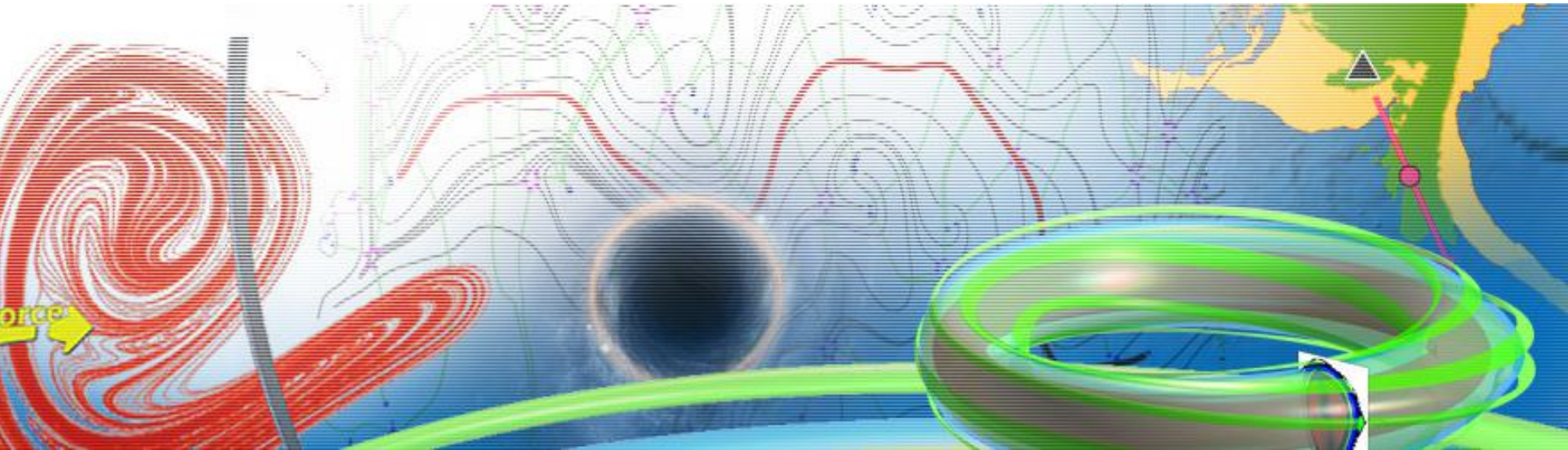
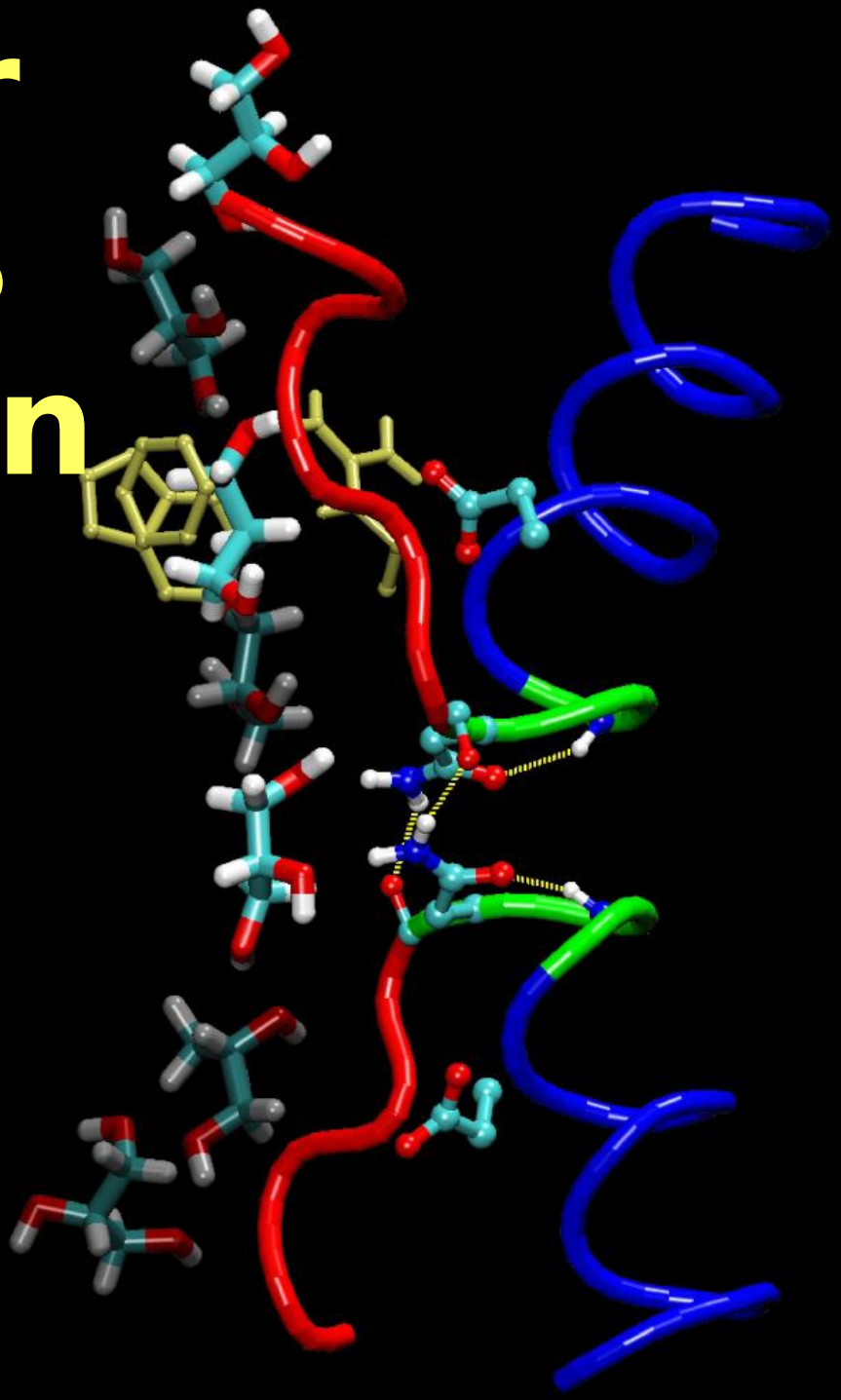
生物动力系统模拟

王冠宇 18665955633
wanggy@sustc.edu.cn

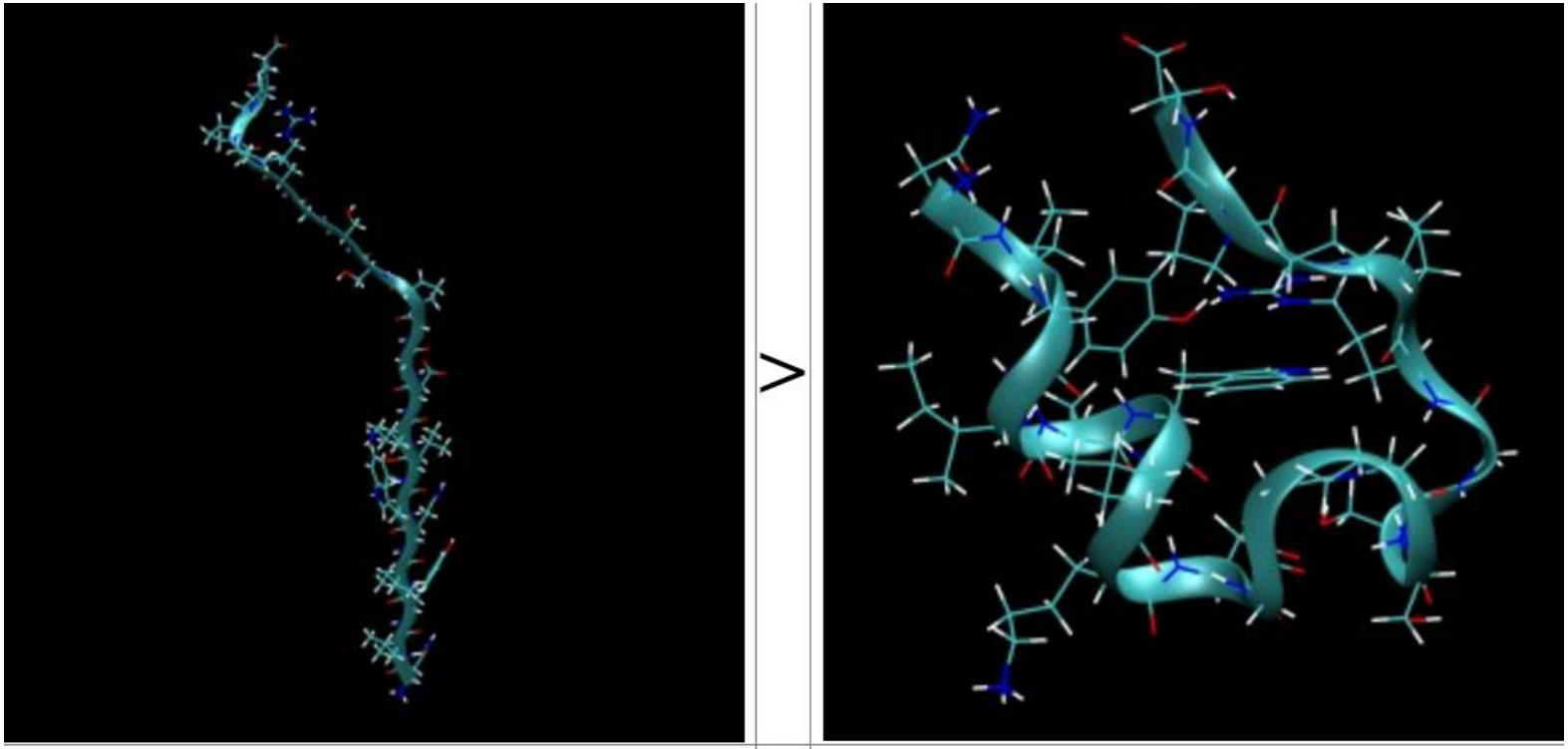
吴韵怡（助教）



Molecular Dynamics Simulation



Example 2: All Atom Structure Prediction and Folding Simulations of a Stable Protein by the software **Amber**



Initial Linear Structure

NMR Structure (1L2Y) of TRP Cage Peptide

The small size (20 aa) of this protein makes it an ideal candidate for computational foldi simulations.

When the original folding simulations were done the experimental structure had not bee solved.

When the experimental structure was solved the predicted structure was within 1.4 angstroms RMSD.

Building the starting structure

Single letter to 3 letter sequence conversion	
G	Glycine (Gly)
P	Proline (Pro)
A	Alanine (Ala)
V	Valine (Val)
L	Leucine (Leu)
I	Isoleucine (Ile)
M	Methionine (Met)
C	Cysteine (Cys)
F	Phenylalanine (Phe)
Y	Tyrosine (Tyr)
W	Tryptophan (Trp)
H	Histidine (His)
K	Lysine (Lys)
R	Arginine (Arg)
Q	Glutamine (Gln)
N	Asparagine (Asn)
E	Glutamic Acid (Glu)
D	Aspartic Acid (Asp)
S	Serine (Ser)
T	Threonine (Thr)

The amino acid sequence is:
NLYIQWLKDGGPSSGRPPPS

So the sequence can be re-written as:

ASN LEU TYR ILE GLN TRP LEU LYS ASP GLY GLY PRO SER SER GLY ARG PRO PRO PRO SER

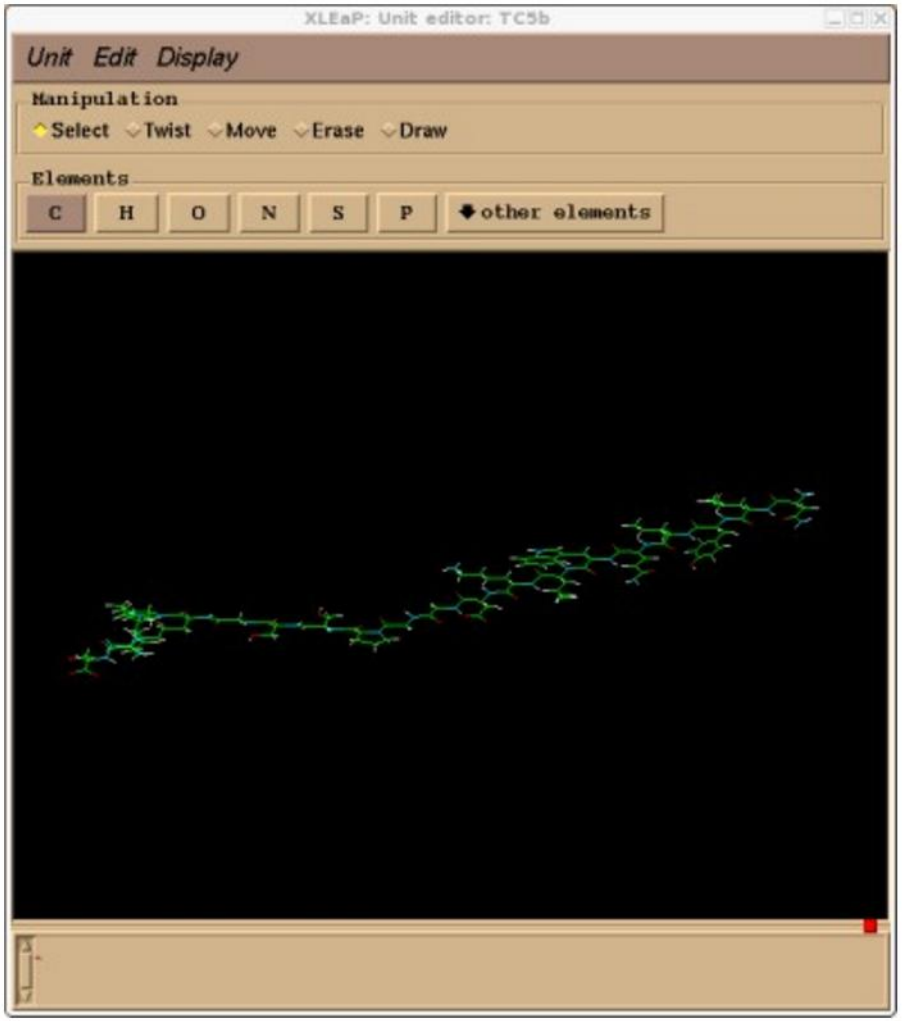
We have to specify our N terminus, by prefixing it with an N, and our C terminus with a C.

NASN LEU TYR ILE GLN TRP LEU LYS ASP GLY GLY PRO SER SER GLY ARG PRO PRO PRO CSER

then create the structure by calling the submodule XLEaP of Amber

> TC5b = sequence { NASN LEU TYR ILE GLN TRP LEU LYS ASP GLY GLY PRO SER SER GLY ARG
PRO PRO PRO CSER }

>edit TC5b

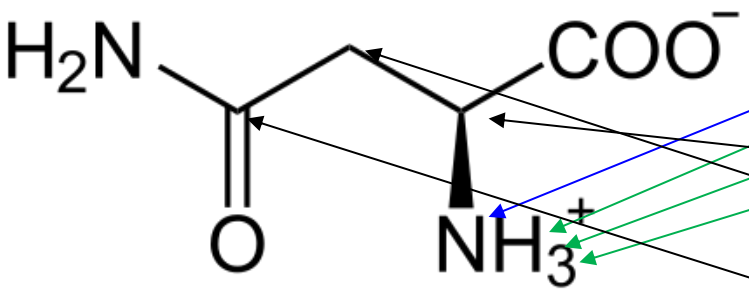

XLEaP: Unit editor: TC5b
Unit Edit Display
Manipulation
Select Twist Move Erase Draw
Elements
C H O N S P other elements

Save this structure as a library file so we can come back to it later

> saveoff TC5b TC5b_linear.lib

We will also save a pdb file of this structure so we can easily visualize it

>savepdb TC5b TC5b_linear.pdb



asparagine

occupancy **temperature factor**

					X	Y	Z		
ATOM	1	N	ASN	1	3.326	1.548	-0.000	1.00	0.00
ATOM	2	H1	ASN	1	4.046	0.840	-0.000	1.00	0.00
ATOM	3	H2	ASN	1	2.823	1.500	-0.875	1.00	0.00
ATOM	4	H3	ASN	1	2.823	1.500	0.875	1.00	0.00
ATOM	5	CA	ASN	1	3.970	2.846	-0.000	1.00	0.00
ATOM	6	HA	ASN	1	3.672	3.400	-0.890	1.00	0.00
ATOM	7	CB	ASN	1	3.577	3.654	1.232	1.00	0.00
ATOM	8	2HB	ASN	1	2.497	3.801	1.241	1.00	0.00
ATOM	9	3HB	ASN	1	3.877	3.116	2.131	1.00	0.00
ATOM	10	CG	ASN	1	4.254	5.017	1.232	1.00	0.00
ATOM	11	OD1	ASN	1	5.005	5.340	0.315	1.00	0.00
ATOM	12	ND2	ASN	1	3.985	5.818	2.266	1.00	0.00
ATOM	13	1HD2	ASN	1	4.408	6.734	2.315		

Creating the prmtop and inpcrd files.

> saveamberparm TC5b TC5b.prmtop TC5b.inpcrd

Minimizing, Heating, Production

$AMBERHOME/bin/sander -O -i heat.in -o heat.out -p TC5b.prmtop -c TC5b.inpcrd -r heat.rst

heat.in

```
Heating of TC5b
&cntrl
imin=1,
maxcyc=1000,
ncyc=500,
cut=999.,
rgbmax=999.,
igb=1,
ntb= ,
ntpr=100 /
```

What does the command do?

imin=1 means minimization!
imin=0 means not minimization!

This is actually minimization!

3: Minimizing

$AMBERHOME/bin/sander -O -i min.in -o min.out -p TC5b.prmtop -c TC5b.inpcrd -r min.rst

min.in

```
Minimizing of TC5b
&cntrl
imin=1,
maxcyc=1000,
ncyc=500,
cut=999.,
rgbmax=999.,
igb=1.
ntb=
ntpr=100 /
```

ntb = 1: periodic box condition (PBC)
ntb = 0: no PBC

What should be the value of ntb?

We have not solvated the peptide into water!
So there are no water molecules in the system!

Because no water, it is unnecessary to use PBC.

cut = 999 too large, actually no cut-off at all

Because of no water, the system is small,
full computation is not a problem,
no need of cut off.

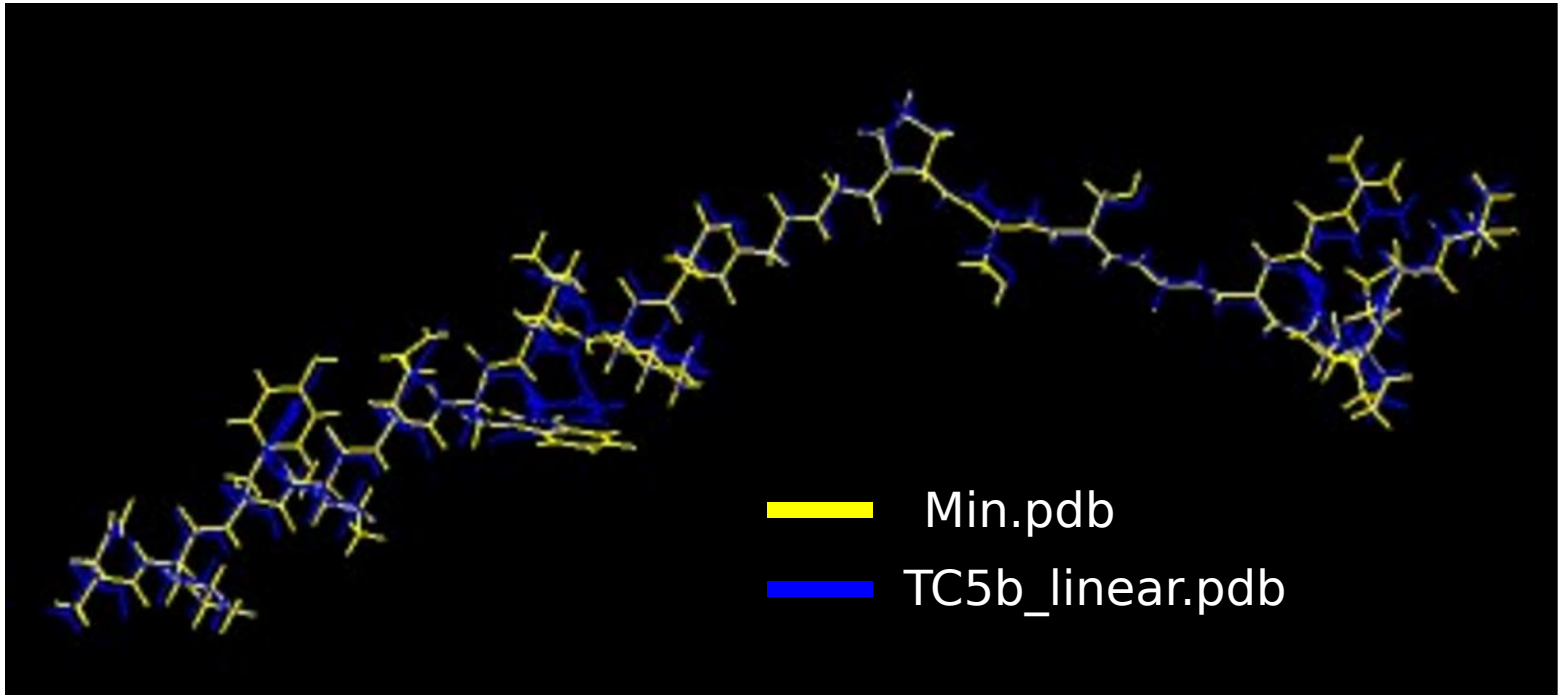
Indeed, it takes about 3.5 seconds on 16 cpus of a 1.3GHz SGI Al

to see the structure after the minimization, which file should we exam

$AMBERHOME/bin/sander -O -i min.in -o min.out -p TC5b.prmtop -c TC5b.inpcrd -r min.rst

cannot be viewed directly, a conversion should be perfromed

$AMBERHOME/bin/ambpdb -p TC5b.prmtop < min.rst
> min.pdb



Heating

et the temperature *T* gradually increase from 0 K to 325 K .

Initial crystal structure is static, which corresponds to 0K

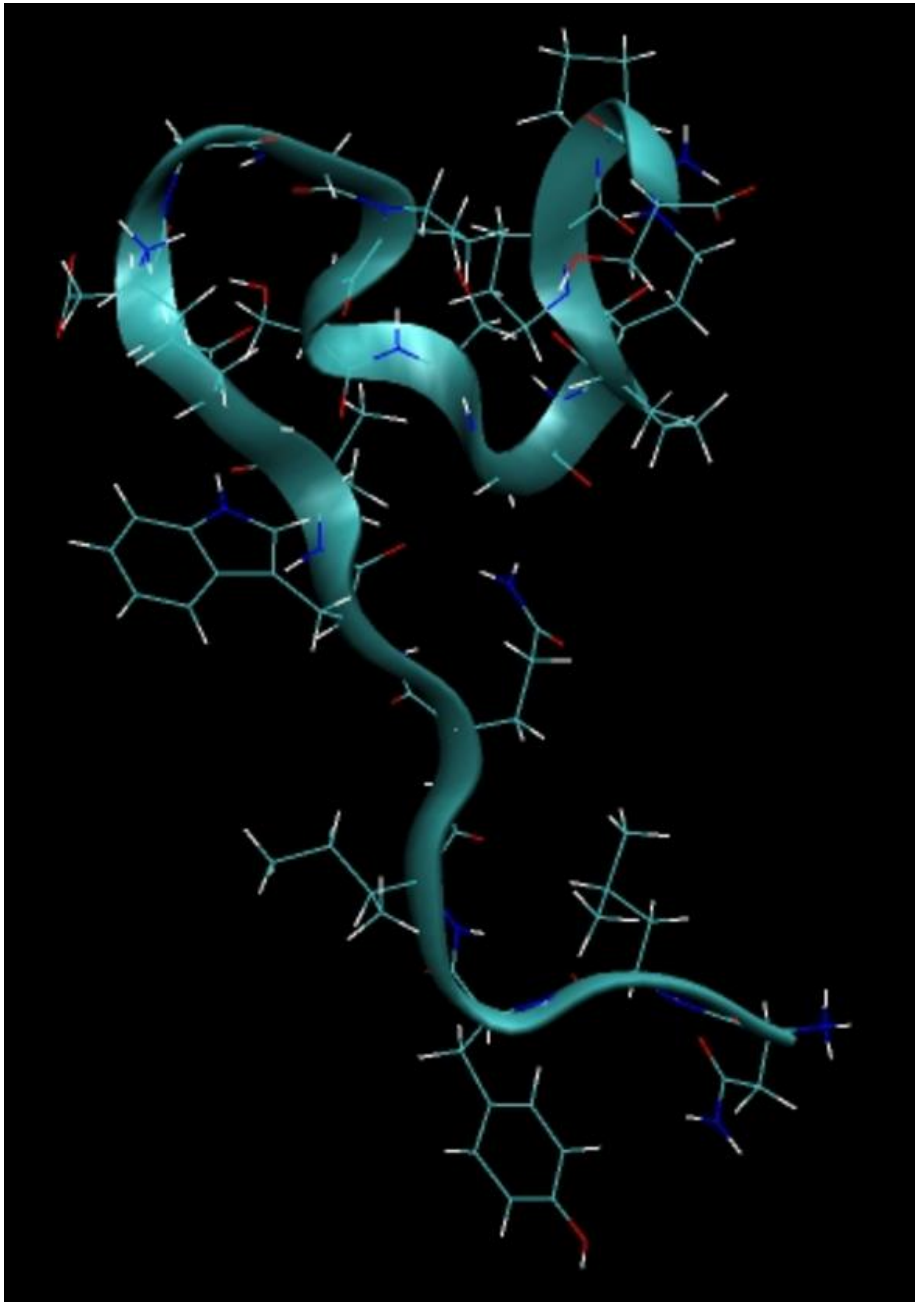
Many target temperatures

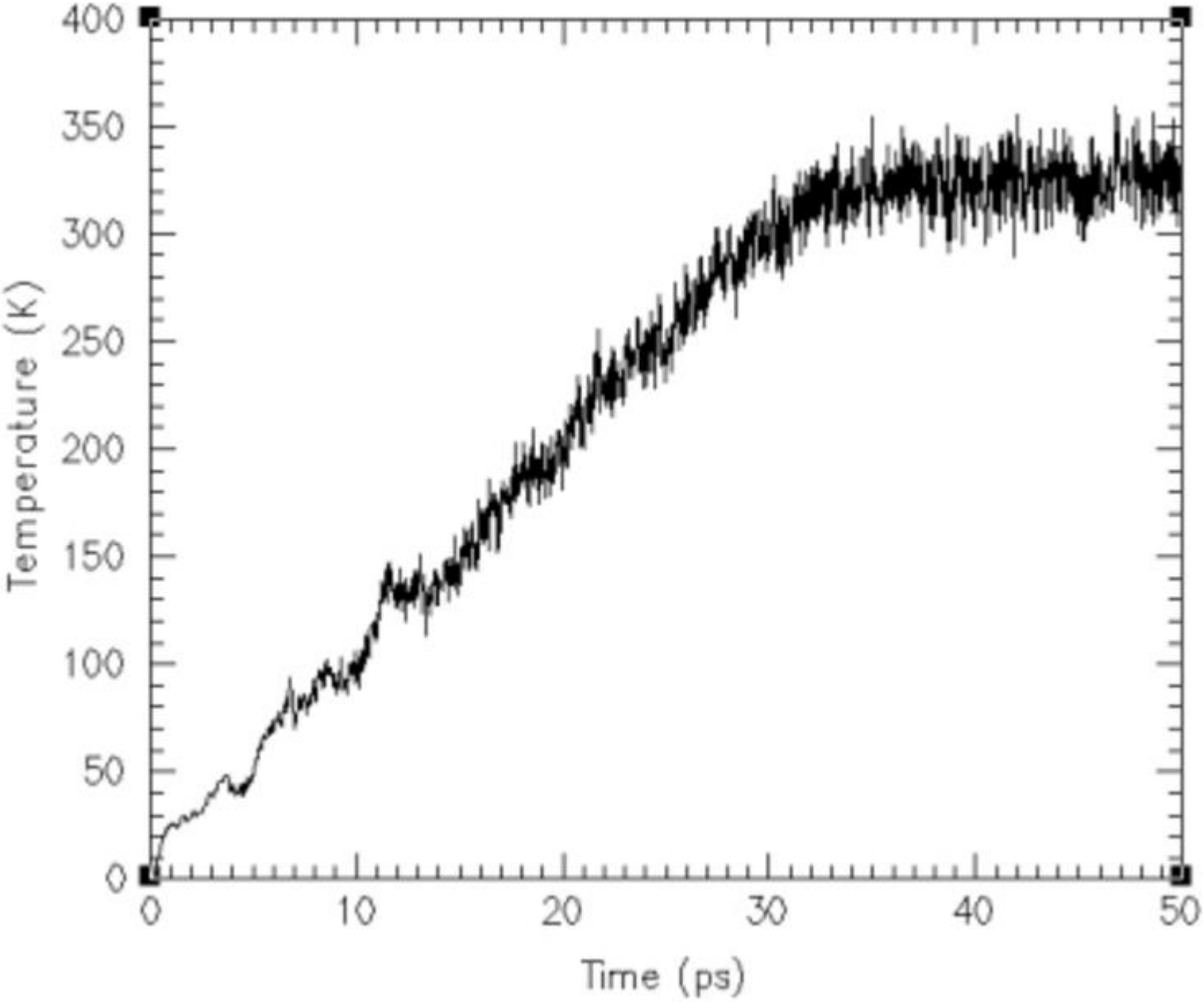
0 50 K 100K 150 K ... 325 K

: Heating the system up.

A Portable Batch System (PBS) script

```
#PBS -l ncpus=16
#PBS -l walltime=500:00:00
#PBS -l cput=2000:00:00
#PBS -j oe
setenv AMBERHOME /usr/people/rcw/amber9
cd ~rcw/initial_heating
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat1.in -p TC5b.prmtop -c min1.rst -r heat1.rst -o heat1.out -x heat1.mdcrd
gzip -9 heat1.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat2.in -p TC5b.prmtop -c heat1.rst -r heat2.rst -o heat2.out -x heat2.mdcrd
gzip -9 heat2.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat3.in -p TC5b.prmtop -c heat2.rst -r heat3.rst -o heat3.out -x heat3.mdcrd
gzip -9 heat3.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat4.in -p TC5b.prmtop -c heat3.rst -r heat4.rst -o heat4.out -x heat4.mdcrd
gzip -9 heat4.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat5.in -p TC5b.prmtop -c heat4.rst -r heat5.rst -o heat5.out -x heat5.mdcrd
gzip -9 heat5.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat6.in -p TC5b.prmtop -c heat5.rst -r heat6.rst -o heat6.out -x heat6.mdcrd
gzip -9 heat6.mdcrd
mpirun -np 16 $AMBERHOME/bin/sander -O -i heat7.in -p TC5b.prmtop -c heat6.rst -r heat7.rst -o heat7.out -x heat7.mdcrd
gzip -9 heat7.mdcrd
echo "DONE"
```



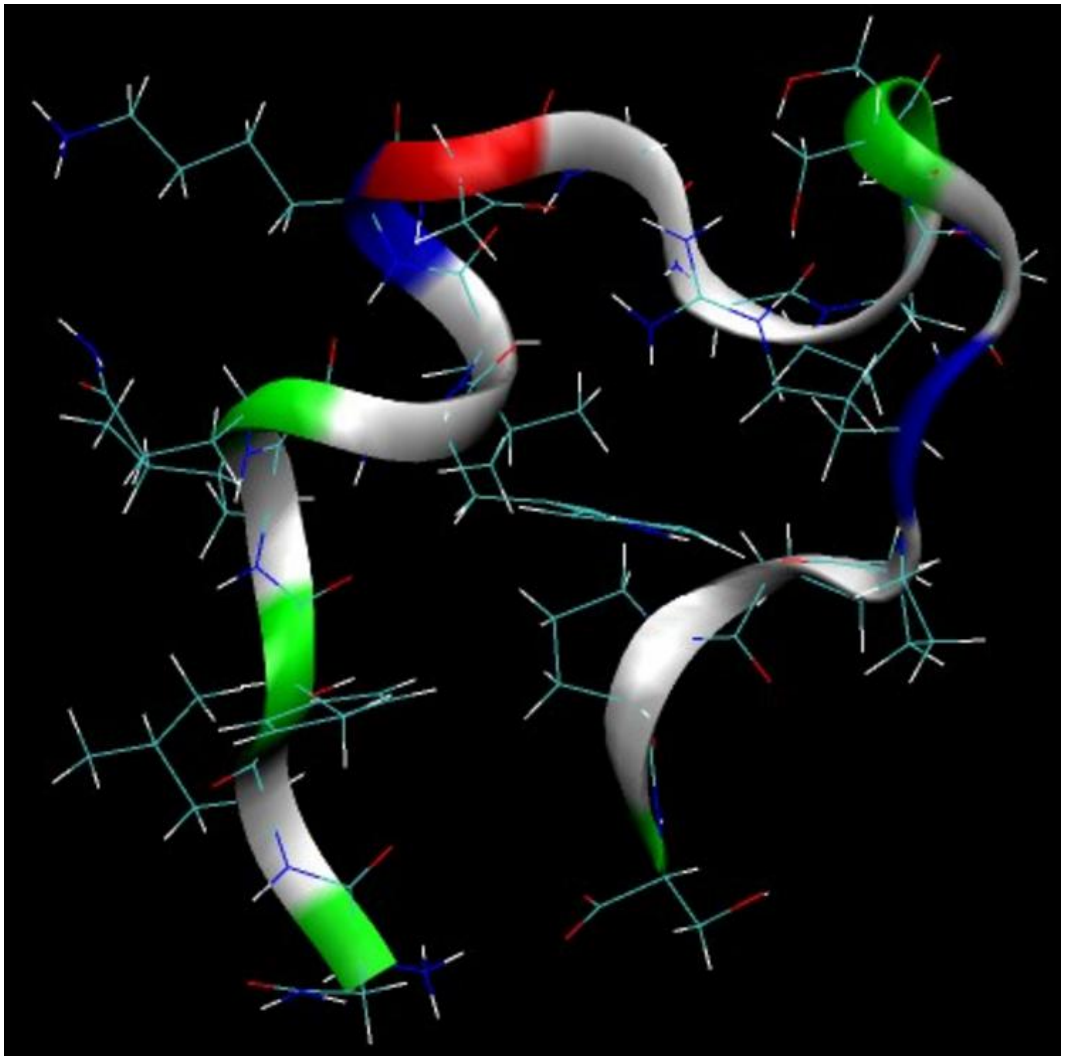


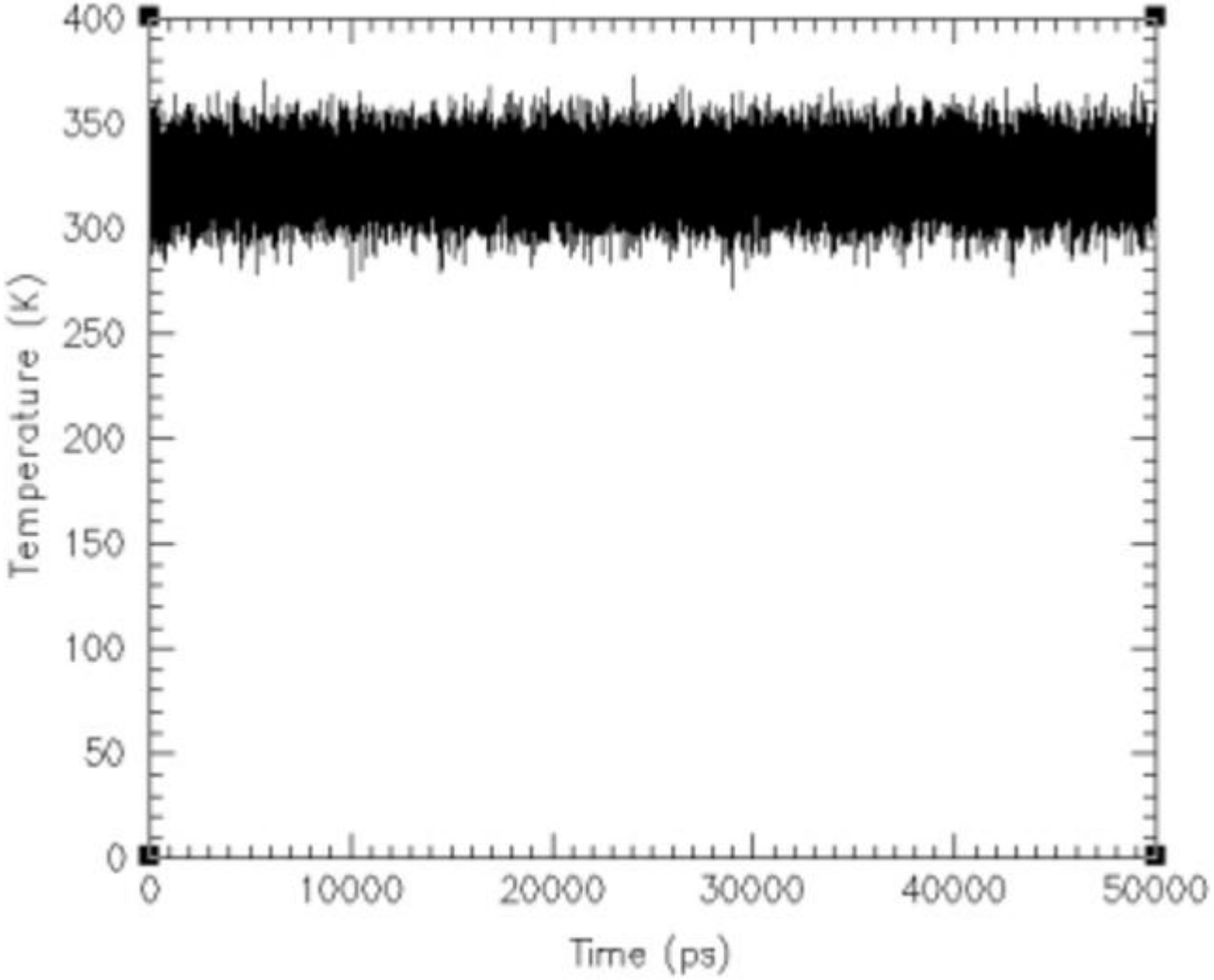
Temperature close up of heating stage

5: Production MD

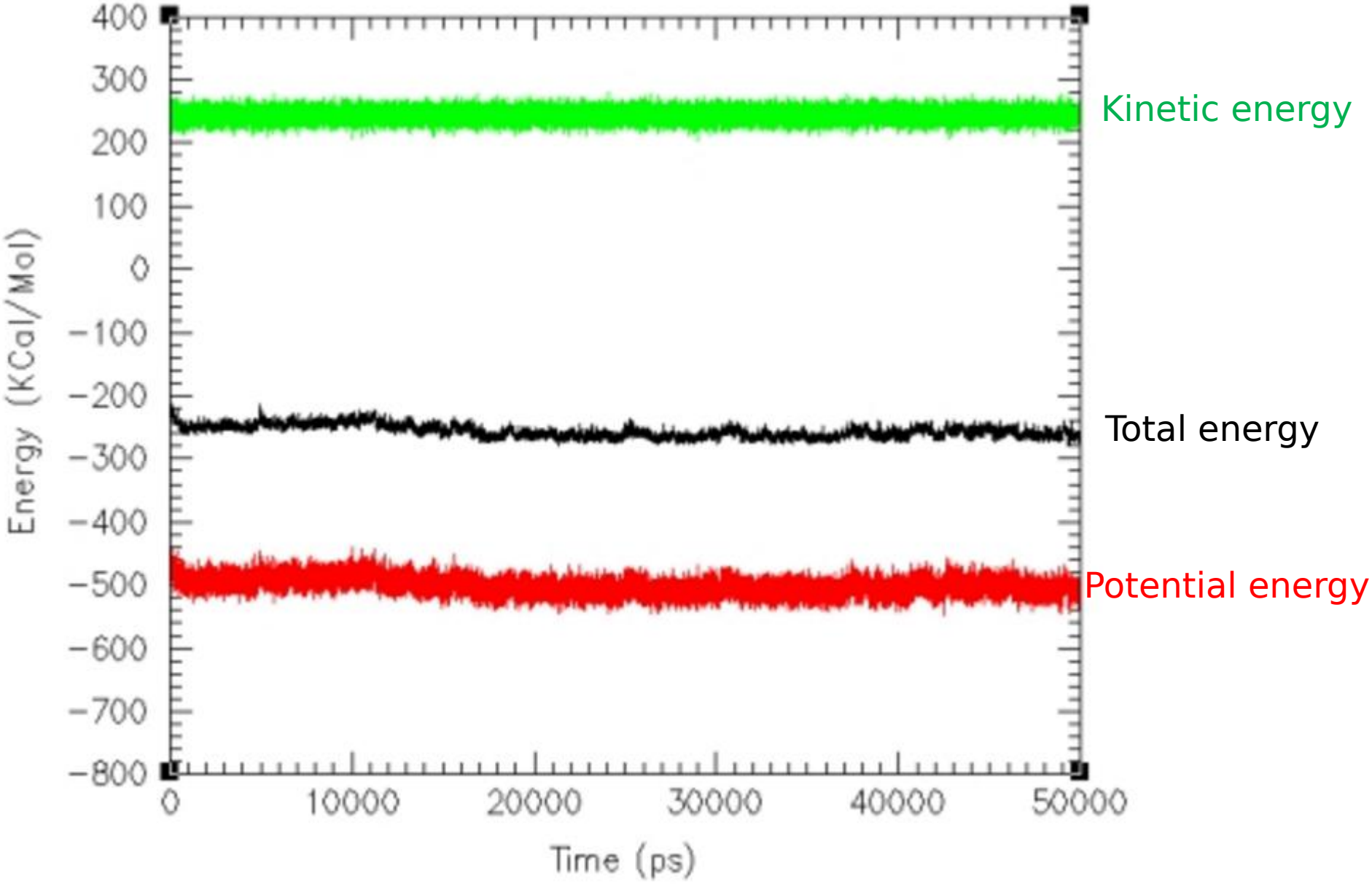
Simulating 50 ns protein folding process

It takes a total of about 27 hours to run on 16 cpus of a 1.3GHz SGI Altix.





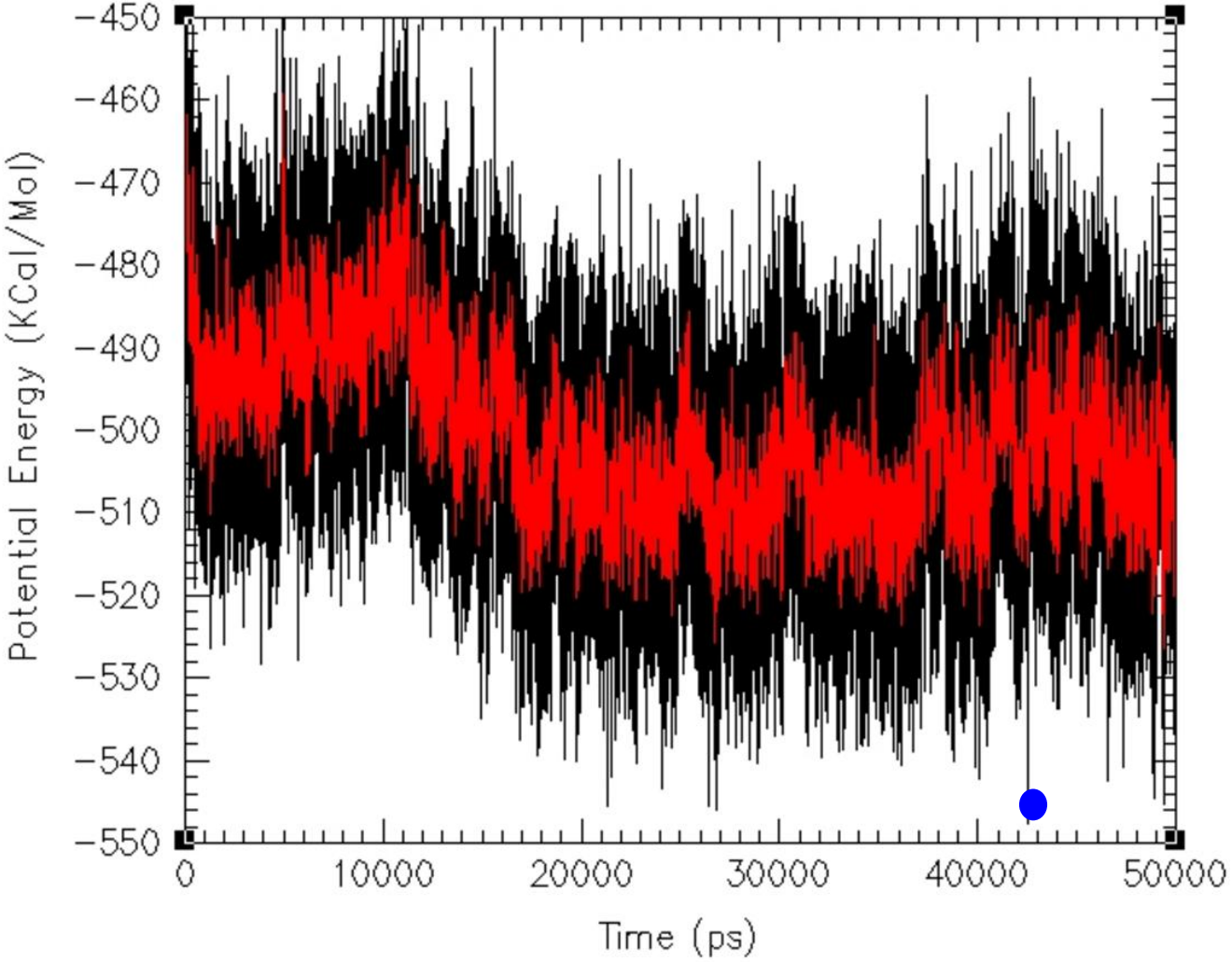
Temperature



he reduction of **total energy** and **potential energy** at about 12000 ps. The m

e structure switches to a more stable one at about 12000 ps.

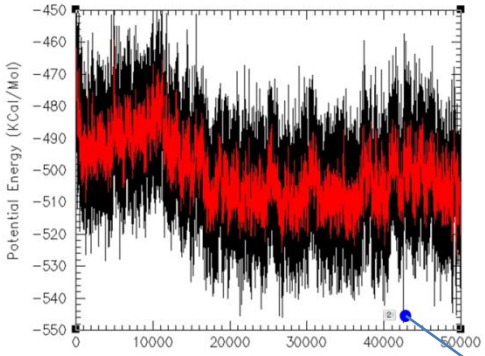
loser view of potential energy only

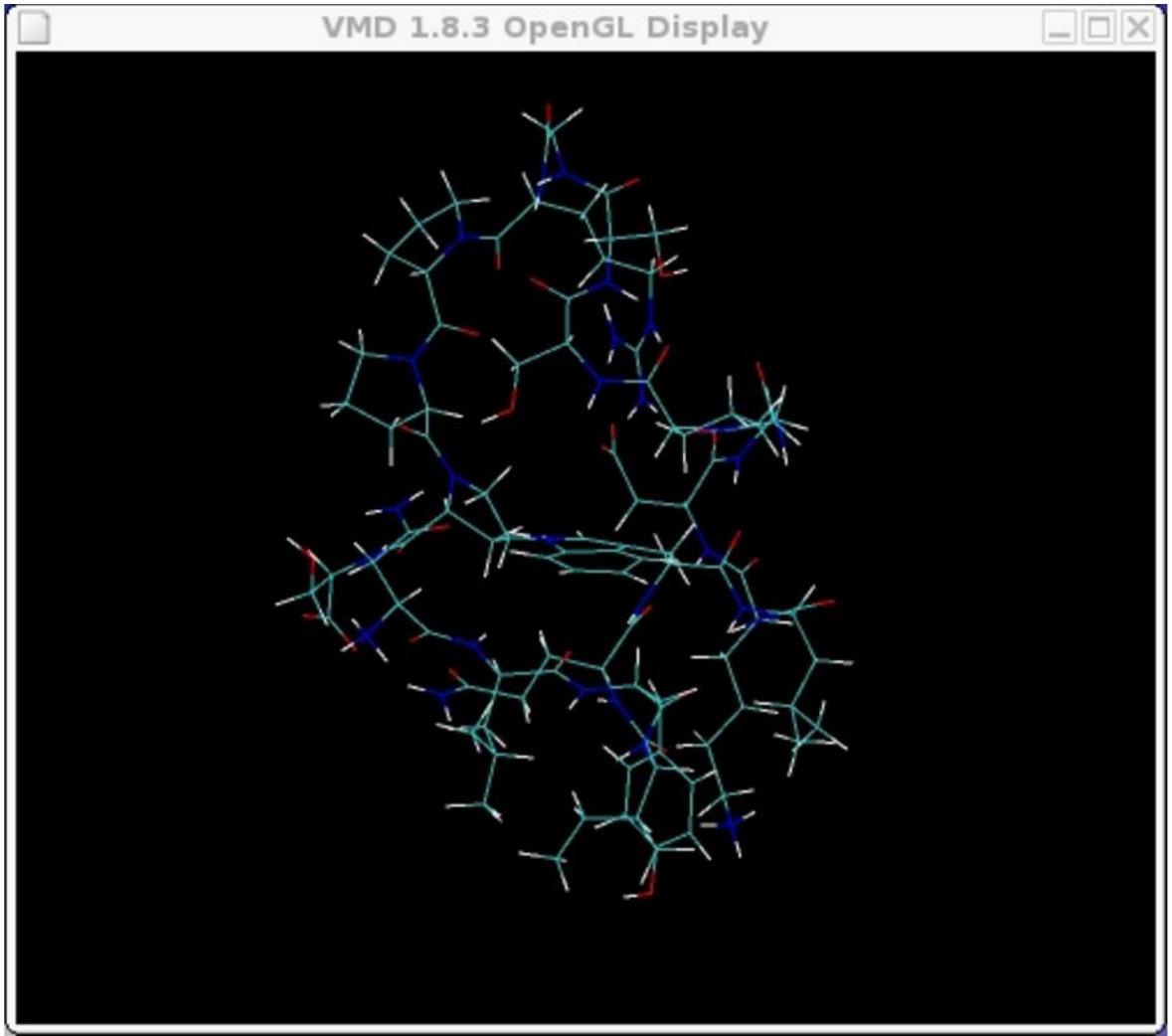


Black = original
Red = running average over 10ps

What is the lowest potential energy? At about which time point?

re interested in the corresponding structure

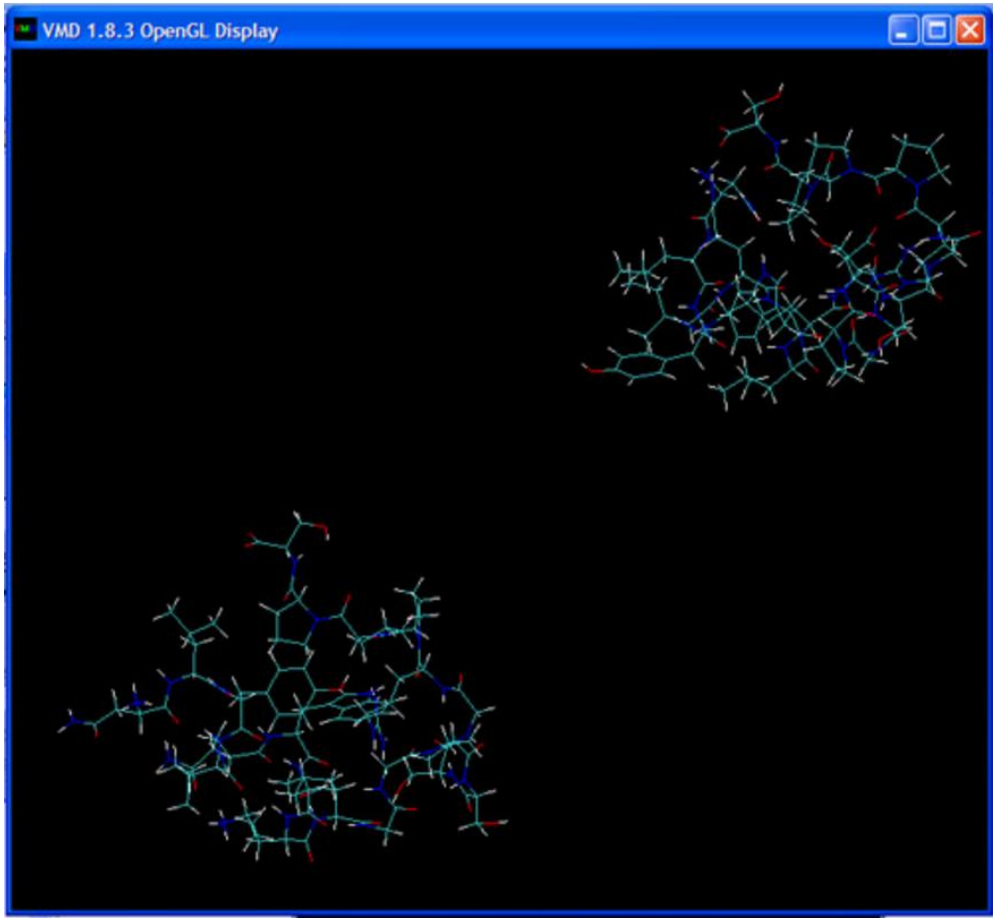




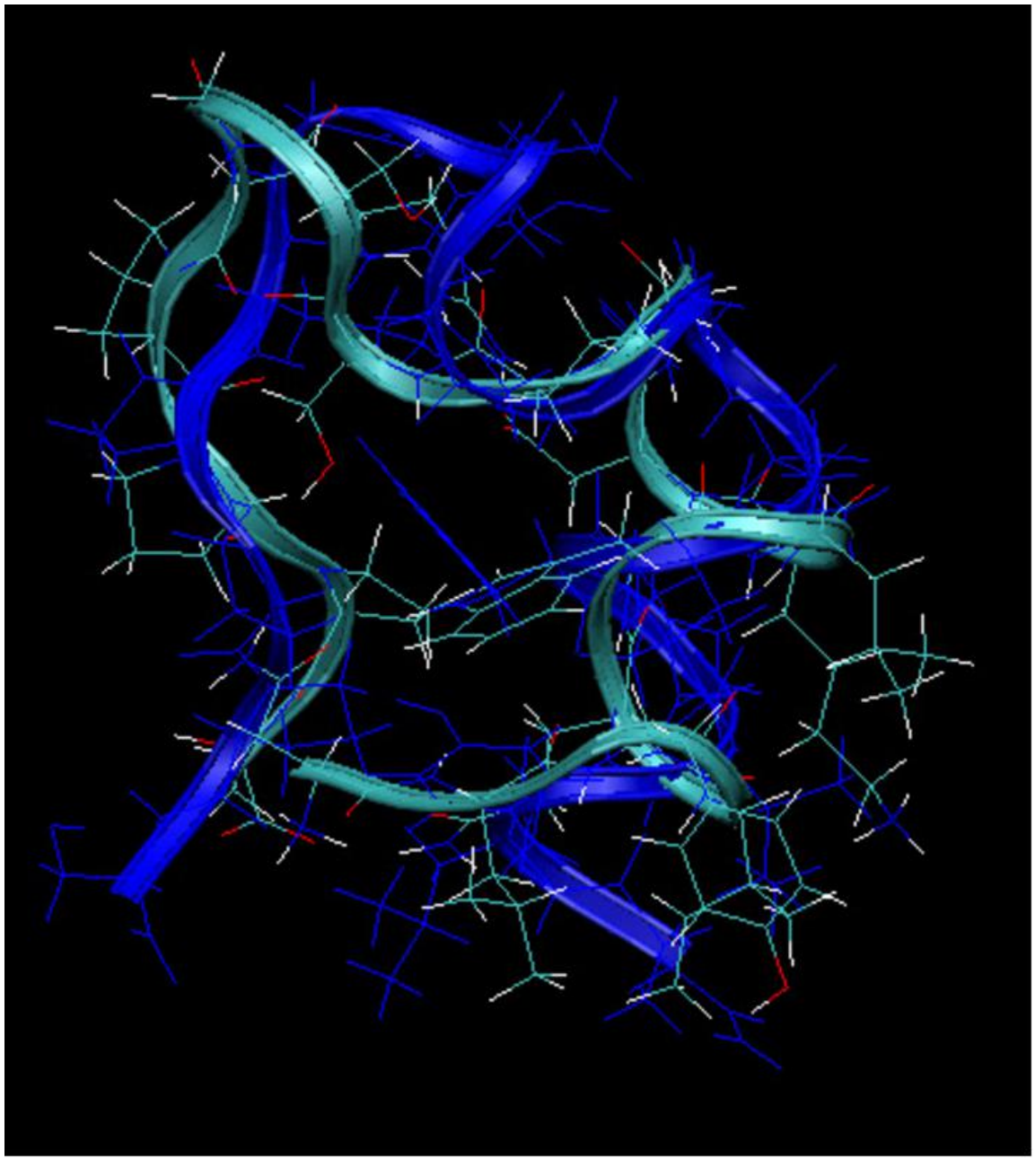
mulation had been done **before** the real NMR structure (1L2Y.pdb) was discover

ow eager to see whether the simulated structure is close to the real structure

We load the structures into VMD environment



ment: making the two structures as close as possible (minimize RMSD)



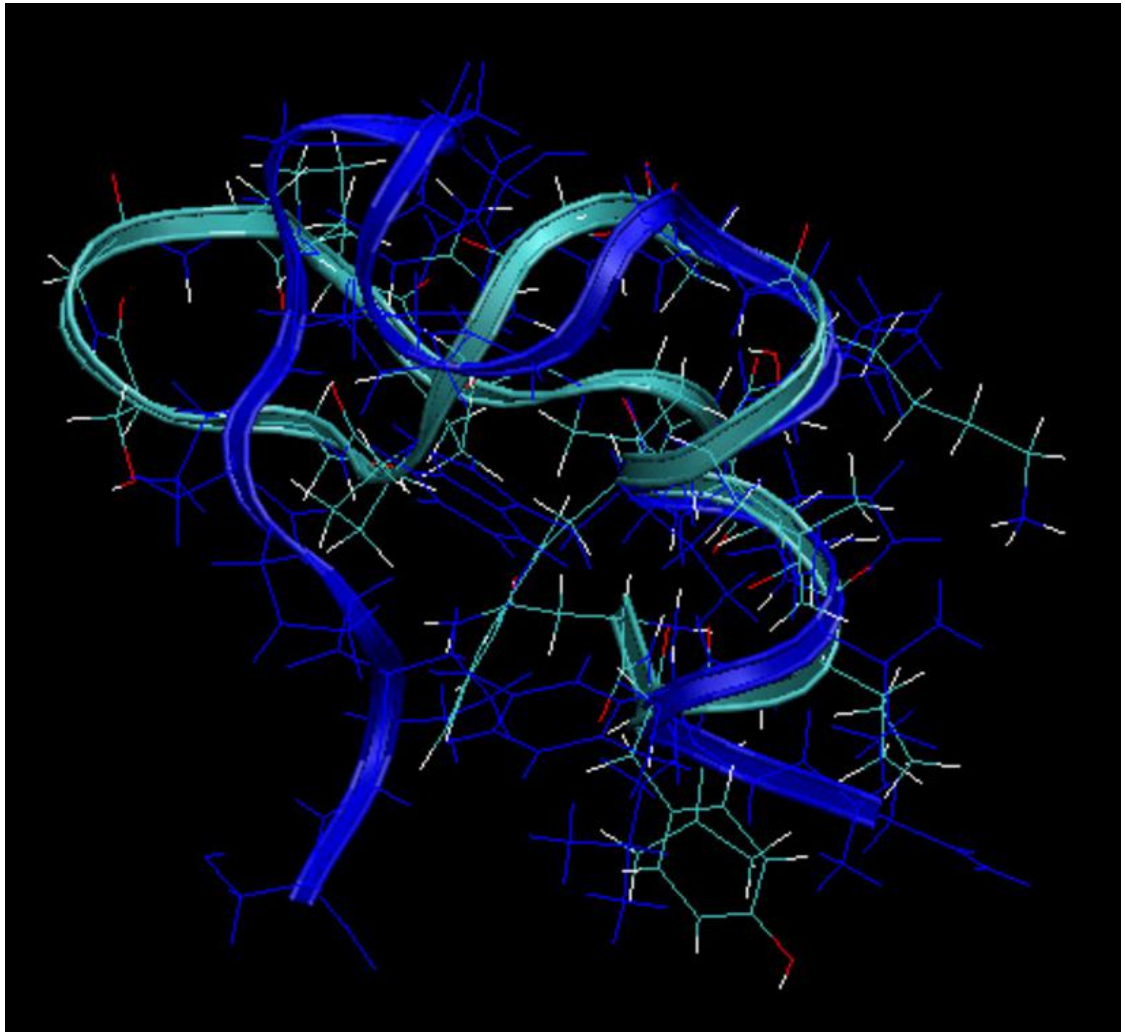
RMSD = 3.72

Not a good fit

The 50ns simulation
may be not long enough

But we want to know if
The major structure,
alpha-helix, is folded well

ial alignment: just align the alpha-helix (residues 3 to 9)

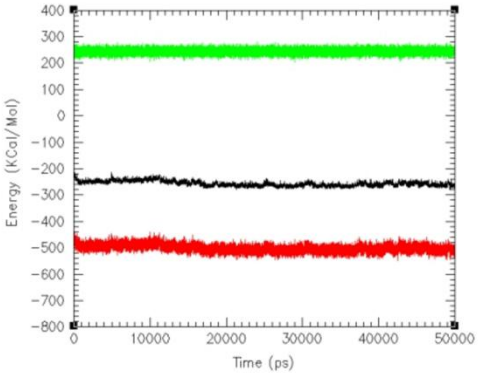


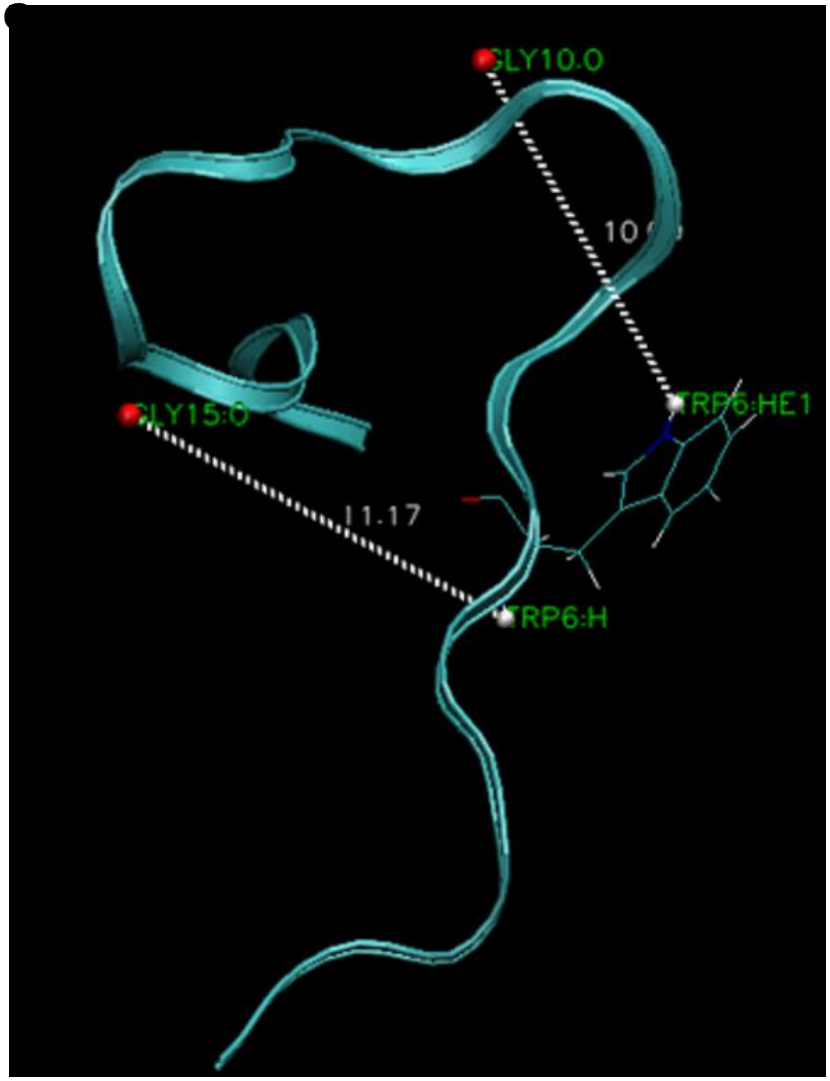
RMSD = 0.7

good fit

So the simulation is meaningfu

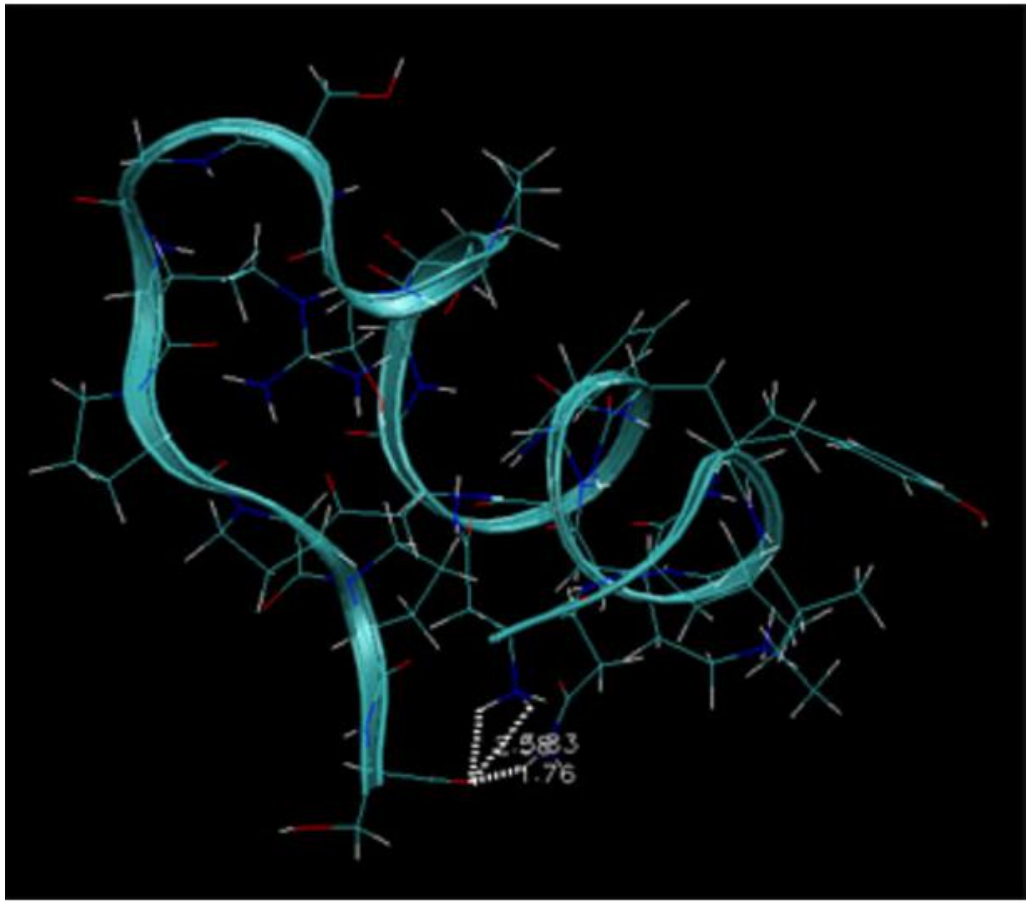
y, it can be explained why a sudden energy reduction
rom one stable structure to another one

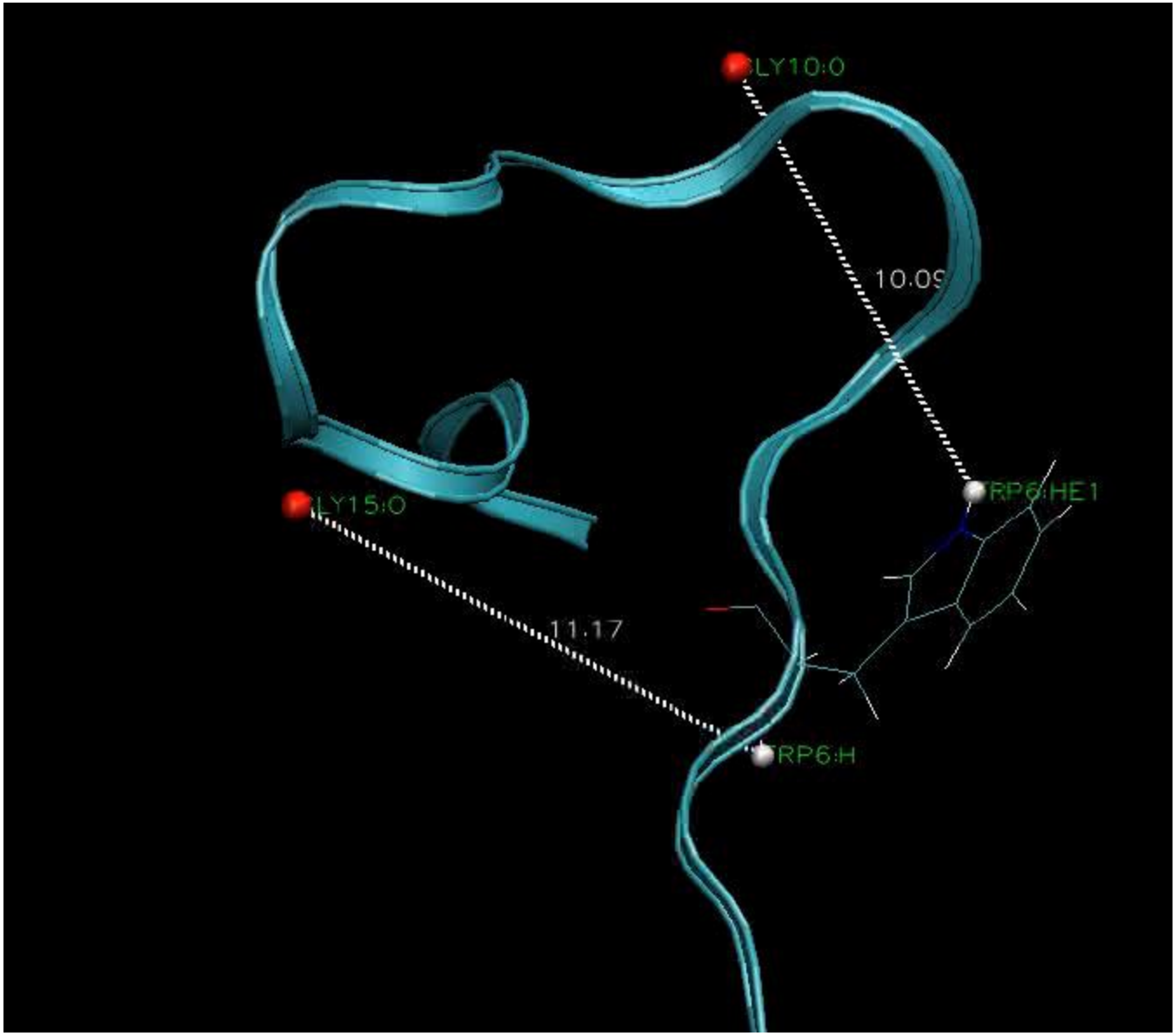




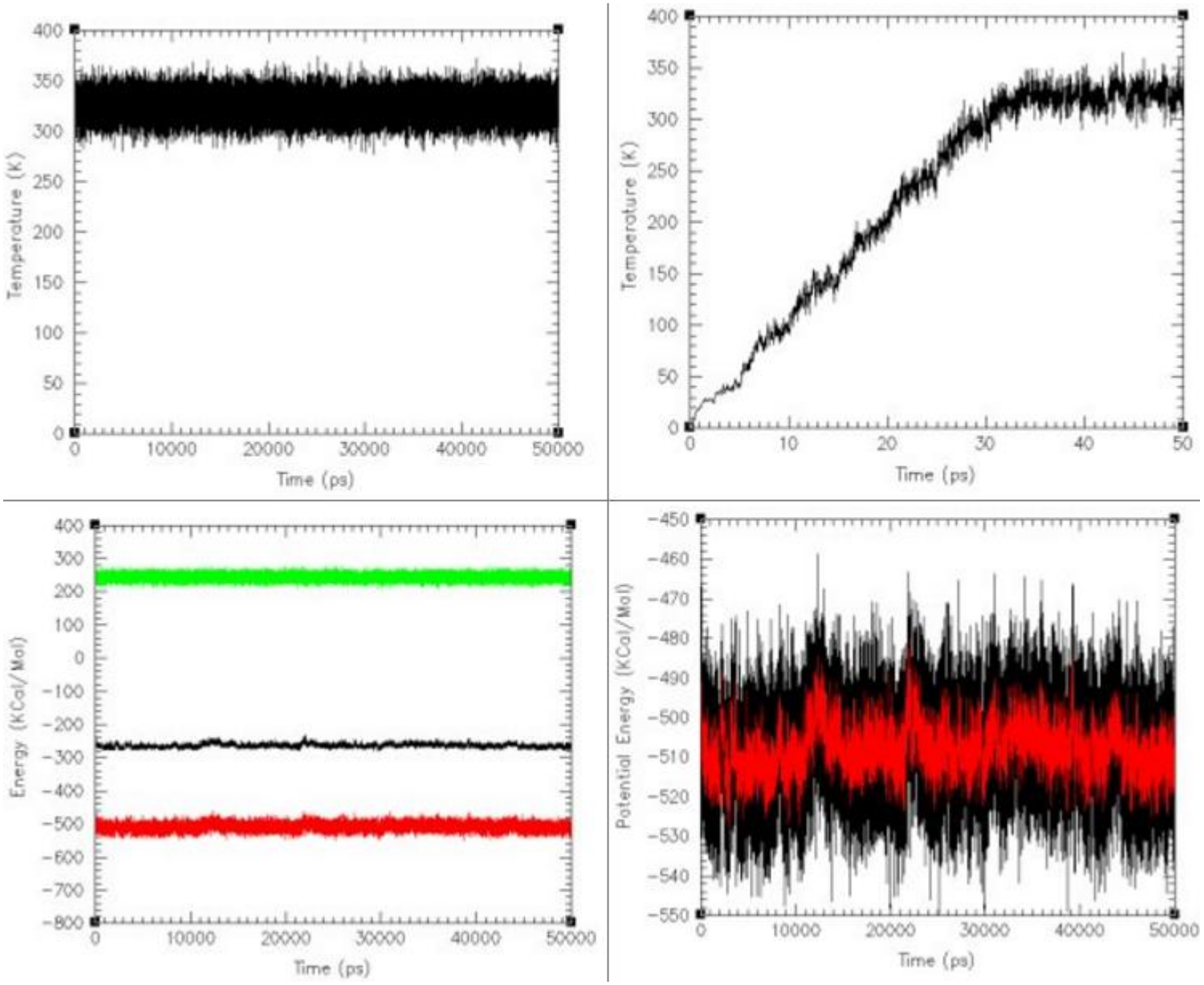
ound that during the first 15ns, two hydrogen bonds formed that
the structure trapped at a higher minimal energy, preventing formation of alph

hydrogen bonds break at 15ns or so, which allows the formation of alpha-helix

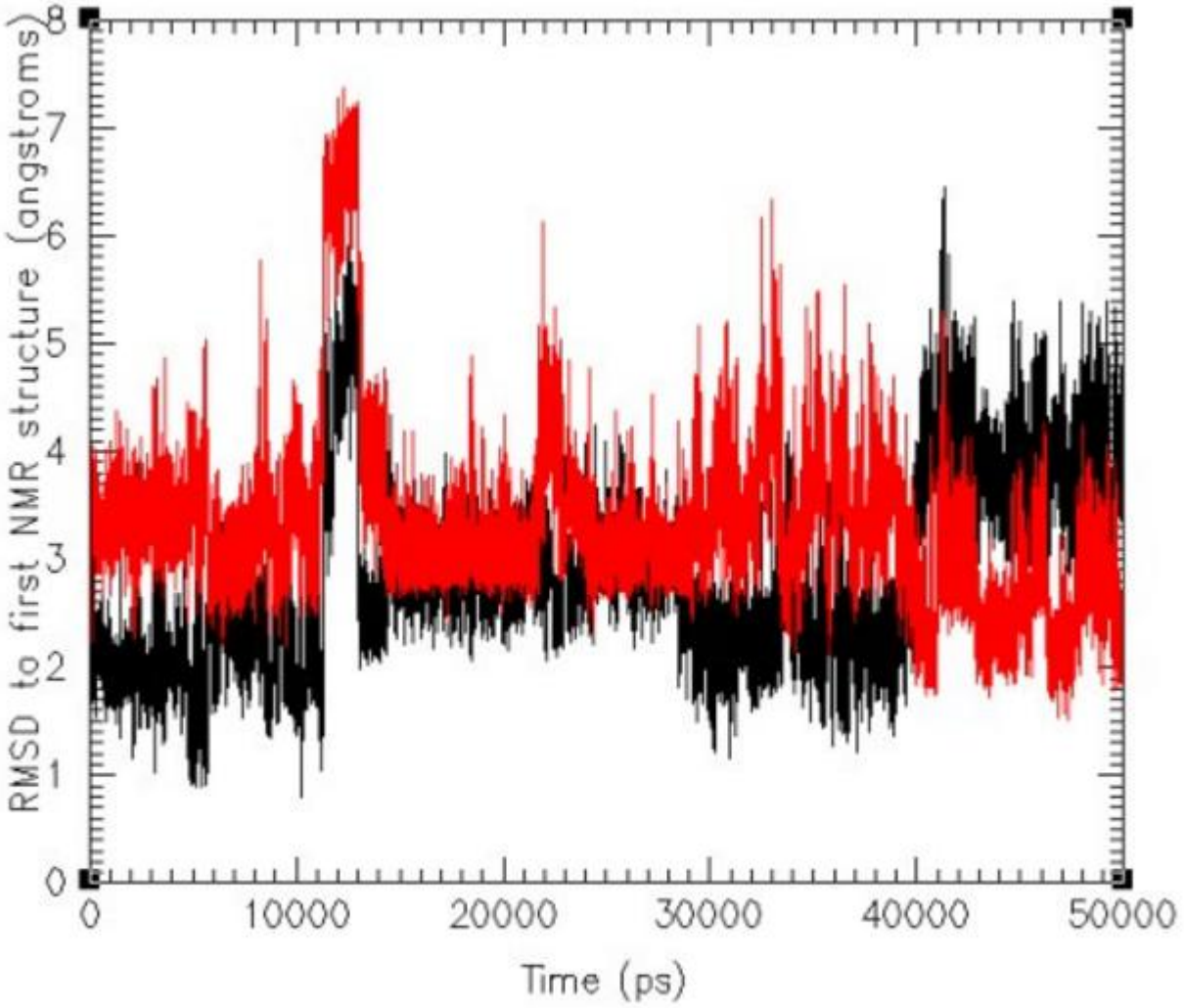



LY10:O
10.09
RP6:HE1
LY15:O
11.17
RP6:H

arting from the NMR structure



RMSD



Plot showing backbone RMSD to 1st NMR structure (Black) and lowest energy structure from original folding calculation (red).