Intermediate Building

CCP-SAS 2015

Atomistic Modeling for Small Angle Scattering:

A Short Course

May 26-28, 2015

Institut Laue-Langevin, France

EPSRC
Engineering and Physical Sciences Research Council

NSF

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?

EPSRC
Engineering and Physical Sciences Research Council

NSF

Overview

Intermediate Building Tricks

Building coordinates for missing atoms/residues
Simulated annealing
Building IDPs
A trick for topological issues

"model building"

EPSRC
Engineering and Physical Sciences Research Council

NSF

Overview

Missing bits

Typical scenarios:

N-terminus & C-terminus missing (random or cloning artifacts)
Single amino acid missing (or small loop)
Large internal loop missing

Two simple solutions:

(1) Use psfgen to add internal coordinates to fill gaps & then minimize

(2) Use CHARMM (MD) to add internal coordinates & minimize

Internal loops: we will then need to run some “simulated annealing” and/or Torsion Angle MD dynamics to sample configurations.

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN III (from yesterday)

protein

Sample input file (here named: build_system_protein.psfgen):

>psfgen build_system_protein.psfgen >& build_protein.out &

```
topology /home/mdschool/toppar/top_all27_prot_na.inp

alias residue HIS HSE

segment PCLN {
 first NTER
 last CTER
 pdb
}

coordpdb

guesscoord

writepsf output_building/

writepdb output_building/
```

sequence file only needs on atom per residue

path to pdb file (sequence)

path coordinates file (pdb)

use IC to add H etc.

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

STEP 1: Create new psf and pdb files where psfgen will add missing bits using edited PDB files as input.

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.

STEP 2: Set up a “fixed-atoms” PDB file so that structure is not pulled apart when you minimize the structure

STEP 3: Write NAMD input file with the addition of:
fixed atoms
IMD (so you can watch)
more minimization steps

STEP 4: Minimize the structure using NAMD (repeat w/o fixed atoms)

STEP 5: If all goes well, run a short NVE dynamics run (10 ps) . . .

If step 5 fails then go back to step 4 or maybe back to step 1 !!!

EPSRC
Engineering and Physical Sciences Research Council

NSF

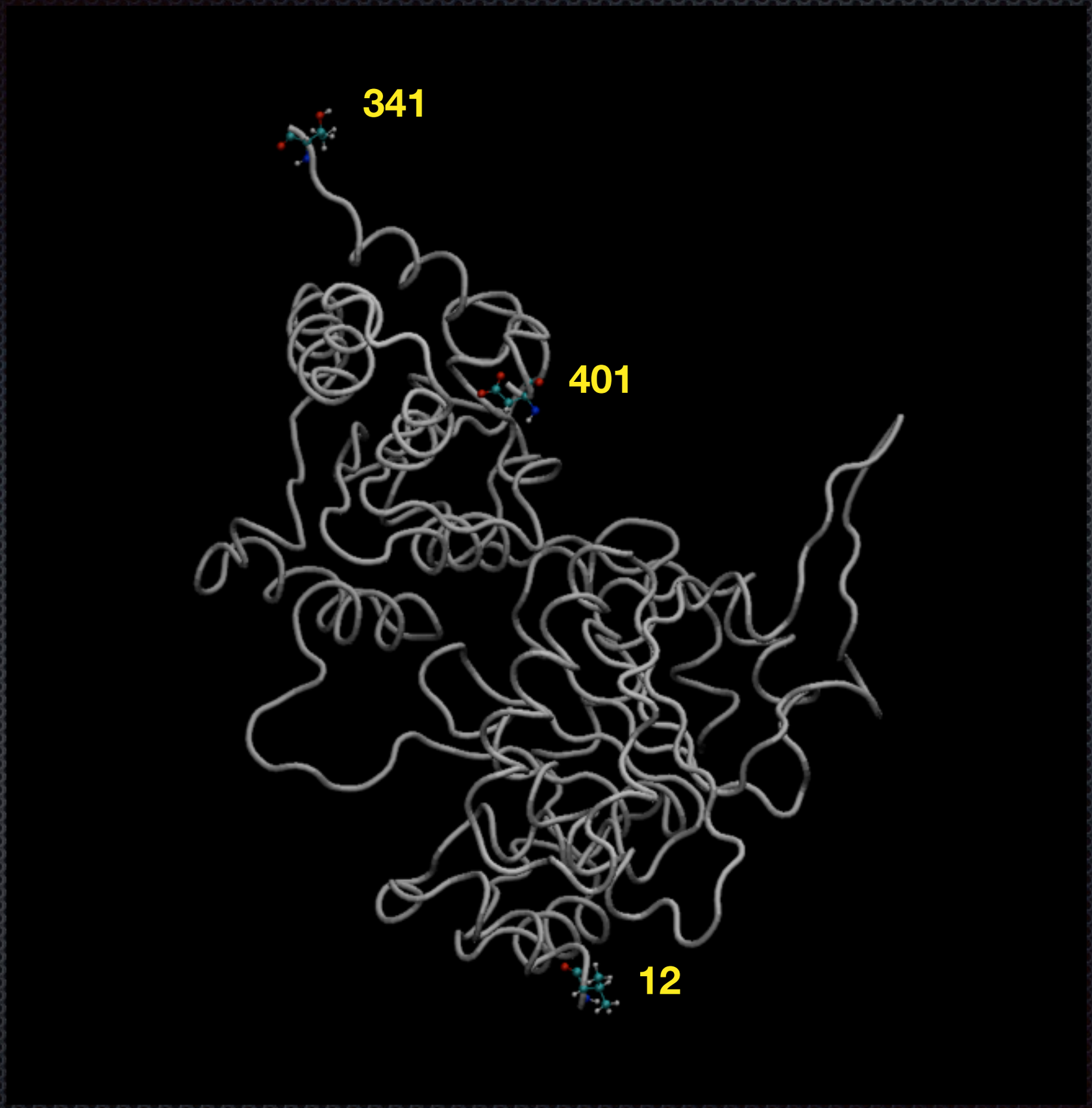
Toy system

a piece of Tral

Missing residues 1 - 11 and 342 - 400 (but we know the sequence)

EPSRC
Engineering and Physical Sciences Research Council

NSF



EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

STEP 1: Create new psf and pdb files where psfgen will add missing bits using edited PDB files as input.

Have the sequence at hand . . .

```
TSGIHVLDEL SVRALSRDIM KQNRVTVHPE KSVPRTAGYS DAVSVLAQDR       50
PSLAIVSGQG GAAGQRERVA ELVMMAREQG REVQIIAADR RSQMNMKQDE      100
RLSGELITGR RQLLEGMAFT PGSTVIVDQG EKLSLKETLT LLDGAARHNV      150
QVLITDSGQR TGTGSALMAM KDAGVNTYRW QGGEQRPATI ISEPDRNVRY      200
ARLAGDFAAS VKAGEESVAQ VSGVREQAIL TQAIRSELKT QGVLGLPEVT      250
MTALSPVWLD SRSRYLRDMY RPGMVMEQWN PETRSHDRYV IDRVTAQSHS      300
LTLRDAQGET QVVRISSLDS SWSLFRPEKM PVADGERLRV TGKIPGLRVS      350
GGDRLQVASV SEDAMTVVVP GRAEPATLPV SDSPFTALKL ENGWVETPGH      400
SVSDSATVFA SVTQMAMDNA TLNGLARSGR DVRLYSSLDE TRTAEKLARH      450
```

(a) you have one pdb file with known coordinates

(b) make a copy of (a) with only the CA atoms, add in unknown residues to make a "sequence.pdb" file. Coordinates in the sequence file do no matter.

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

STEP 1: Create new psf and pdb files where psfgen will add missing bits using edited PDB files as input.

> cat coords.pdb | grep CA > ca_sequence.pdb

then, copy known coordinate line (say, ATOM 163 CA VAL ...) and manually change the index, resname, and resid fields. Do this for all missing atoms/AA.

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

now, create a temporary pdb/psf to build the structure

Sample input file (here named: build_system_protein.psfgen):

>psfgen build_system_protein.psfgen >& build_protein.out &

```
topology /home/mdschool/research/toppar/top_all27_prot_na.inp

alias residue HIS HSE

segment
 first NTER
 last CTER
 pdb output_building/
}

coordpdb

guesscoord

writepsf output_building/

writepdb output_building/
```

path to pdb file (sequence)

path coordinates file (pdb)

use IC to add H and missing AA

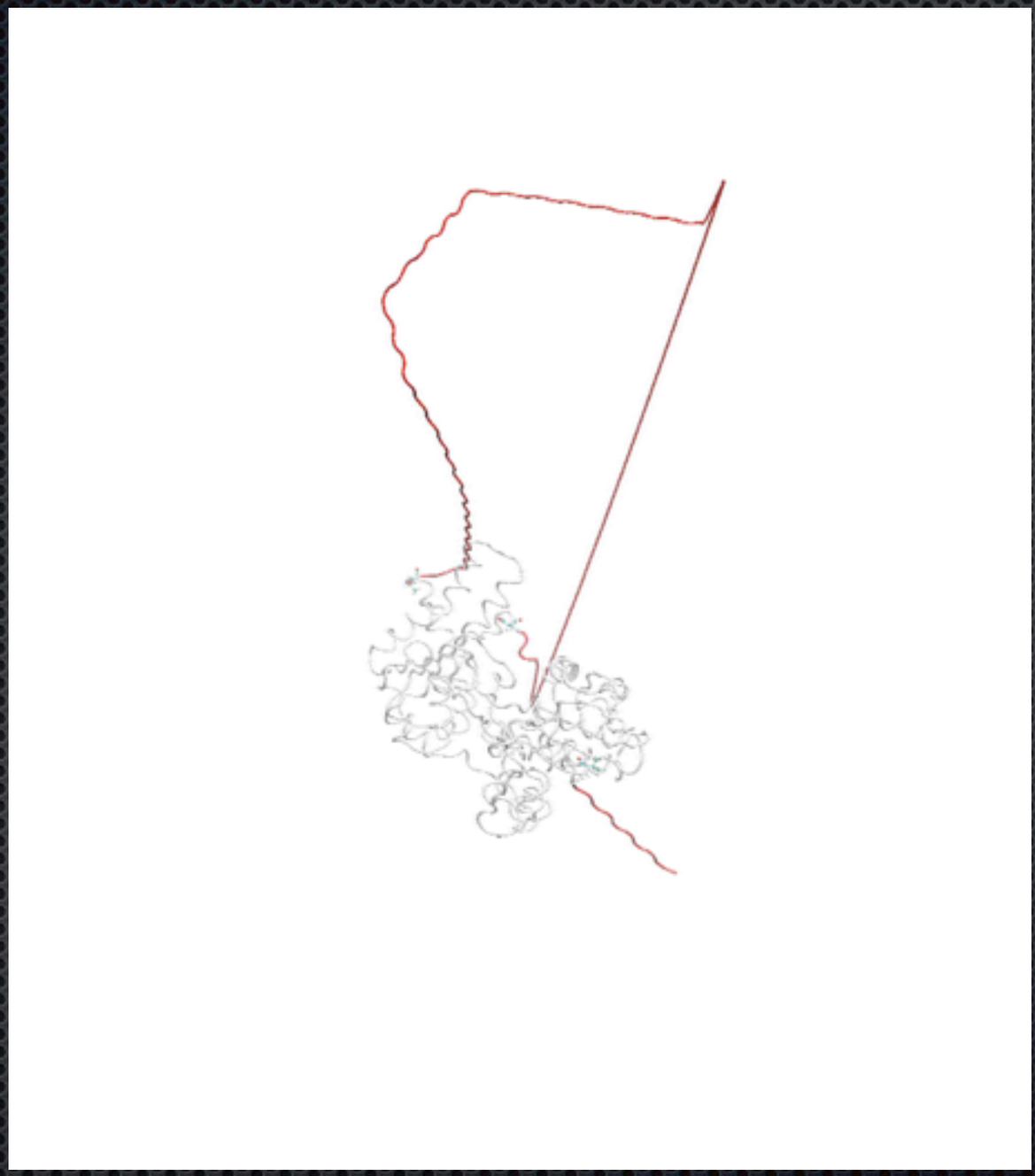
IC will “grow” the chain **blindly** . . .

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.

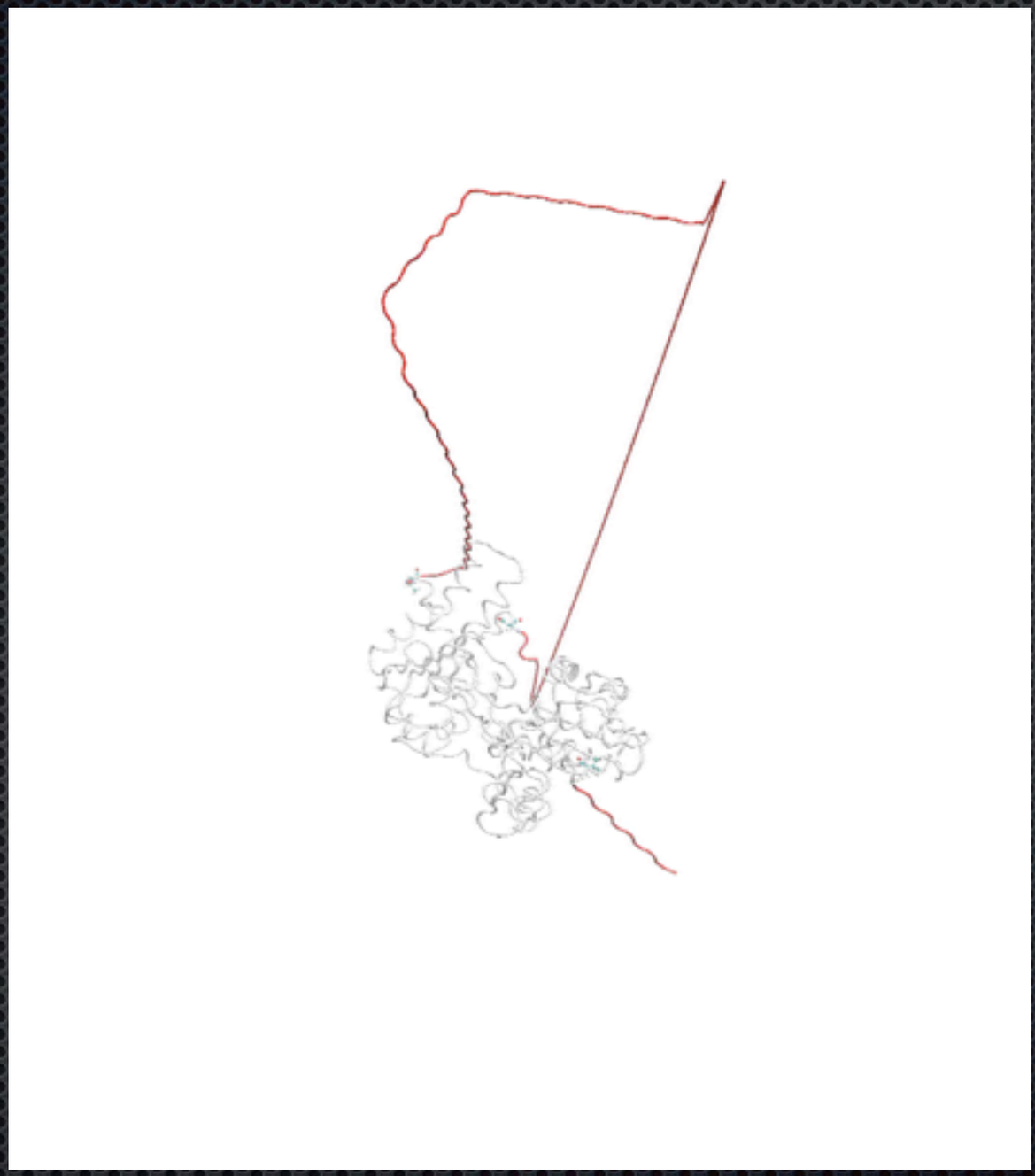


EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.

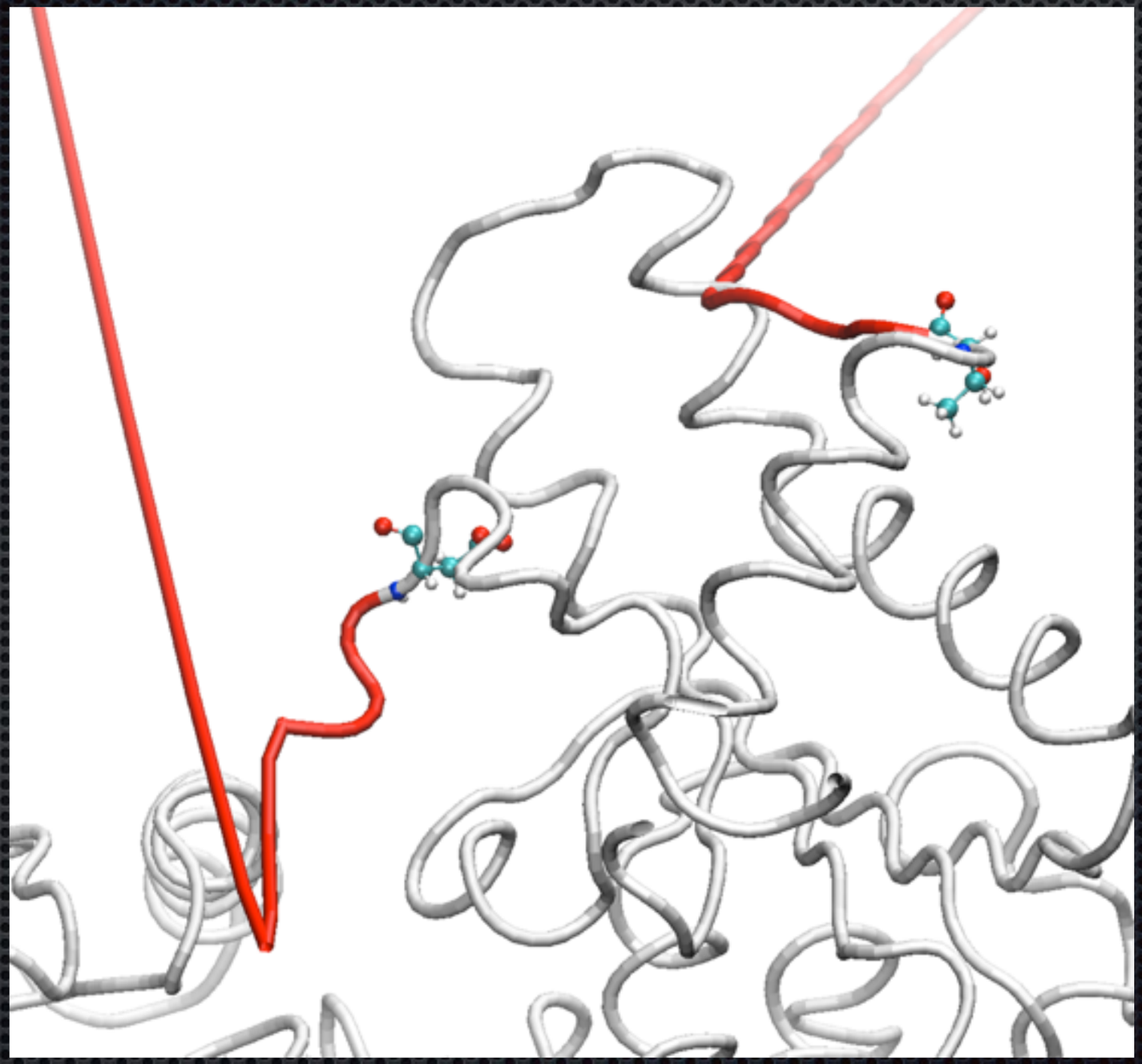


EPSRC
Engineering and Physical Sciences Research Council

NSF

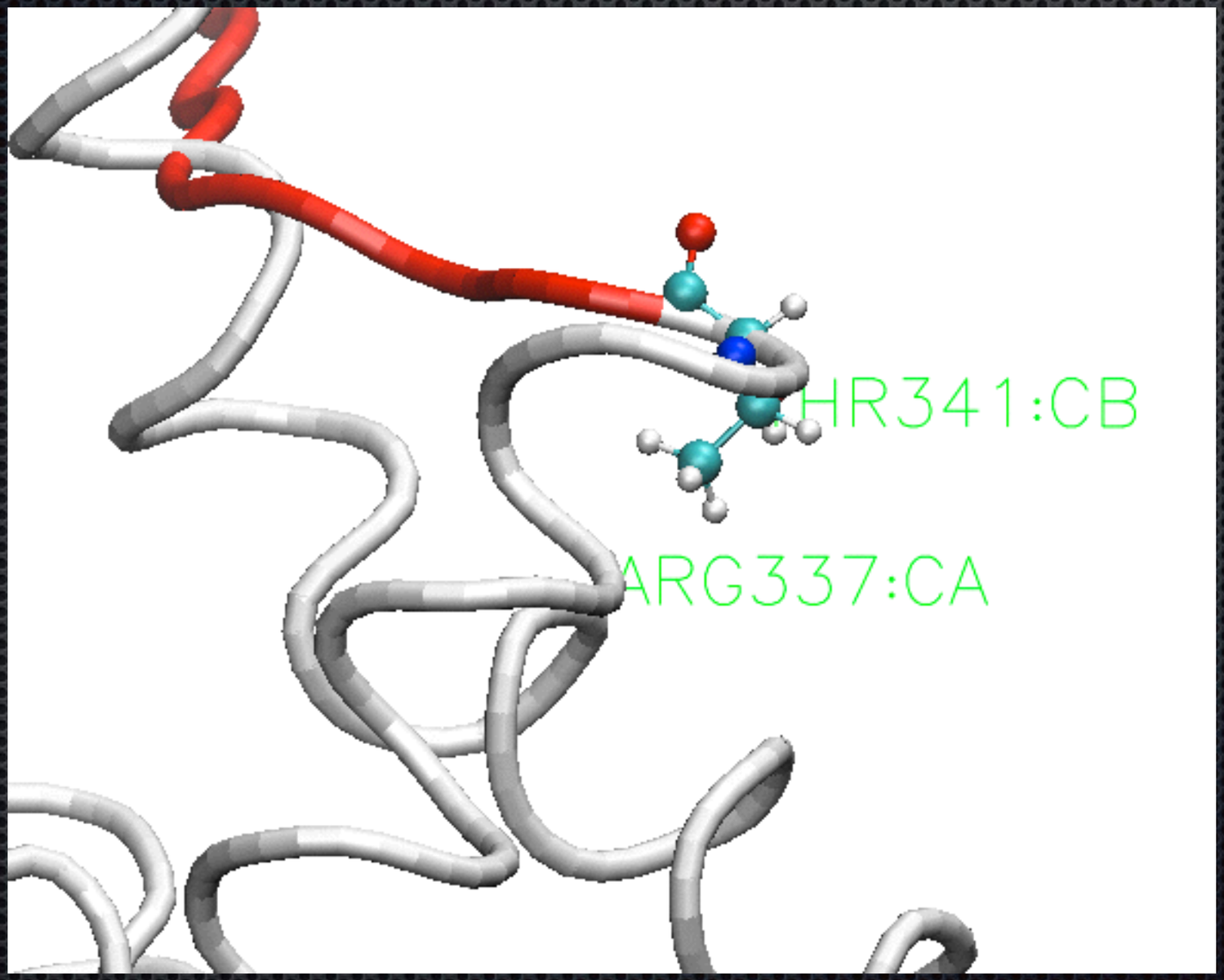
PSFGEN (filling gaps)

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.



PSFGEN (filling gaps)

YIKES! **Welcome to the real world!!!** **We have overlap. Let's remove some known coordinates near the overlap and re-build the structure. You don't need to change the sequence file. Call the new coordinate file: coords_tral_10_gap_481.pdb**



EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

YIKES! Welcome to the real world!!! We have overlap. Let's remove some known coordinates near the overlap and re-build the structure. You don't need to change the sequence file. Call the new coordinate file: coords_tral_10_gap_481.pdb

```
. . .

ATOM   1621    N LEU D 338     -62.123 -15.211 -65.673      207.57           N
ATOM   1622   CA LEU D 338     -62.146 -15.489 -64.287      207.57           C
ATOM   1623    C LEU D 338     -62.455 -14.115 -63.597      207.57           C
ATOM   1629    N ARG D 339     -63.258 -13.329 -64.312      274.04           N
ATOM   1630   CA ARG D 339     -63.760 -11.991 -63.952      274.04           C
ATOM   1631    C ARG D 339     -65.237 -12.041 -64.220      274.04           C
ATOM     49    N ALA D 406     -47.335 -18.204 -62.224      192.55           N
ATOM     50   CA ALA D 406     -47.752 -17.411 -63.397      192.55           C
ATOM     51    C ALA D 406     -46.685 -16.573 -64.078      192.55           C
ATOM     59    N THR D 407     -46.016 -17.138 -65.075      244.03           N
ATOM     60   CA THR D 407     -45.050 -16.361 -65.852      244.03           C
ATOM     61    C THR D 407     -45.999 -16.025 -66.999      244.03           C

. . .
```

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

now, create a temporary pdb/psf to build the structure

Sample input file (here named: build_system_protein.psfgen):

>psfgen build_system_protein_II.psfgen >& build_protein_II.out &

```
topology /home/mdschool/toppar/top_all27_prot_na.inp

alias residue HIS HSE

segment
 first NTER
 last CTER
 pdb output_building/ca_
}

coordpdb

guesscoord

writepsf output_building/

writepdb output_building/
```

path to pdb file (sequence)

path coordinates file (pdb)

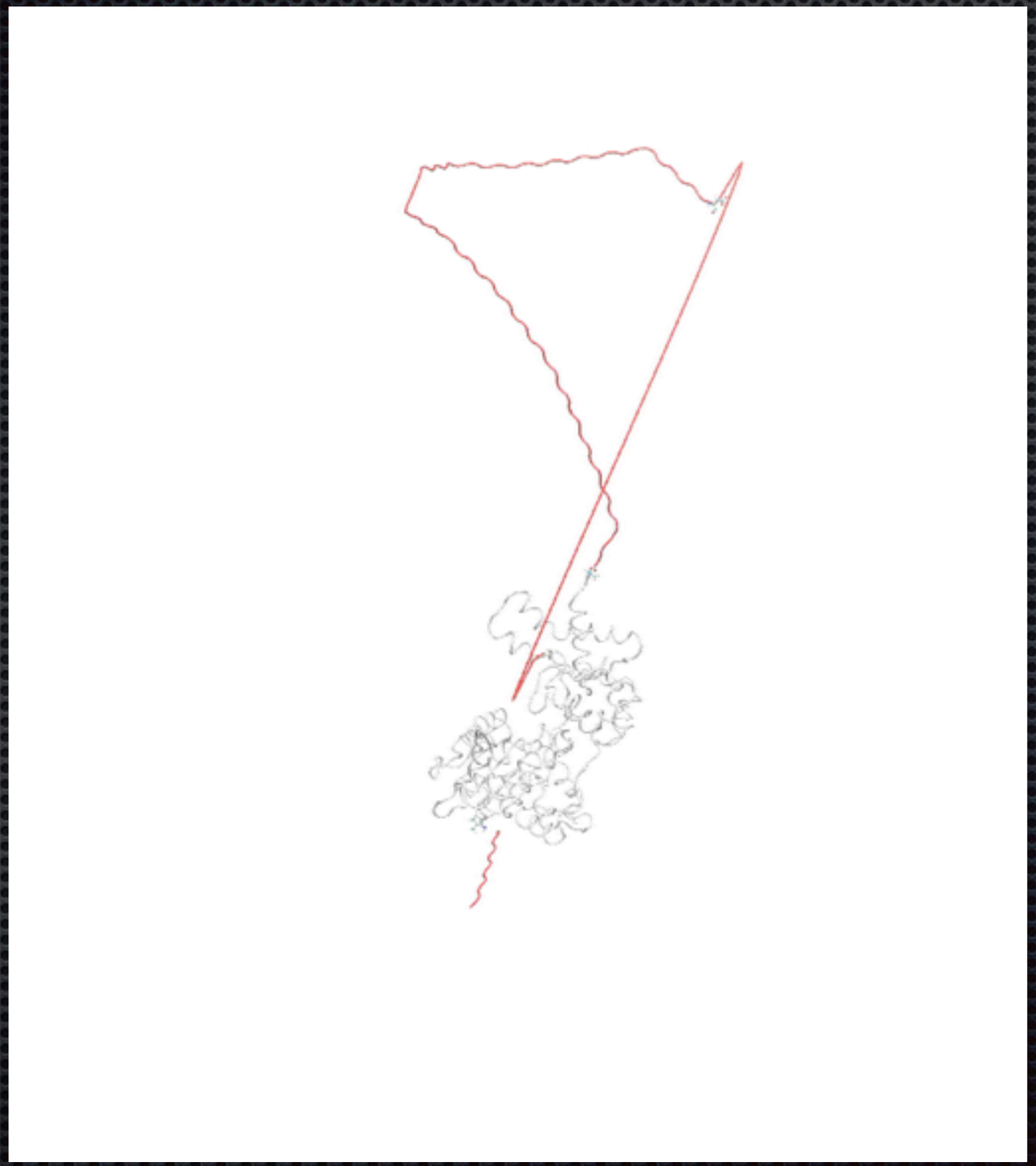
use IC to add H and missing AA

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

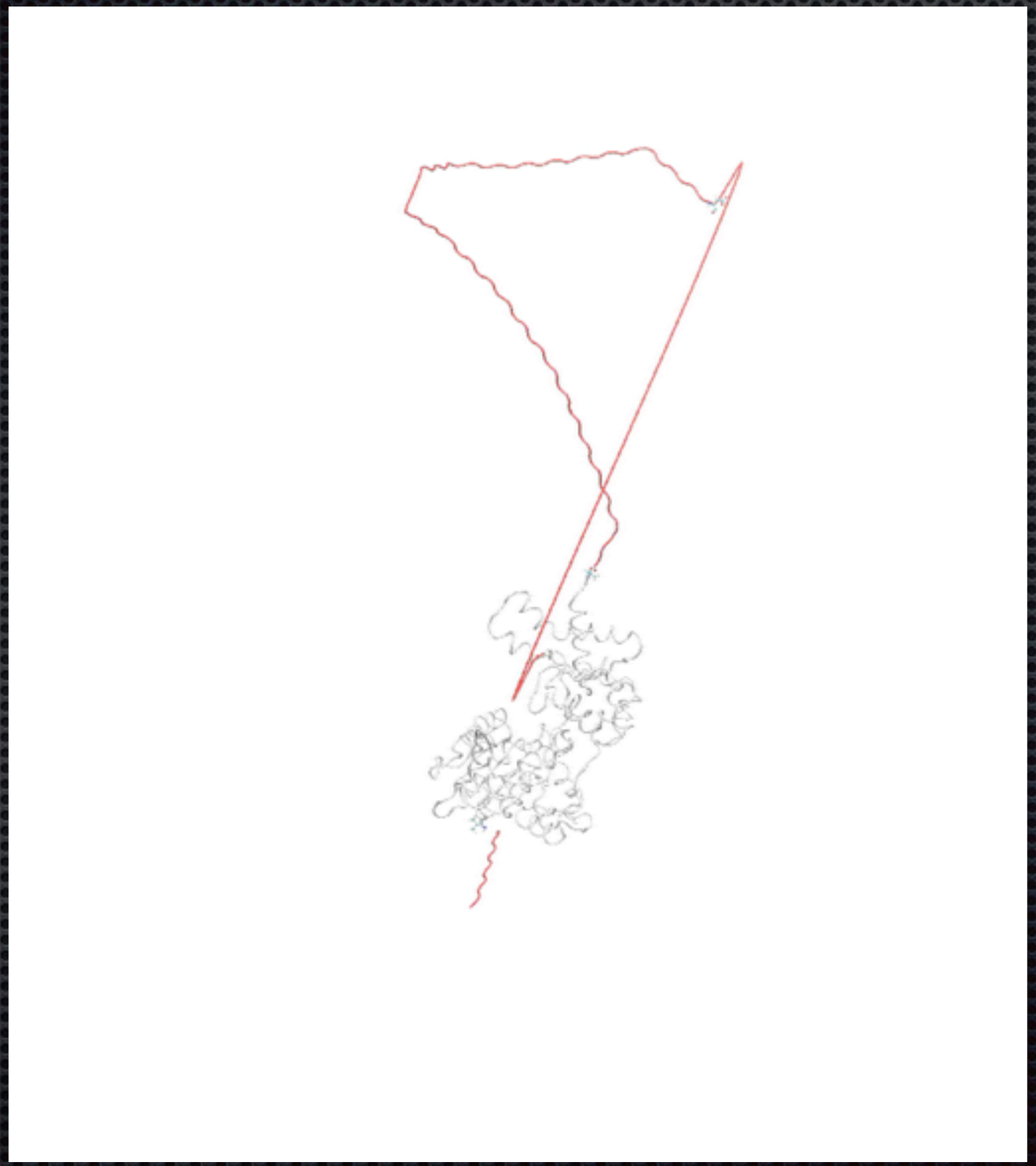
LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.



EPSRC
Engineering and Physical Sciences Research Council

PSFGEN (filling gaps)

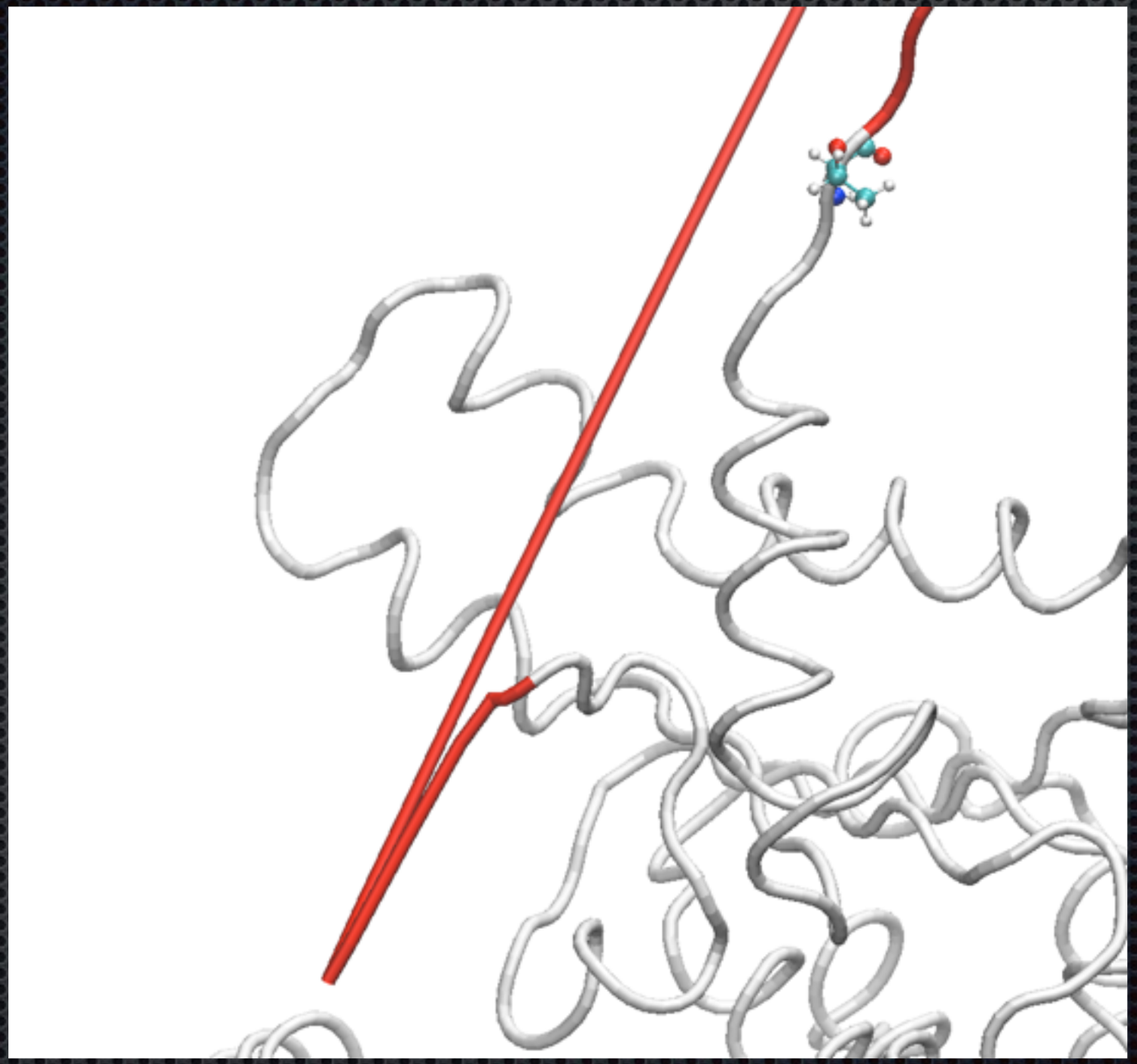
LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.



EPSRC
Engineering and Physical Sciences Research Council

PSFGEN (filling gaps)

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.



EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

STEP 1: Create new psf and pdb files where psfgen will add missing bits using edited PDB files as input.

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.

STEP 2: Set up a "fixed-atoms" PDB file so that structure is not pulled apart when you minimize the structure

STEP 3: Write NAMD input file with the addition of:
fixed atoms
IMD (so you can watch)
more minimization steps

STEP 4: Minimize the structure using NAMD (repeat w/o fixed atoms)

STEP 5: If all goes well, run a short NVE dynamics run (10 ps) . . .

If step 5 fails then go back to step 4 or maybe back to step 1 !!!

EPSRC
Engineering and Physical Sciences Research Council

NSF

New PDB I

With the **temp_2_tral_piece.pdb (psf)** file open in VMD . . . at the console type:

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

vmd > $everyone set beta 1.00

vmd > set sel1 [atomselect top “(resid > 341 and resid < 406) or (resid <
12)”]
atomselect33

vmd > $sel1 set beta 0.00
```

```
vmd > $everyone writepdb fixed_atoms.pdb

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

New PDB I

With the **temp_2_tral_piece.pdb (psf)** file open in VMD . . . at the console type:

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

vmd > $everyone set beta 1.00

vmd > set sel1 [atomselect top “(resid > 341 and resid < 406) or (resid <
12)”]
atomselect33

vmd > $sel1 set beta 0.00
```

```
vmd > $everyone writepdb fixed_atoms.pdb

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

```
. . .

ATOM   6203  C   SER P 405     -46.159 -18.127 -61.621  1.00  0.00       PCLN C
ATOM   6204  O   SER P 405     -45.252 -17.351 -61.914  0.00  0.00       PCLN O
ATOM   6205  N   ALA P 406     -47.335 -18.204 -62.224  1.00  1.00       PCLN N
ATOM   6206  HN  ALA P 406     -48.017 -18.846 -61.875  0.00  1.00       PCLN H

. . .
```

PSFGEN (filling gaps)

STEP 1: Create new psf and pdb files where psfgen will add missing bits using edited PDB files as input.

LOOK AT THE STRUCTURE BEFORE MOVING TO STEP 2.

STEP 2: Set up a “fixed-atoms” PDB file so that structure is not pulled apart when you minimize the structure

STEP 3: Write NAMD input file with the addition of:
fixed atoms
IMD (so you can watch)
more minimization steps

STEP 4: Minimize the structure using NAMD (repeat w/o fixed atoms)

STEP 5: If all goes well, run a short NVE dynamics run (10 ps) . . .

If step 5 fails then go back to step 4 or maybe back to step 1 !!!

EPSRC
Engineering and Physical Sciences Research Council

NSF

NAMD INPUT FILE (min0)

```
# sample NAMD configuration file for Minimization

# molecular system
coordinates       output/temp_2_tra_piece.pdb
structure         output/temp_2_tra_piece.psf
temperature       300

# restraints

fixedAtoms                on
fixedAtomsForces          on
fixedAtomsFile            output/fixed_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

# output

outputname        output/min0
   aryoutput      no

# interactive MD
```

```
fixedAtoms              on
fixedAtomsForces        on
fixedAtomsFile          output/fixed_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

# output
outputname        output/min0
binaryoutput      no

# interactive MD

IMDon                   yes
IMDport                 1032
IMDfreq                 10
IMDwait                 no

# run control

minimize          10000
```

you will get ALOT of warnings !!!

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 &

EPSRC
Engineering and Physical Sciences Research Council

NSF

PSFGEN (filling gaps)

```
ETITLE:      TS           BOND         ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG

ENERGY:       0   3959878.3422      2109.1105      1825.6380        25.3996
-9900.8184  99999999.9999          0.0000          0.0000          0.0000
99999999.9999          0.0000  99999999.9999  99999999.9999          0.0000

INITIAL STEP: 1e-06
GRADIENT TOLERANCE: 1.07792e+08
Warning: Bad global exclusion count, possible error!
Warning: Increasing cutoff during minimization may avoid this.

. . .

GRADIENT TOLERANCE: 5.17843
ETITLE:      TS           BOND         ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG

ENERGY:   10000     14400.1513      4244.3536      1948.5108        37.2996
-10260.9634  99999999.9999          0.0000          0.0000          0.0000
99999999.9999          0.0000  99999999.9999  99999999.9999          0.0000

WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 10000
WRITING COORDINATES TO OUTPUT FILE AT STEP 10000
WRITING VELOCITIES TO OUTPUT FILE AT STEP 10000
====================================================
```

EPSRC
Engineering and Physical Sciences Research Council

NSF

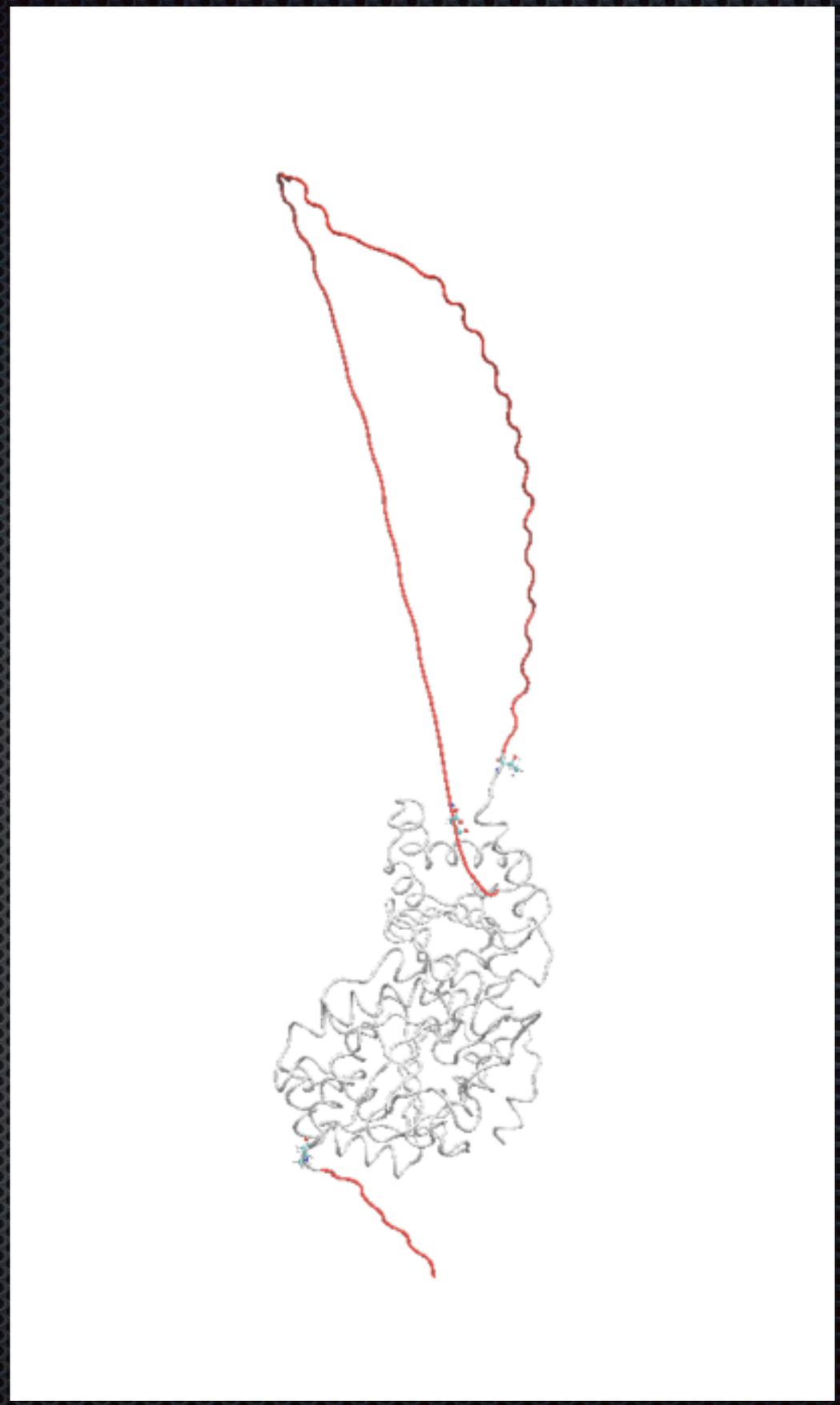
PSFGEN (filling gaps)

```
ETITLE:      TS           BOND         ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG

ENERGY:       0   3959878.3422      2109.1105      1825.6380        25.3996
-9900.8184  99999999.9999            0.0000         0.0000         0.0000
99999999.9999          0.0000  99999999.9999  99999999.9999         0.0000

INITIAL STEP: 1e-06
GRADIENT TOLERANCE: 1.07792e+08
Warning: Bad global exclusion count, possible error!
Warning: Increasing cutoff during minimization may avoid this.

. . .

GRADIENT TOLERANCE: 5.17843
ETITLE:      TS           BOND         ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG

ENERGY:   10000     14400.1513      4244.3536      1948.5108        37.2996
-10260.9634  99999999.9999            0.0000         0.0000         0.0000
99999999.9999          0.0000  99999999.9999  99999999.9999         0.0000

WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 10000
WRITING COORDINATES TO OUTPUT FILE AT STEP 10000
WRITING VELOCITIES TO OUTPUT FILE AT STEP 10000
====================================================
```

try continuing minimization w/o fixed atoms

EPSRC Engineering and Physical Sciences Research Council

NSF



EPSRC
Engineering and Physical Sciences Research Council

MIN1 (w/o fixed atoms)

```
TCL: Minimizing for 10000 steps
ETITLE:      TS           BOND          ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG
ENERGY:       0     14400.1137      4244.3463      1948.5042        37.3339
-10260.8912  99999999.9999          0.0000          0.0000          0.0000
99999999.9999          0.0000  99999999.9999  99999999.9999          0.0000

INITIAL STEP: 1e-06
GRADIENT TOLERANCE: 2.44629e+09

. . .

GRADIENT TOLERANCE: 0.048755
TIMING: 10000  CPU: 1535.9, 0.150397/step  Wall: 1481.5, 0.144865/step, 0
hours remaining, 752124 kB of memory in use.
ETITLE:      TS           BOND          ANGLE          DIHED          IMPRP
ELECT             VDW        BOUNDARY           MISC        KINETIC
TOTAL          TEMP          TOTAL2         TOTAL3        TEMPAVG

ENERGY:   10000       471.9604      1640.2784      2439.1727        77.2777
-17646.7352     -1541.1823          0.0000          0.0000          0.0000
-14559.2282          0.0000    -14559.2282    -14559.2282          0.0000

WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 10000
WRITING COORDINATES TO OUTPUT FILE AT STEP 10000
WRITING VELOCITIES TO OUTPUT FILE AT STEP 10000
==========================================
```

DYN0: (IN LAB)

EPSRC Engineering and Physical Sciences Research Council

NSF

DYN0: (IN LAB)

EPSRC
Engineering and Physical Sciences Research Council

NSF

Method 2 (CHARMM)

Missing bits

(2) Use CHARMM (MD) to add internal coordinates & minimize in fewer steps

STEP 1: Create new a pdb file with the the full-length sequence and a pdb file with the known coordinates (we know which one to use)

STEP 2: Write a CHARMM input file to read sequence file, add bits, constrain AA, and minimize

STEP 3: If all goes well, run a short NVE dynamics run (10 ps) . . .

EPSRC
Engineering and Physical Sciences Research Council

NSF

CHARMM INPUT FILE (filling gaps)

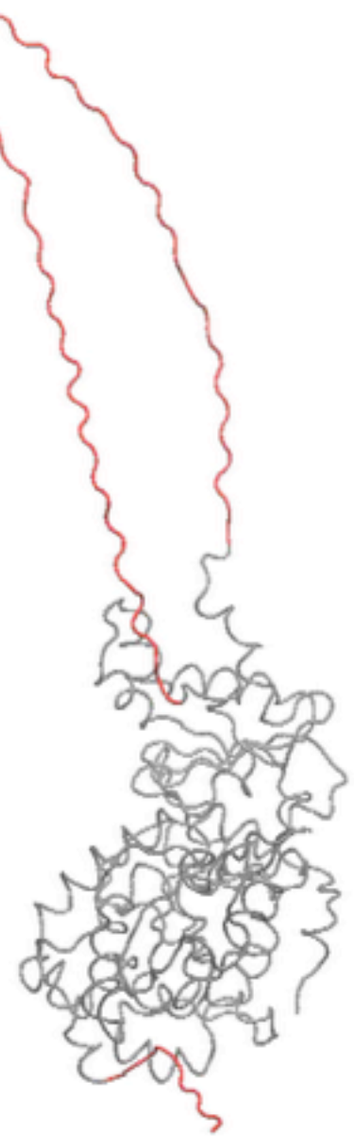
fill gaps, minimize w/ constraints, minimize w/o constaints, then 10 ps NVE MD

Sample input file (here named: nbuild_system.inp):

>charmm < nbuild_system.inp >& cbuild.out &

```
* fill in missing amino acids, minimize with constraints, then
* minimize without constraints

open unit 1 read form name ~/toppar/top_all27_prot_na.inp
read rtf card unit 1
close unit 1

open unit 1 read form name ~/toppar/par_all27_prot_na.inp
read param card unit 1
close unit 1

open unit 1 read card name
read sequence pdb unit 1
close unit 1

generate

open unit 1 read card name
read unit 1 coor pdb
close unit 1

!ic seed
ic fill preserve
ic param
ic build

open unit 1 write form name
write coor pdb unit 1
* non-minimized traI piece
* by jc 06/26/11

close unit 1
```

```
Minimise

! Save the optimised coordinates.
Open write card unit 17 name
Write coor pdb unit 17
* minimized region 2
* by jc 6/26/11
*

close unit 17
CONS FIX SELE NONE END

Energy ihbfrq 0 inbfrq 10 imgfrq 10 cutim 999.0

! Perform the minimisation.
Minimise

! Save the optimised coordinates.
Open write card unit 17 name
Write coor pdb unit 17
* fully minimized region 2
* by jc 6/26/11
*
close unit 17

! now run a short NVE dynamics run
! open the trajectory file
open unit 50 write unform name trajectory.trj

! IHTFRQ & IEQFRQ are, by default 0, so since
! we don't set them here, this runs with the
! NVE ensemble
dyna leap start
     nprint 1000 nsavc 50 iuncrd 50 ntrfrq 5000

! Save the optimised coordinates.
Open write card unit 17 name
Write coor pdb unit 17
* fully minimized and 10 ps of NVE dynamics region 2
* by jc 6/26/11
*
close unit 17

stop
```



EPSRC
Engineering and Physical Sciences Research Council

Overview

Intermediate Building Tricks

Building coordinates for missing atoms/residues

Simulated annealing

Building IDPs
A trick for topological issues

EPSRC
Engineering and Physical Sciences Research Council

NSF

SA

Simulated Annealing (SA)

May wish to relax the structure further

May wish to explore configurations (NMR, FRET, other constraints)

For IDPs our “software” can not handle INTERNAL disordered loops

Write a TCL script (a list of commands) at the end of one of our “dyn1” NAMD input files to carry out SA.

The idea is simple. Run a series of heating and cooling cycles of your system and at the end of each cooling cycle you minimize and save the structure.

Heating allows you to cross energetic barriers.

An ensemble of such structures is representative of what might exist . . .

EPSRC
Engineering and Physical Sciences Research Council

NSF

SA Protocol (100 cycles)

```
. . .

for { set x 0} {$x <
if {$x == 0} {
checkpoint
print "setting checkpoint structure"
}
for { set TEMP
  run 2000
  reassignTemp $TEMP
  langevinTemp $TEMP
}
run 20000
for { set TEMP
  reassignTemp $TEMP
  langevinTemp $TEMP
  run 2000
}
run 2000
minimize 500
output output/sa0_min_$x
#revert
}
```

at bottom of regular namd input file

save initial structure

HEAT UP

COOL DOWN

Minimize and save structure

you probably want to used fixed atoms!!!

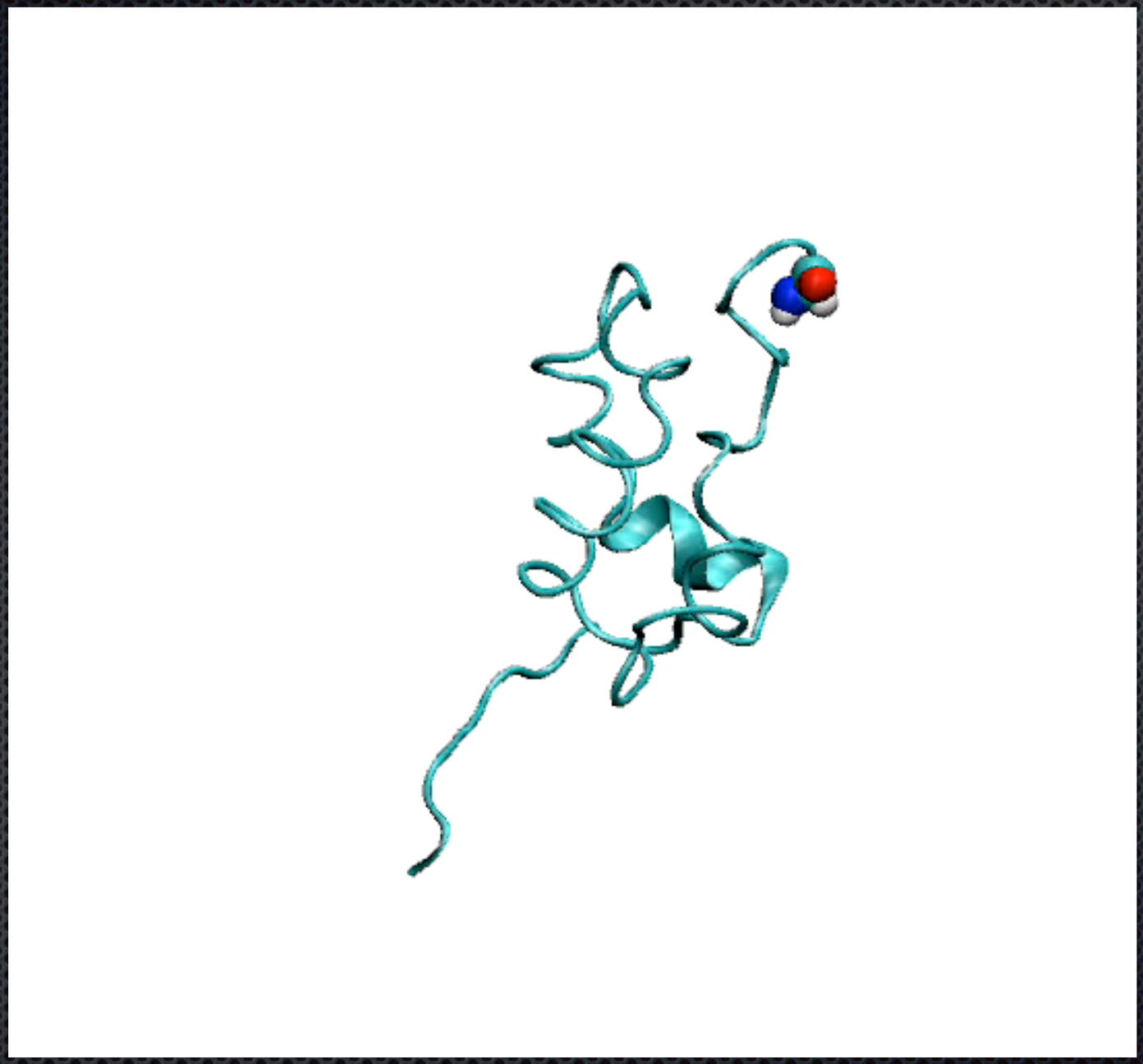
EPSRC
Engineering and Physical Sciences Research Council

NSF

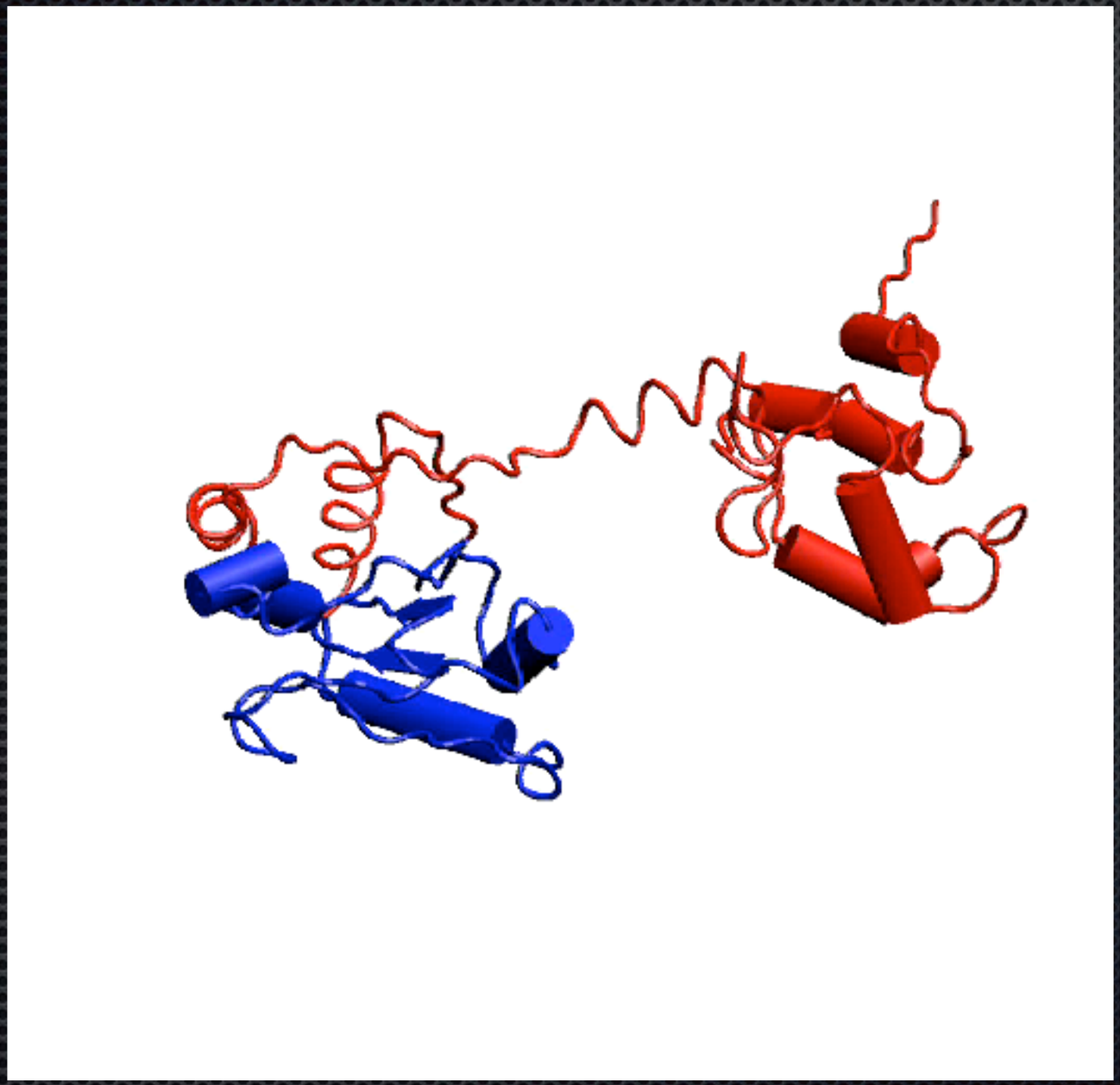
SA Examples

SA Examples

SA Examples



SA Examples



EPSRC
Engineering and Physical Sciences Research Council

SA

Runs take a long time . . . adding solvent is even more CPU intensive (implicit solvent : GBSA)

Need to concatenate all the “minimized structures”. Here is a command to read in the individual sa_min0*.coor files and make a single DCD file of the ensemble.

For this example, we’ll use the PSF file and assume that the individual sa_min0*coor files are in a directory called “output/”

```
> catdcd -o all_sa_min0.dcd -otype dcd -stype psf -s output/   .psf -filetype
namdbin output/sa_min0*coor
```

When you forget the options for this handy program, just type “catdcd” at the command line and the usage will print out.

EPSRC
Engineering and Physical Sciences Research Council

NSF

SA

```
> catdcd
CatDCD 4.0
   catdcd -o outputfile [-otype <filetype>] [-i indexfile]
      [-stype <filetype>] [-s structurefile]
      [-first firstframe] [-last lastframe] [-stride stride]
      [-<filetype>] inputfile1 [-<filetype>] inputfile2 ...


Allowed input file types:
cpmd pdb dcd Alchemy AMBERPREP BallStick MSIBGF BiosymCAR Boogie
Cacao CADPAC CHARMm Chem3d-1 Chem3d-2 CSSR FDAT GSTAT Dock
DockPDB Feature Fractional GAMESSoutput GaussianZmatrix
Gaussian92output Gaussian94output Gromos96A Gromos96N
HyperchemHIN IsisSDF M3D MacMolecule Macromodel MicroWorld
MM2Input MM2Output MM3 MMADS MDLMOL MOLIN MopacCartesian
MopacInternal MopacOutput PCModel PSGVBin PSGVBout QuantaMSF
Schakal ShelX SMILES Spartan SpartanSE SpartanMM SybylMol
SybylMol2 Conjure UniChemXYZ XYZ XED gro g96 trr trj xtc crd
crdbox namdbin binpos cube rst7 tinker POSCAR OUTCAR XDATCAR xml
dlpolyhist dlpoly3hist lammpstrj vcf vtf xyz cor molden pqr mol2
car gamess xsf bgf xbgf webpdb netcdf

Allowed output file types:
pdb dcd trr crd crdbox namdbin binpos rst7 POSCAR xyz pqr mol2
bgf xbgf
```

EPSRC
Engineering and Physical Sciences Research Council

NSF

Overview

Intermediate Building Tricks

Building coordinates for missing atoms/residues

Simulated annealing

Building IDPs

A trick for topological issues

EPSRC
Engineering and Physical Sciences Research Council

NSF

IDPs

Piece structures together (FUTURE)

1. Load individual PDB files into VMD. Use “8” to grab one and sequentially move along X, rotate system, move along Y, rotate system, move along Z. Save coordinates using VMD. ((will probably have to execute rotate TCL commands))

“align by eye” & hope for the best. Concatentate files, renumber amino acids and atoms, make new PDB/PSF pair, minimize, 10 ps NVE MD . . .

2. Add 3 to 5 amino acids to N-terminal of the C-terminal part, then overlap these amino acids (using backbone as a basis).

Then you align the two structures (C-terminal of “A” to modified N-terminal of “B”)

Concatentate PDB files, remove the “bridging” 3 to 5 AA, renumber remaining amino acids and atoms, make new PDB/PSF pair, minimize, 10 ps NVE MD . . .

3. Insert your method here . . . etc. (pymol, mdtools, etc.)

EPSRC
Engineering and Physical Sciences Research Council

NSF

IDPs

Many IDPs that we have built have been a combination of X-ray, NMR coordinates and homology models. Many regions are evaluated by predicting if a region has the potential to have SS or disorder etc.

SS:

http://cib.cf.ocha.ac.jp/bitool/MIX/
http://bioinf.cs.ucl.ac.uk:80/psipred/

Disorder:

http://dis.embl.de/
http://mbs.cbrc.jp/poodle/poodle.html

Homology modeling:

http://www.sbg.bio.ic.ac.uk/~phyre/

Do more experiments!

CD

Enzyme digests

NMR

etc.

http://www.smallangles.net/sassie/tral/Tral/Outline.html

EPSRC
Engineering and Physical Sciences Research Council

NSF

Overview

Intermediate Building Tricks

Building coordinates for missing atoms/residues
Simulated annealing
Building IDPs

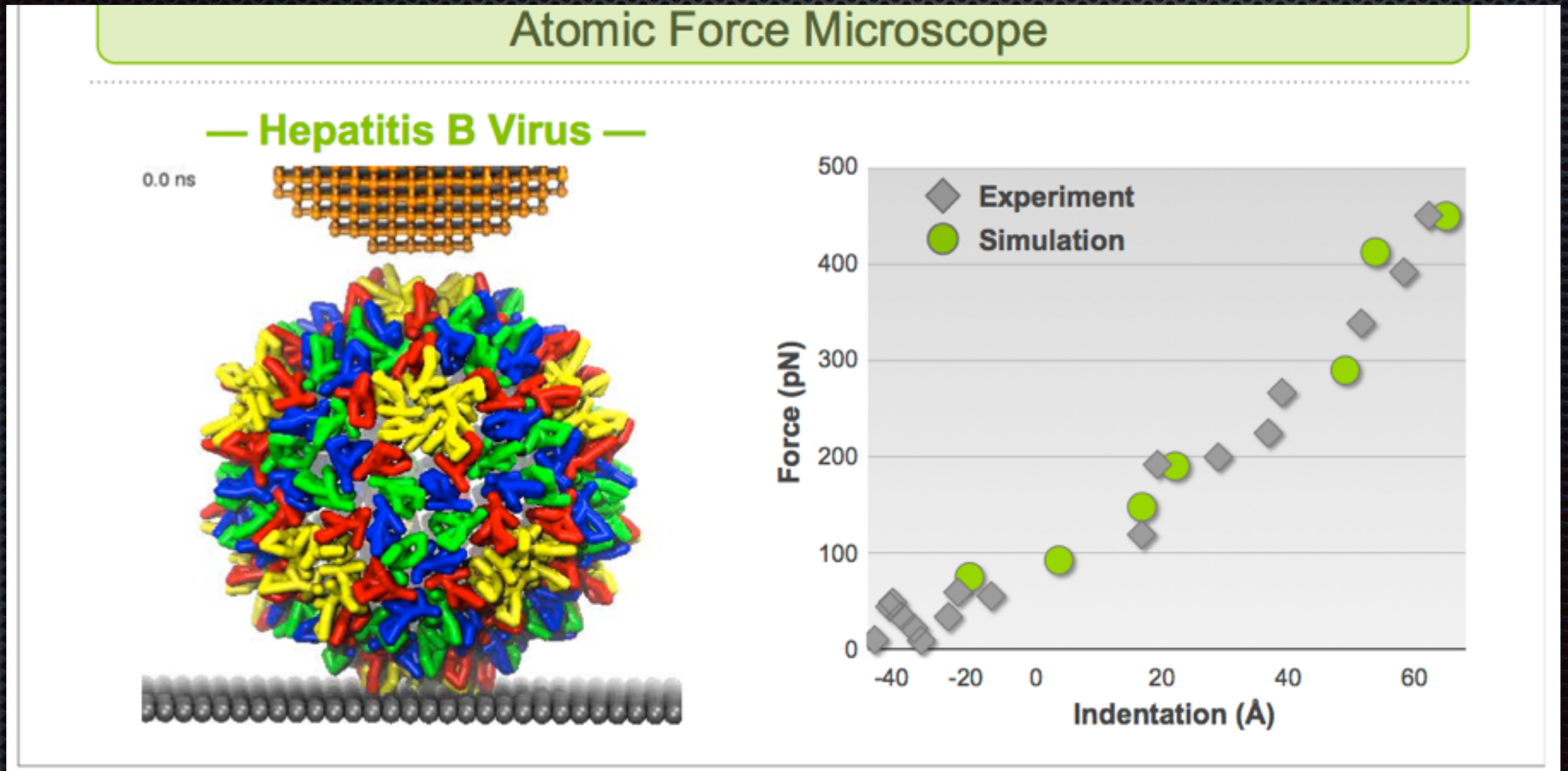
A trick for topological issues

EPSRC
Engineering and Physical Sciences Research Council

NSF

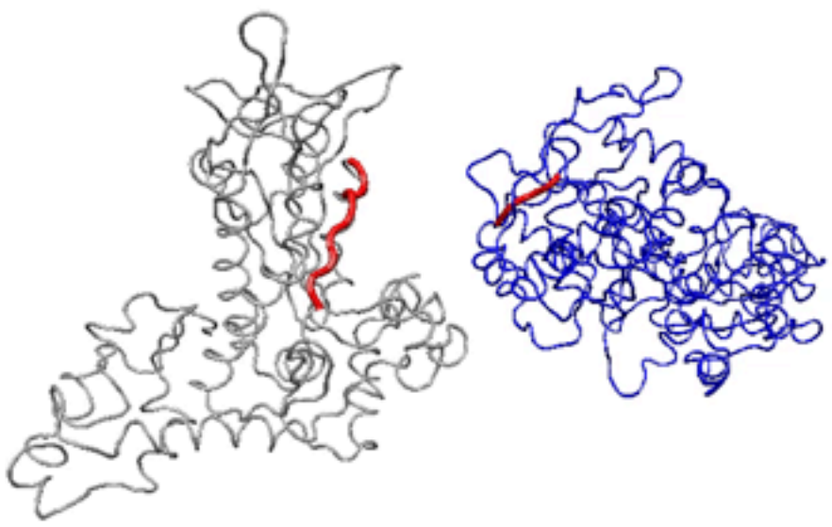
Steered MD

http://www.ks.uiuc.edu/Research/smd_imd/



Steered MD for building

Buried disordered tails ... have to pull them out to link two structures together

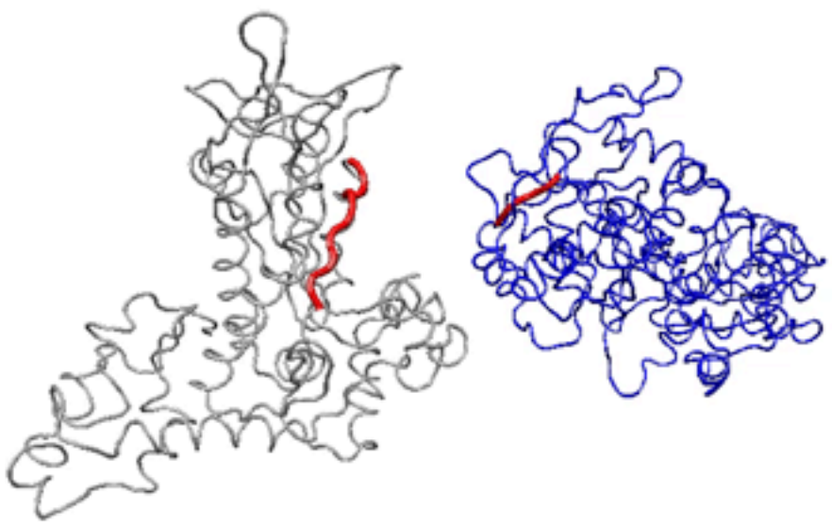


EPSRC
Engineering and Physical Sciences Research Council

NSF

Steered MD for building

Buried disordered tails ... have to pull them out to link two structures together



EPSRC
Engineering and Physical Sciences Research Council

NSF

Steered MD for building

So, in the following movie you will see:

1. files that we are working with
2. a view of the input file (min0)
3. loading the psf and then pdb into vmd
4. initiating namd2 min0
5. starting the run via IMD in VMD
6. applying the residue specifc force to pull the N-terminal out
7. a few thousand steps of MD

NOTE:

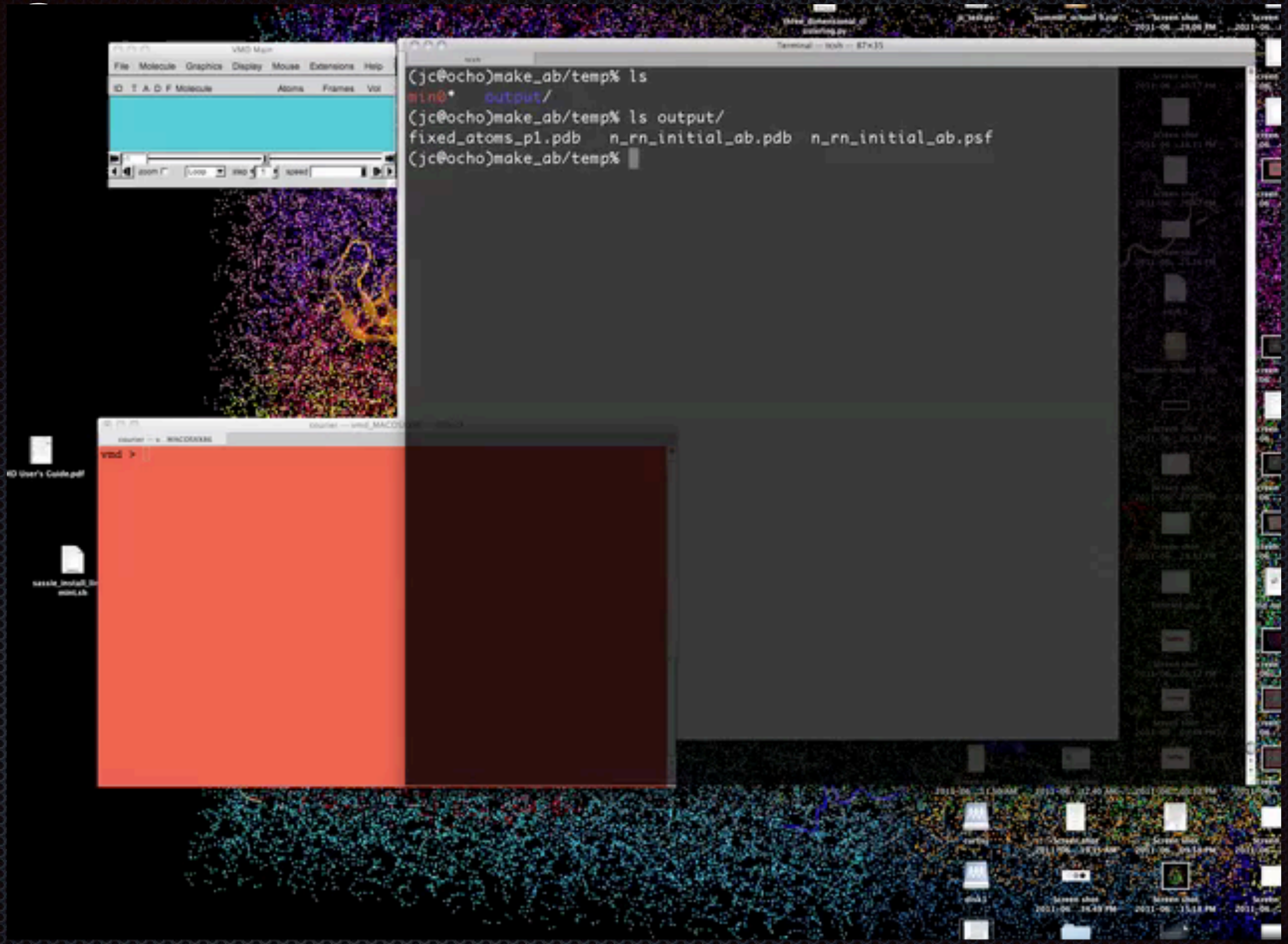
This N-terminal was known NOT to be buried and that it was unstructured. The resulting structure was minimized prior to it's use in further building steps!

This example was done in the process of building a larger IDP with many known and unknown bits.

It allowed us to overlap the N-terminal AA with the "other" piece etc.

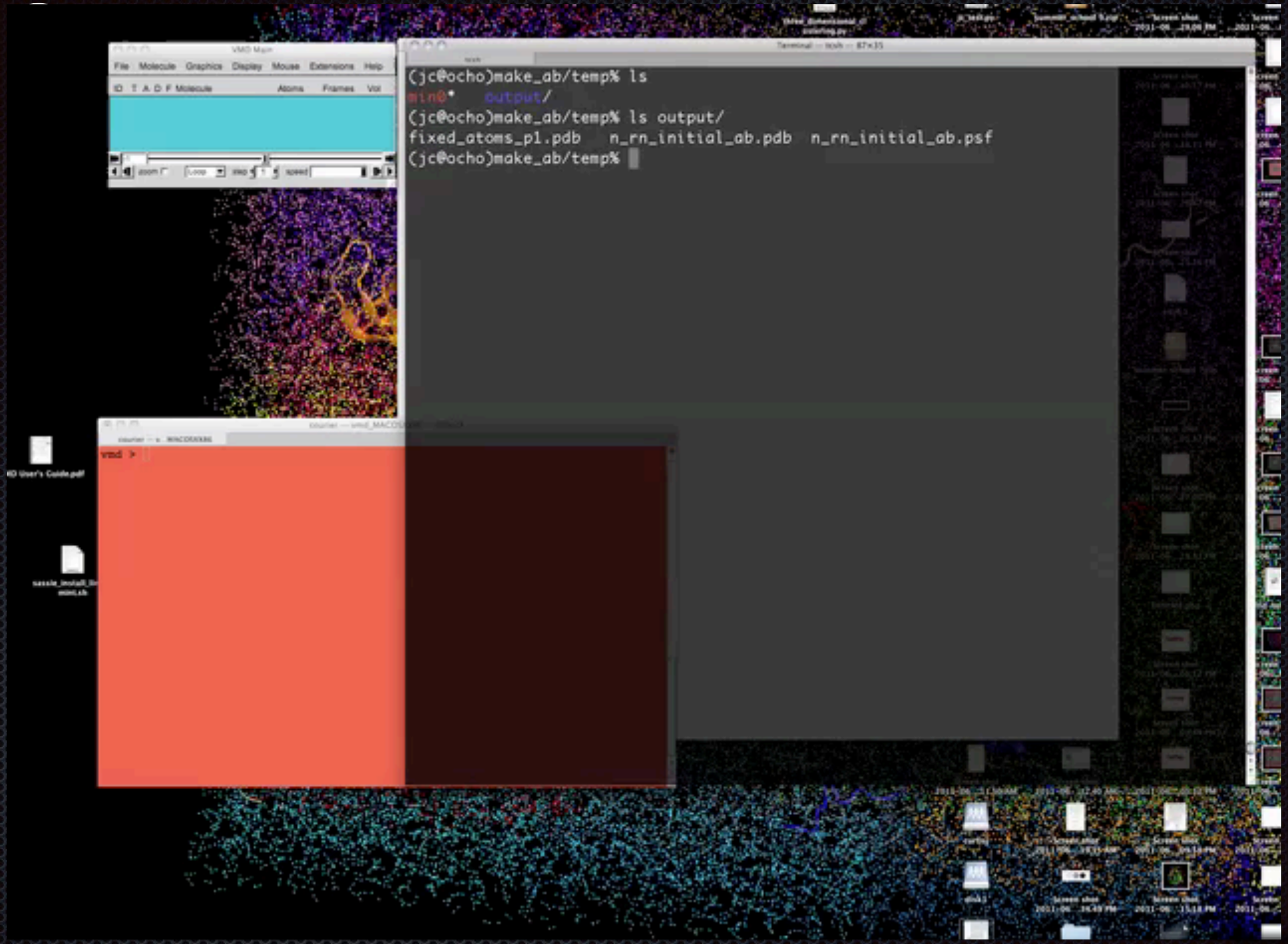
EPSRC
Engineering and Physical Sciences Research Council

NSF



EPSRC
Engineering and Physical Sciences Research Council

NSF



EPSRC
Engineering and Physical Sciences Research Council

NSF

SUMMARY

What have we accomplished?

We showed some examples of how to complete structures and covered some ways to build larger structures/models from bits of known structure.

What is next?

LAB III --> MDII & Analysis

LAB IV --> Building IDPs and STUDENT PROJECTS

EPSRC
Engineering and Physical Sciences Research Council

NSF