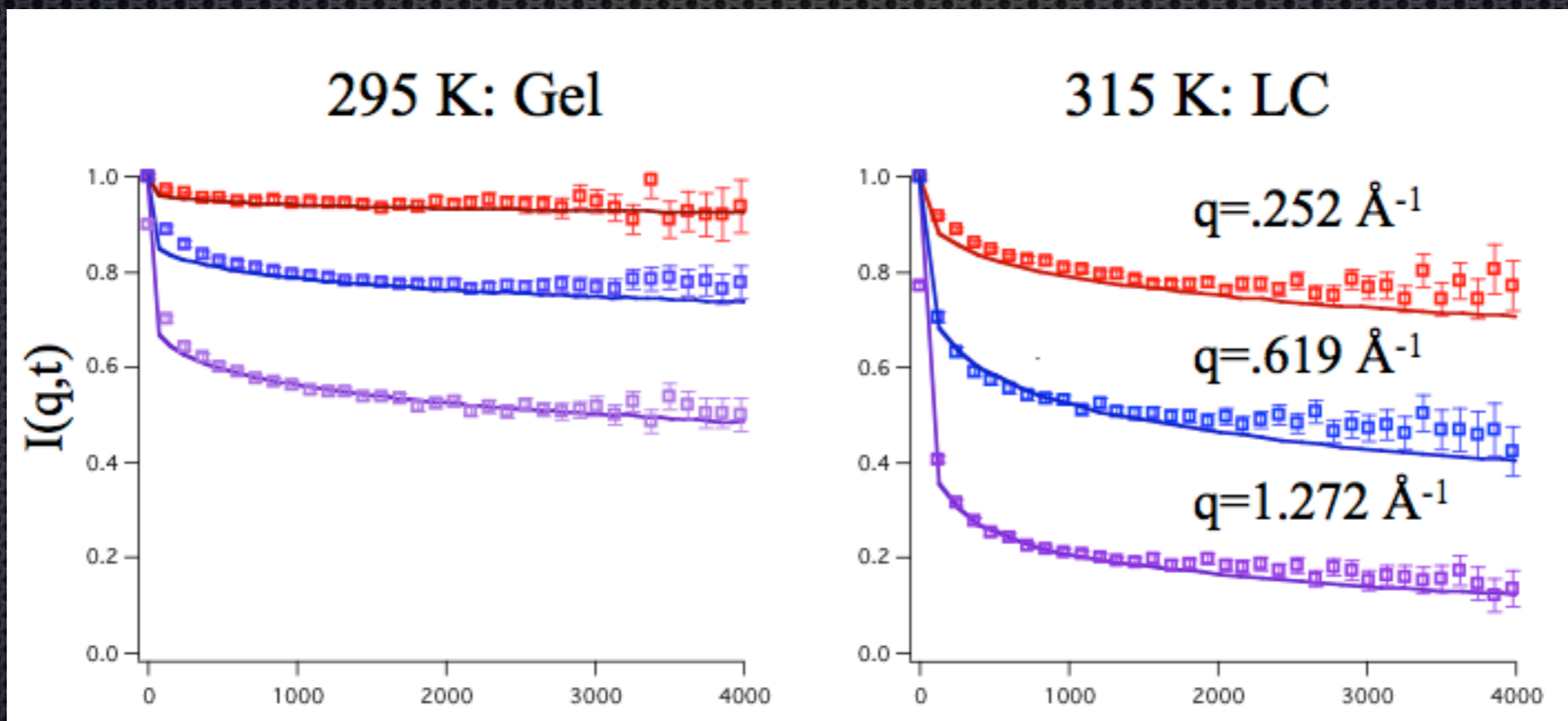
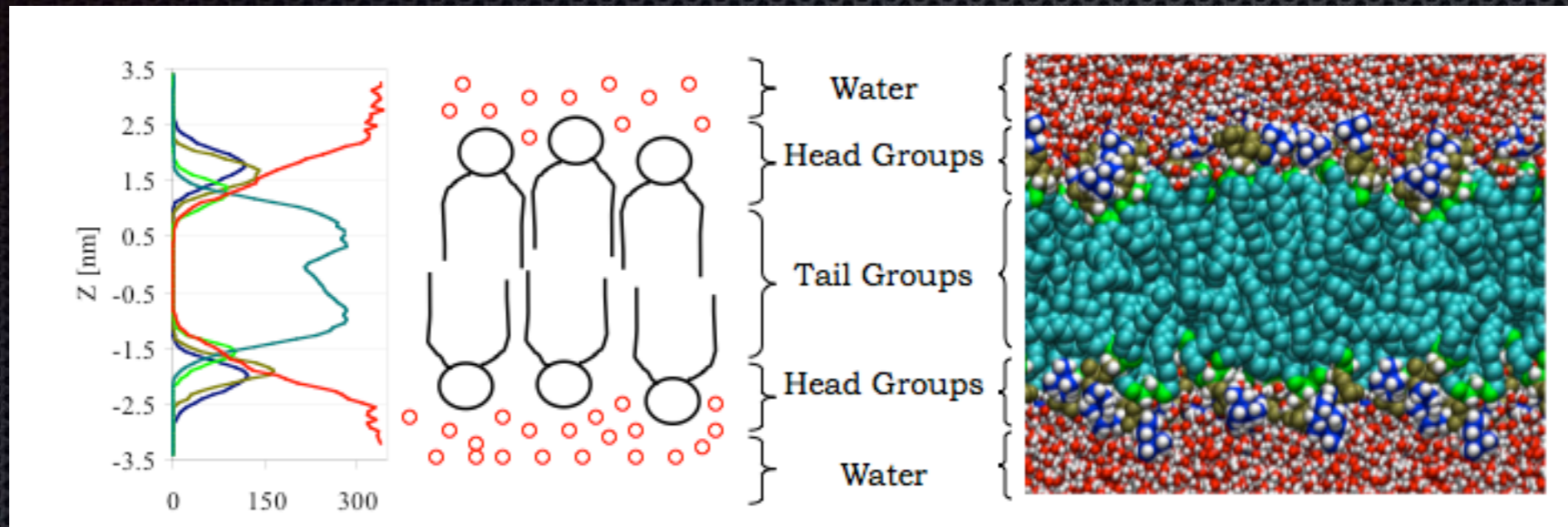


# Examples

## Atomic details ... pressure



# Examples

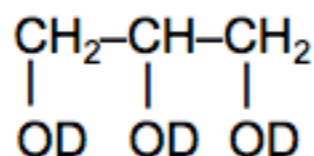
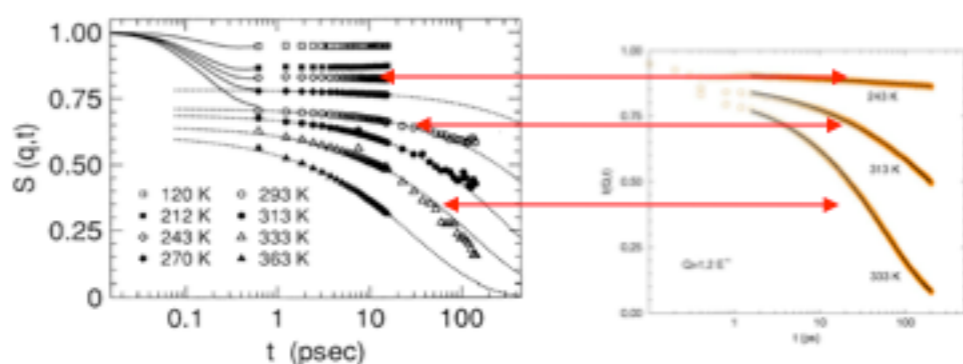


# Examples

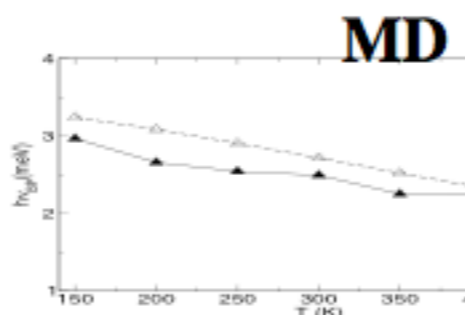
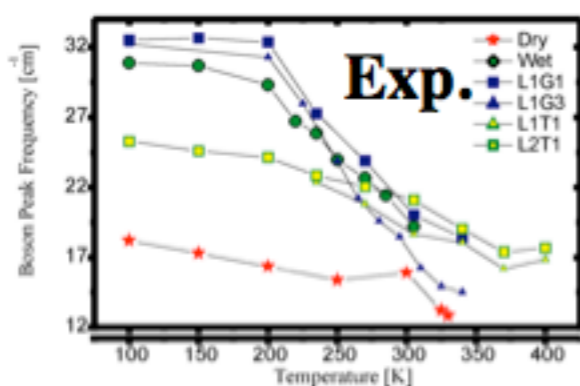
## Validation of FF and Pure Glass Systems

### Glycerol

Fourier inversion of neutron scattering spectra  
[Wuttke et al., PRE 52, 4026 (1995)]

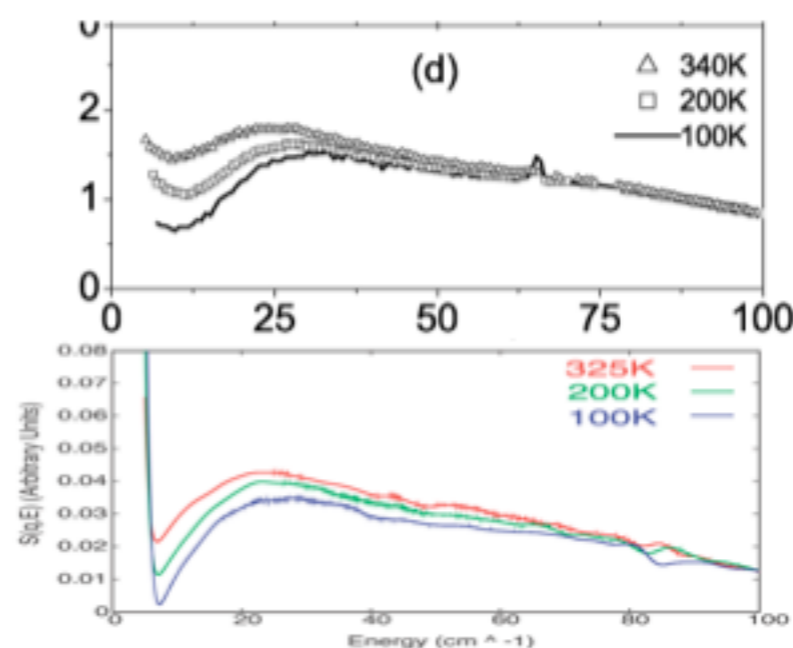


100s of ps



### Trehalose

Raman spectra of lysozyme in dehydrated trehalose  
[Caliskan et al., J. Non-Cryst. Solids (2002)]



ps

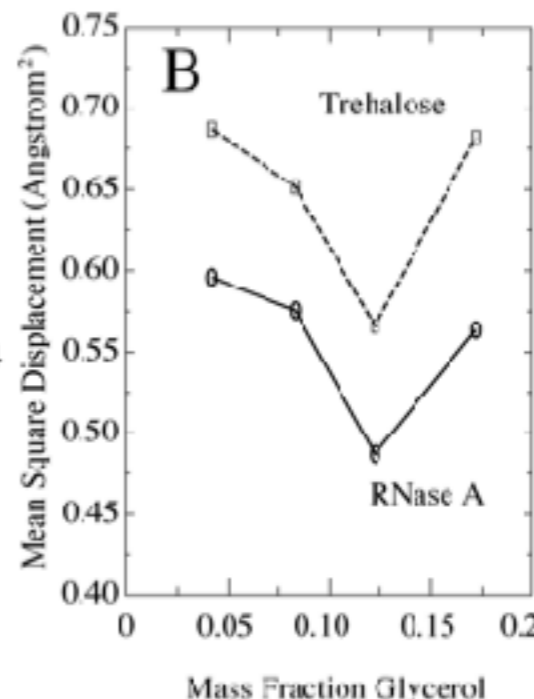
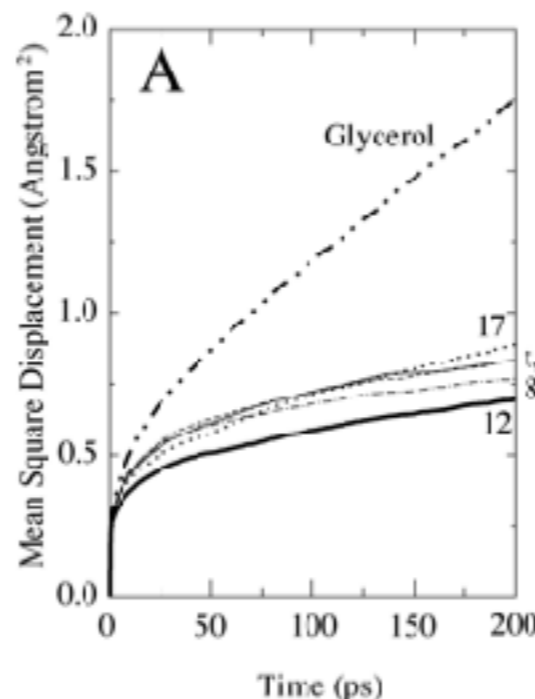
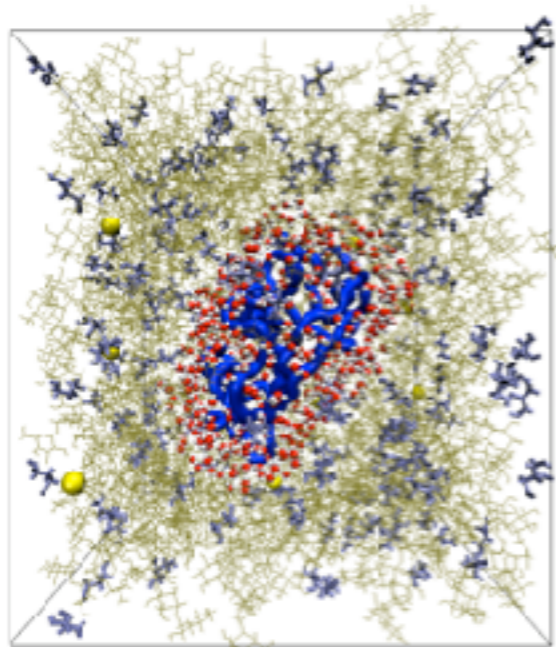
$$S(q, \nu) = \frac{A v_0}{v_0^2 + \nu^2} + B \exp \left\{ - \frac{[\ln(\nu / \nu_{BP})]^2}{2[\ln(W / \nu_{BP})]^2} \right\}$$

ICP.122, 244910 (2005); ICP.124, 1-8 (2006)

# Examples

Binary glasses and protein dynamics: glycerol in trehalose

Moderate Protein Density (7 mM)



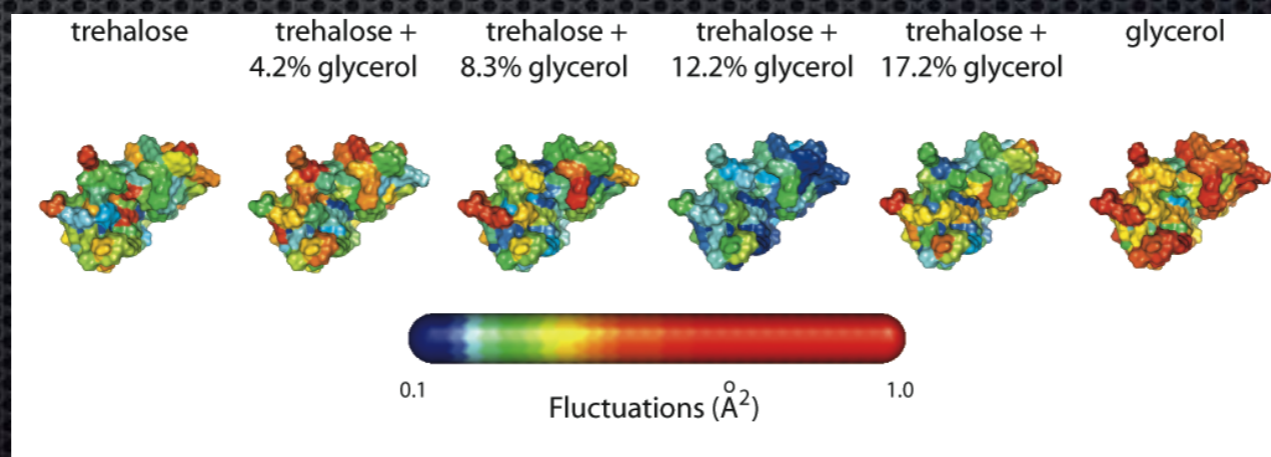
**RMSD (Ca) < 2Å**  
**MSD (H2O) --> 8Å**

**~30,000 atoms, 10 ns**

**6 protein/glass combinations**

**“What distinguishes one glass formulation from another?”**

Dehydrated binary glass is rigidified by glycerol-trehalose hydrogen bonds (JCP **122**, 114505 (2005), JCP **122**, 244910 (2005), JCP **124**, 1-8 (2006)).



# Further Reading

## NAMD

<http://www.ks.uiuc.edu/Training/Tutorials/>

## CHARMM & AMBER

<http://www.charmm.org/>

<http://ambermd.org/>

## MD

Frenkel & Smit : Understanding Molecular Simulation, ISBN: 0122673514

Allen & Tildesley : Computer Simulation of Liquids, ISBN: 9780198556459

Leach : Molecular Modelling: Principles and Applications, ISBN: 0582382106

Schlick : Molecular Modeling and Simulation, ISBN: 038795404X

## Key Papers

Electrostatics calculations: latest methodological advances. P. Koehl. *Current Opinion in Structural Biology*, 16:142-151, 2006.

Empirical force fields for biological macromolecules: Overview and issues. A. D. MacKerell. *Journal of Computational Chemistry*, 25:1584-1604, 2004.

Molecular dynamics simulation of nucleic acids: successes, limitations, and promise. T. E. Cheatham and M. A. Young. *Biopolymers*, 56:232-256, 2000.

Molecular dynamics simulations of lipid bilayers. S. E. Feller. *Current Opinion in Colloid & Interface Science*, 5:217-223, 2000.

The MARTINI forcefield: coarse grained model for biomolecular simulations. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *Journal of Physical Chemistry B*, 111:7812-7824, 2007.

# So many options . . .

Bio-simulation (All-atom MD packages):

Amber\*  
CHARMM\*  
NAMD\*  
GROMACS\*  
GROMOS  
LAMPPS\*  
PINYMD\*  
HIPPO  
GPIUTMD  
DL\_POLY\*  
ESPReso  
MacroModel\*  
MACSIMUS  
MOLDY  
MOSCITO  
ProtoMol  
TINKER\*  
MDGrape  
Materials Studio (InsightII)\*

Classical force fields:

Amber\*  
CHARMM\*  
CVFF  
COSMOS-NMR  
GROMACS\*  
GROMOS  
OPLS\*  
ENZMIX  
ECP/2  
QCFF/PI  
CFF\*  
MMFF  
MM2, MM3, MM4\*  
XPOL  
SIBFA  
AMOEB  
VALBOND  
DRF90  
CG MD\*

Methods & details\*:

Thermostats  
Barostats  
Electrostatics (PBC)  
Polarizability  
Implicit Solvent  
Langevin Dynamics  
Replica Exchange  
Parallel Tempering  
Steered MD  
Free-Energy Calculations  
Umbrella Sampling  
Normal Mode Analysis  
VMD  
Pymol  
Chimera, O  
APBS  
Gaussian, Gamess, CPMD  
Beowulf, GPU Clusters  
Rosetta  
PHYRE  
Folding @ Home  
TIP3, TIP3P, TIP4P, SPC, ST2

“HOW TO” . . .  
Start you on the path