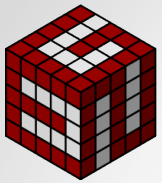


GENESIS

Generalized-ensemble simulation system

理化学研究所 計算科学研究機構
粒子系生物物理研究チーム
Computational Biophysics Research Team
RIKEN Advanced Institute for Computational Science

GENESIS Tutorial 4
String Method

Lecturer: Yasuhiro Matsunaga
TA: Jaewoon Jung, Chigusa Kobayashi,
Koichi Tamura, Motoshi Kamiya

2017/03/01

Contents

I. Most probable pathway with the string method

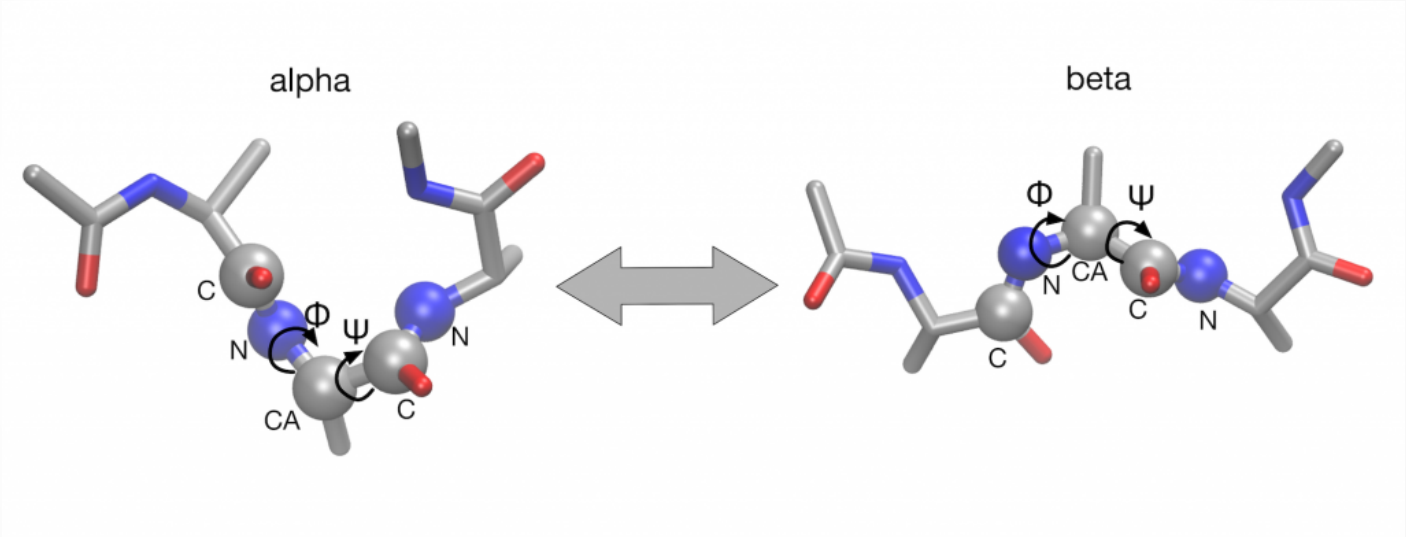
- ✓ Equilibration along an initial pathway
- ✓ String method
- ✓ Visualization of the pathways

II. Free energy profile along the pathway

- ✓ Equilibration along the pathway
- ✓ Umbrella sampling along the pathway
- ✓ Analysis of free energy profile

Target system

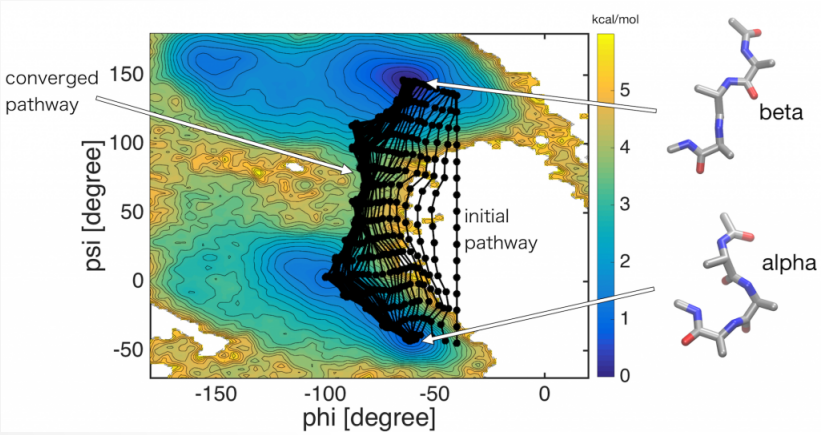
Conformational change of alanine-tripeptide (same as REMD's)



How can we characterize this conformational change by MD simulation?

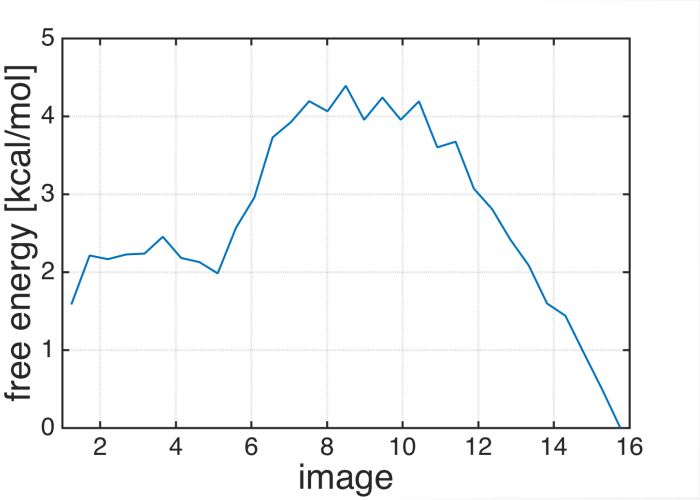
Two main results obtained by this tutorial

Pathways (=most probable pathway) in the dihedral angle space



Gets insights into causalities

Free energy profile along the pathway

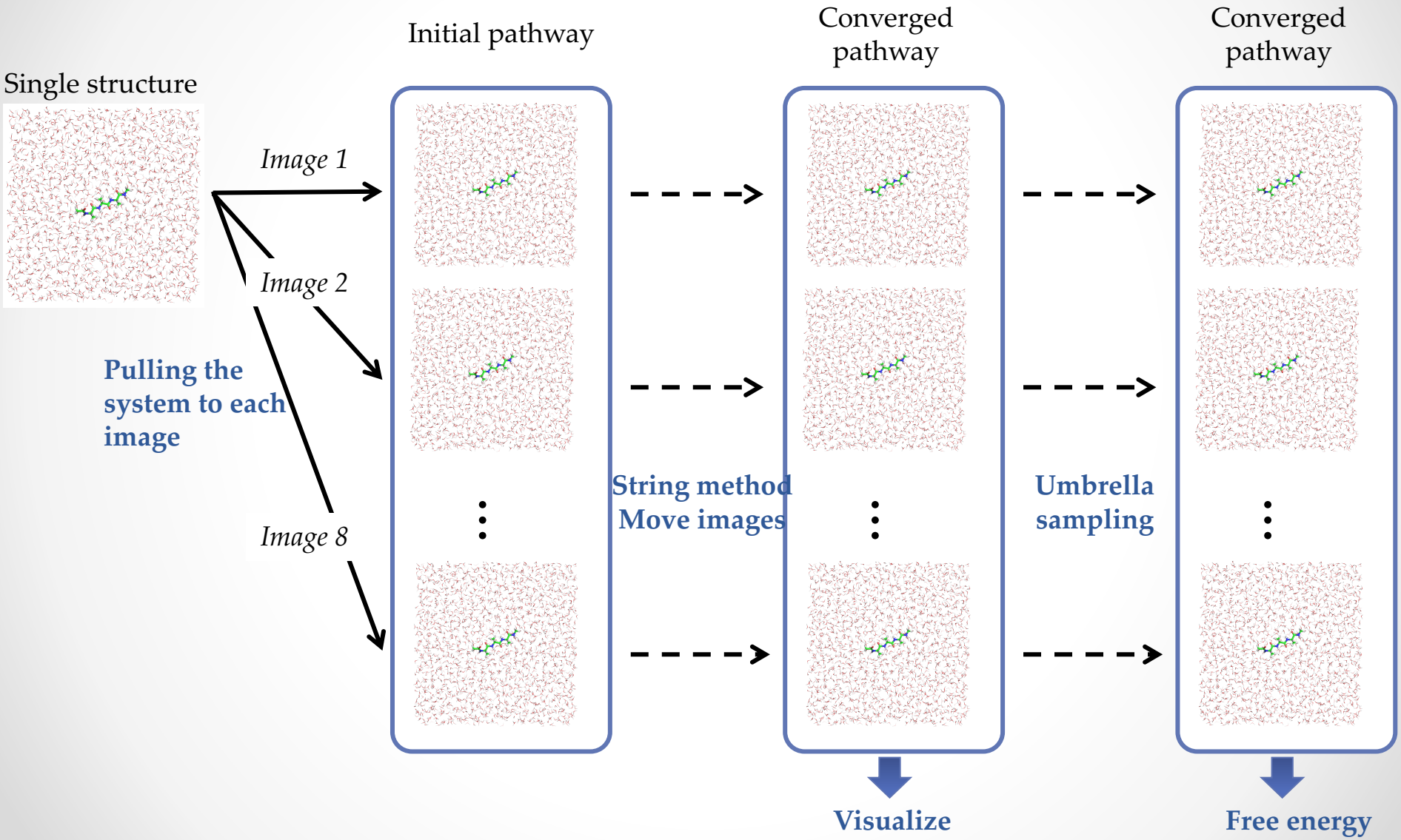


Identifies transition state

Basic usage

```
# enter the login node
$ ssh –Y -l userXX XXXXXX

# enter a computational node
$ ssh –Y XXX

# change the current directory to Tutorial_4
$ cd ~/Tutorial_4/
```

Flow of string method simulations



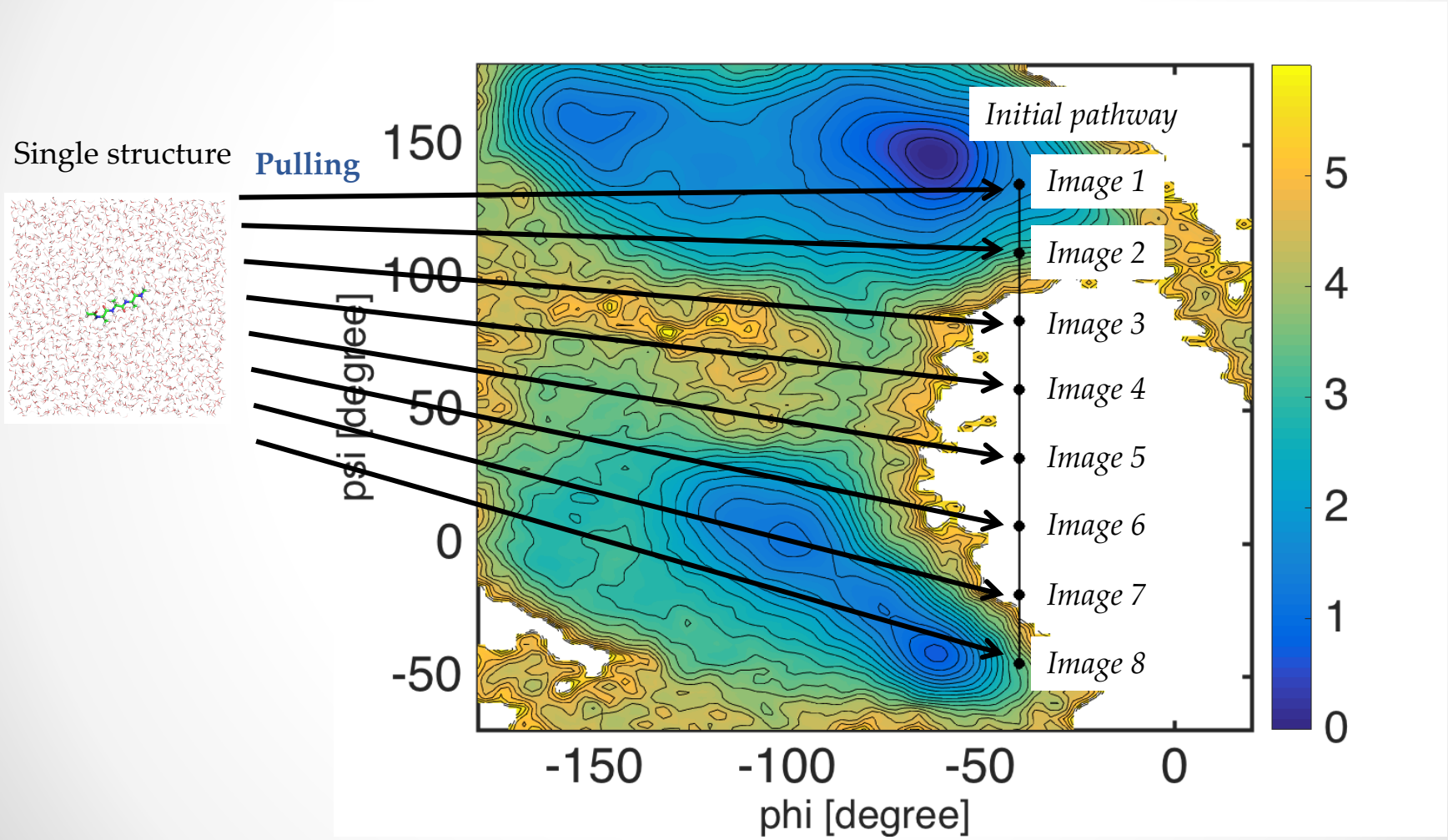
Preparation

```
$ cd ~/Tutorial_4/


$ ls
0_setup       2_pre-equilibration  4_string_method  6_umbrella
1_minimize    3_equilibration      5_visualize      7_analysis
```

We will skip step 0 to 3 because these are same as REMD's

We will change the directory from 3_initial_path to 6_visualization step by step

Step3. Pulling the system toward images along the initial pathway



Control file for GENESIS spdyn

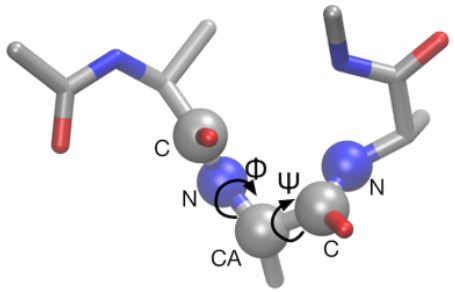
The RPATH module in GENESIS can simulate replicas **without string method** performed (rpath_period = 0).

```
[INPUT]
rstfile = ../2_minimize_pre-equi/step2.4.rst

[OUTPUT]
logfile     =  {}.log
dcdfile     =  {}.dcd
rstfile     =  {}.rst
rpathfile   =  {}.rpath

[ENERGY]
forcefield          = CHARMM
electrostatic       = PME
switchdist          = 10.0
cutoffdist          = 12.0
pairlistdist        = 13.5

[DYNAMICS]
integrator          = LEAP
nsteps              = 2000
timestep            = 0.002
eneout_period       = 1000
crdout_period       = 1000
rstout_period       = 2000

[CONSTRAINTS]
rigid_bond          = YES      # use SHAKE/SETTLE
```

Replica index is automatically inserted into {} of the output filename.

```
[ENSEMBLE]
ensemble            = NPT
tpcontrol           = LANGEVIN
temperature         = 300.00
pressure            = 1.0

[RPATH]
nreplica        = 8
rpath_period    = 0
rest_function = 1 2

[SELECTION]
group1          = atomindex:15
group2          = atomindex:17
group3          = atomindex:19
group4          = atomindex:25
group5          = atomindex:27

[RESTRAINTS]
nfunctions      = 2

function1       = DIHED
constant1       = 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0
reference1      = -40 -40 -40 -40 -40 -40 -40 -40
select_index1 = 1 2 3 4   # PHI

function2       = DIHED
constant2       = 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0
reference2      = -45.0 -19.3 6.4 32.1 57.9 83.6 109.3 135.0
select_index2 = 2 3 4 5   # PSI
```

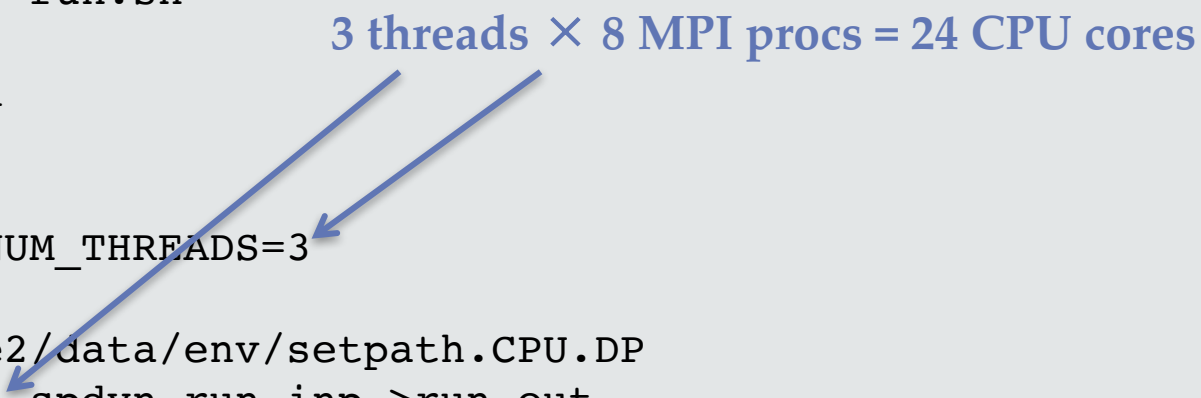
Simulate replicas without string method



Images along the initial pahtway

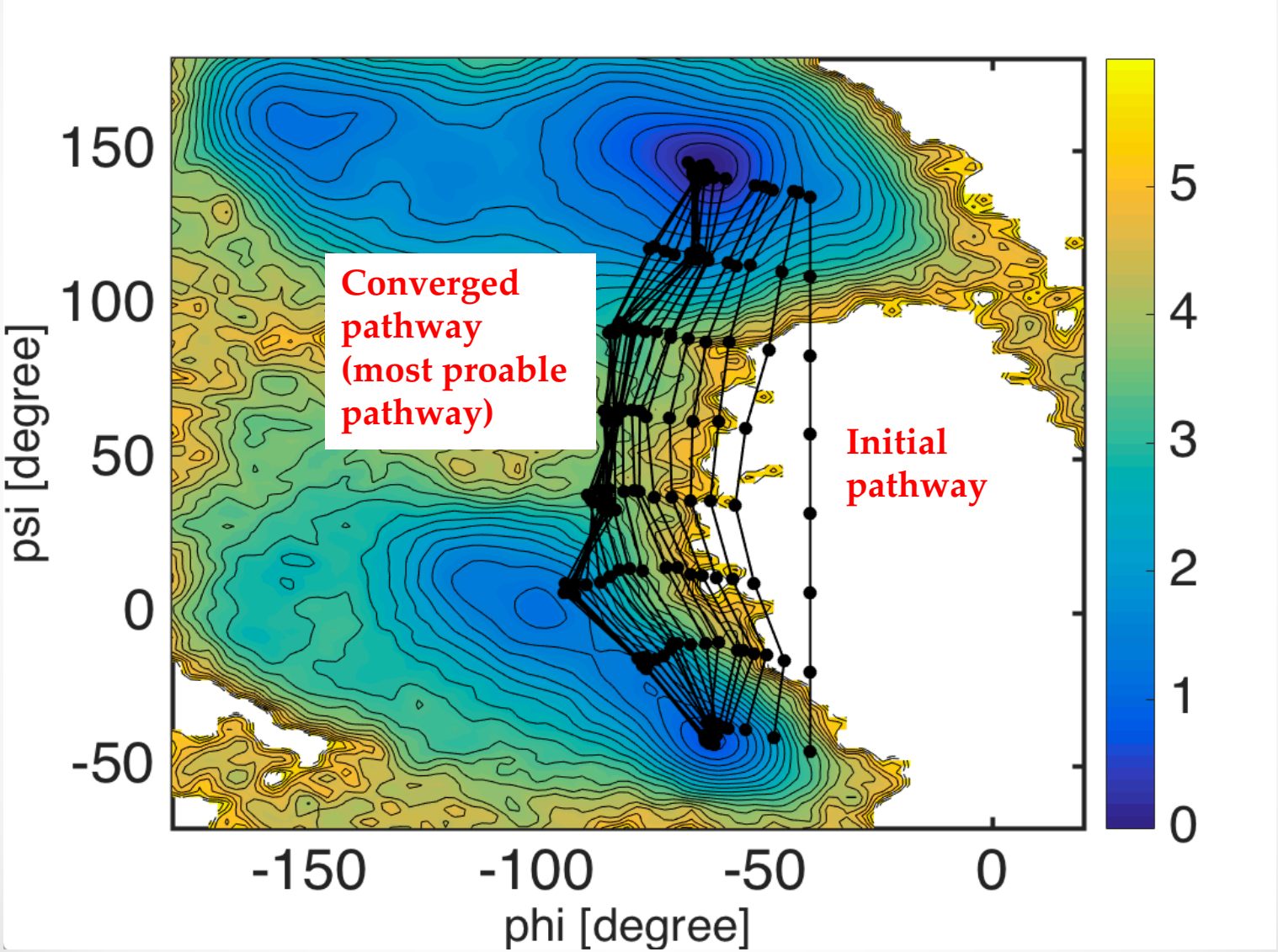
Running a job

```
$ cd 3_initial_path/
$ ls
run.inp      run.sh

$ cat run.sh
#!/bin/bash

export OMP_NUM_THREADS=3

source /home2/data/env/setpath.CPU.DP
mpirun -np 8 spdyn run.inp >run.out

$ ./run.sh
```



This will finish in about 2 minutes

Step4. String method



Control file for GENESIS spdyn

The RPATH module updates images every rpath_period = 500

```
[INPUT]
rstfile = ../3_initial_path/{}.rst

[OUTPUT]
logfile    =  {}.log
dcdfile    =  {}.dcd
rstfile    =  {}.rst
rpathfile  =  {}.rpath

[ENERGY]
forcefield       = CHARMM
electrostatic    = PME
switchdist       = 10.0
cutoffdist       = 12.0
pairlistdist     = 13.5

[DYNAMICS]
integrator       = LEAP
nsteps           = 10000
timestep         = 0.002
eneout_period    = 250
crdout_period    = 250
rstout_period    = 10000

[CONSTRAINTS]
rigid_bond       = YES     # use SHAKE/SETTLE
```

```
[ENSEMBLE]
ensemble          = NPT
tpcontrol         = LANGEVIN
temperature       = 300.00
pressure          = 1.0

[RPATH]
nreplica           = 8
rpath_period       = 500
delta              = 0.035
rest_function      = 1 2

[SELECTION]
group1         = atomindex:15
group2         = atomindex:17
group3         = atomindex:19
group4         = atomindex:25
group5         = atomindex:27

[RESTRAINTS]
nfunctions     = 2

function1      = DIHED
constant1      = 50.0 50.0 50.0 50.0 50.0 50.0 50.0 50.0
reference1     = 0 0 0 0 0 0 0 0
select_index1 = 1 2 3 4   # PHI

function2      = DIHED
constant2      = 50.0 50.0 50.0 50.0 50.0 50.0 50.0 50.0
reference2     = 0 0 0 0 0 0 0 0
select_index2 = 2 3 4 5   # PSI
```

Stronger force constant

Images are automatically read from rst files

Running a job

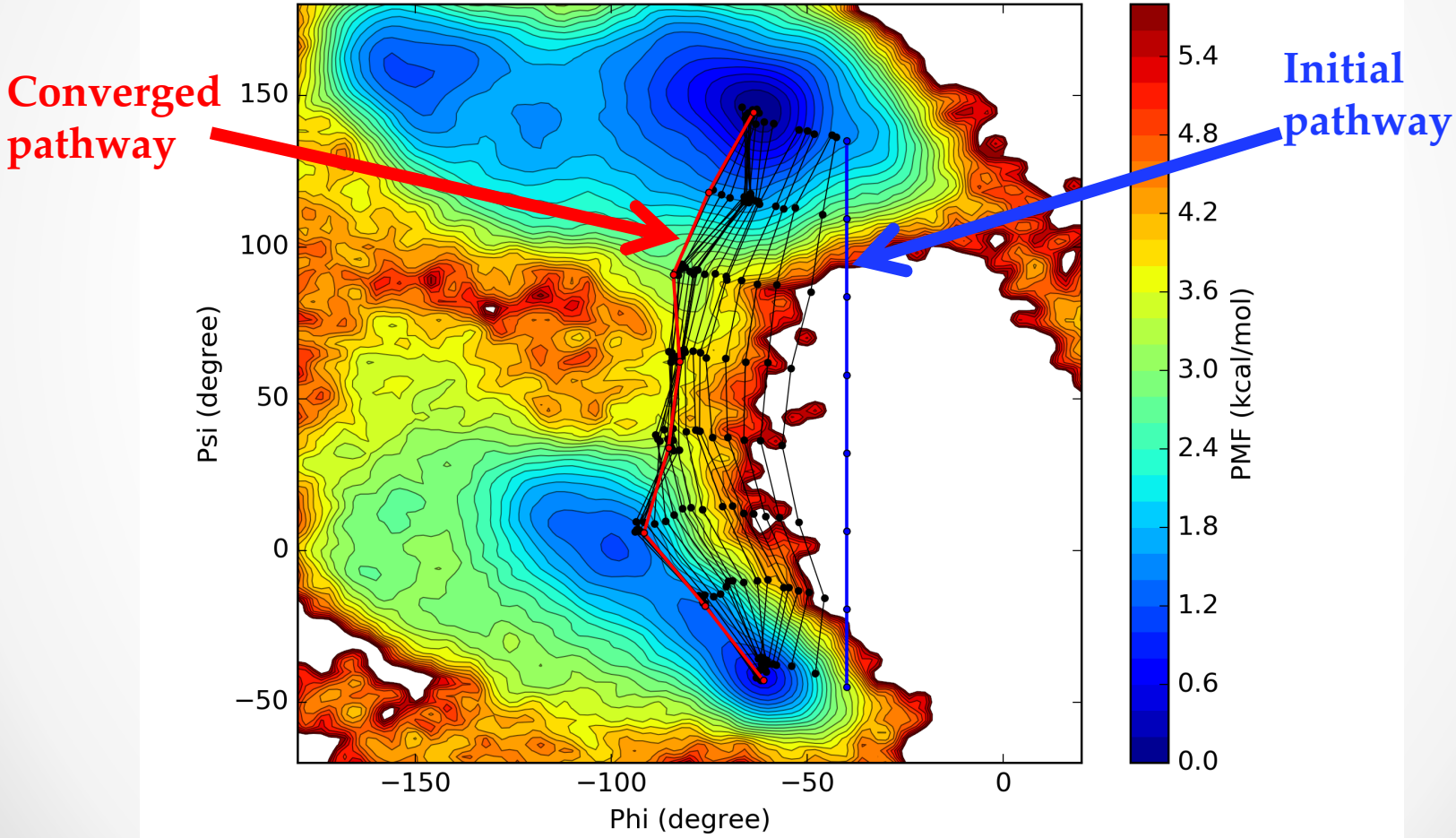
```
$ cd 4_string_method/
$ ls
run.inp       run.sh

$ cat run.sh
#!/bin/bash

export OMP_NUM_THREADS=3

source /home2/data/env/setpath.CPU.DP
mpirun -np 8 spdyn run.inp >run.out

$ ./run.sh
```

This will finish in about **13 minutes**

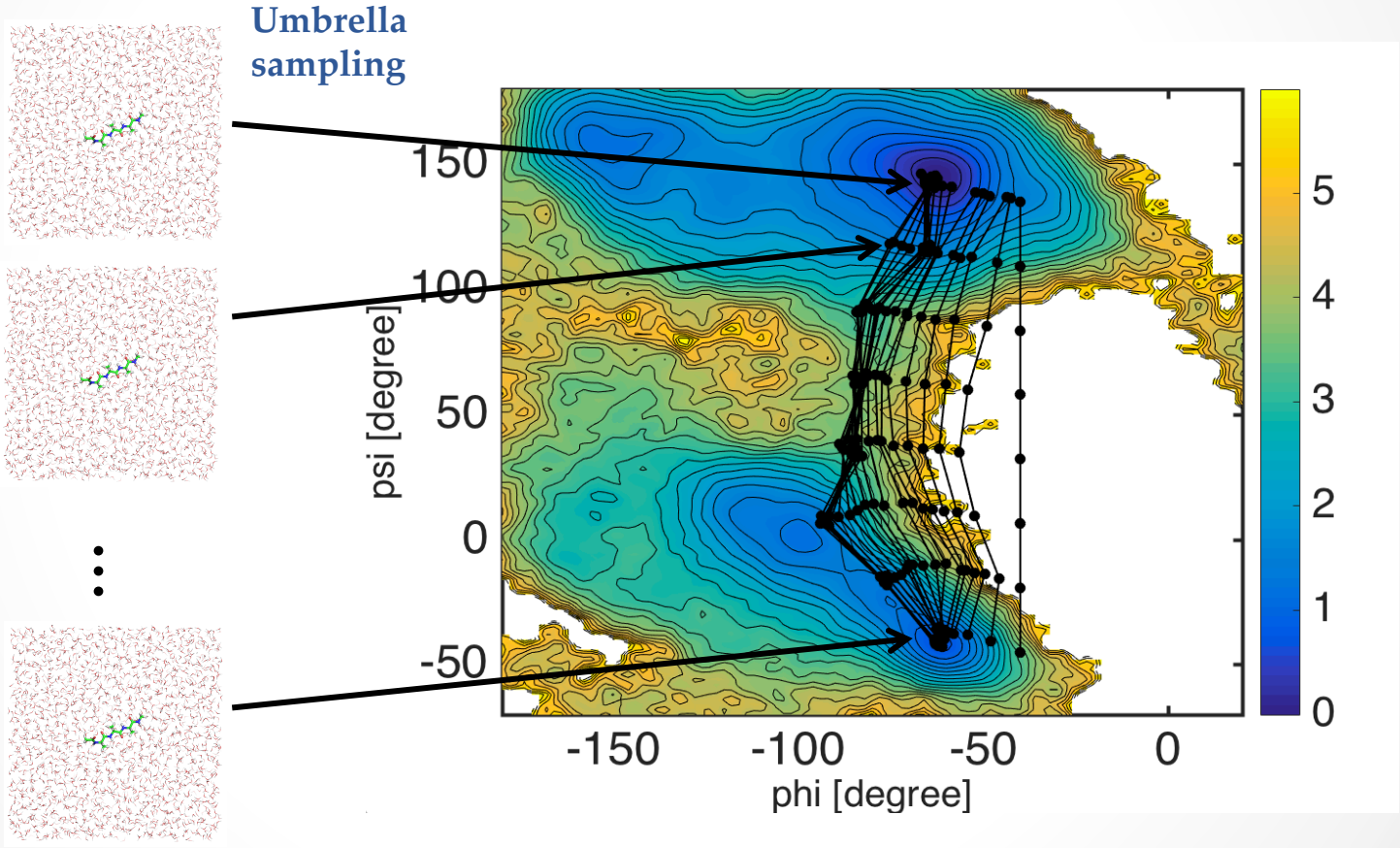
Step5. Visualization of pathways using a Python script

```
$ cd 5_visualize/
$ ls
plot.py   pmf.mat   xi.mat    yi.mat

# plot.py visualize ../4_string_method/*.rpath
$ python3 plot.py

# display plot.png or please scp it to your local machine
$ display plot.png
```

plot.png


Converged pathway
Initial pathway
Psi (degree)
Phi (degree)
PMF (kcal/mol)
150
100
50
0
−50
−150
−100
−50
0
5.4
4.8
4.2
3.6
3.0
2.4
1.8
1.2
0.6
0.0

Step6. Umbrella sampling around the converged pathway



Control file for GENESIS spdyn

The RPATH module in GENESIS can simulate replicas **without string method** performed (rpath_period = 0).

```
[INPUT]
rstfile = ../4_string_method/{}.rst

[OUTPUT]
logfile    =  {}.log
dcdfile    =  {}.dcd
rstfile    =  {}.rst
rpathfile  =  {}.rpath

[ENERGY]
forcefield        = CHARMM
electrostatic     = PME
switchdist        = 10.0
cutoffdist        = 12.0
pairlistdist      = 13.5

[DYNAMICS]
integrator        = LEAP
nsteps            = 10000
timestep          = 0.002
eneout_period     = 50
crdout_period     = 50
rstout_period     = 10000

[CONSTRAINTS]
rigid_bond        = YES     # use SHAKE/SETTLE
```

```
[ENSEMBLE]
ensemble          = NPT
tpcontrol         = LANGEVIN
temperature       = 300.00
pressure          = 1.0

[RPATH]
nreplica           = 8
rpath_period       = 0
rest_function      = 1 2

[SELECTION]
group1         = atomindex:15
group2         = atomindex:17
group3         = atomindex:19
group4         = atomindex:25
group5         = atomindex:27

[RESTRAINTS]
nfunctions     = 2

function1      = DIHED
constant1      = 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0
reference1     = 0 0 0 0 0 0 0 0
select_index1  = 1 2 3 4   # PHI

function2      = DIHED
constant2      = 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0
reference2     = 0 0 0 0 0 0 0 0
select_index2  = 2 3 4 5   # PSI
```

Stronger force constant

Images are automatically read from rst files

Running a job

```
$ cd 6_umbrella/
$ ls
run.inp      run.sh

$ cat run.sh
#!/bin/bash

export OMP_NUM_THREADS=3

source /home2/data/env/setpath.CPU.DP
mpirun -np 8 spdyn run.inp >run.out

$ ./run.sh
```

This will finish in about **13 minutes**

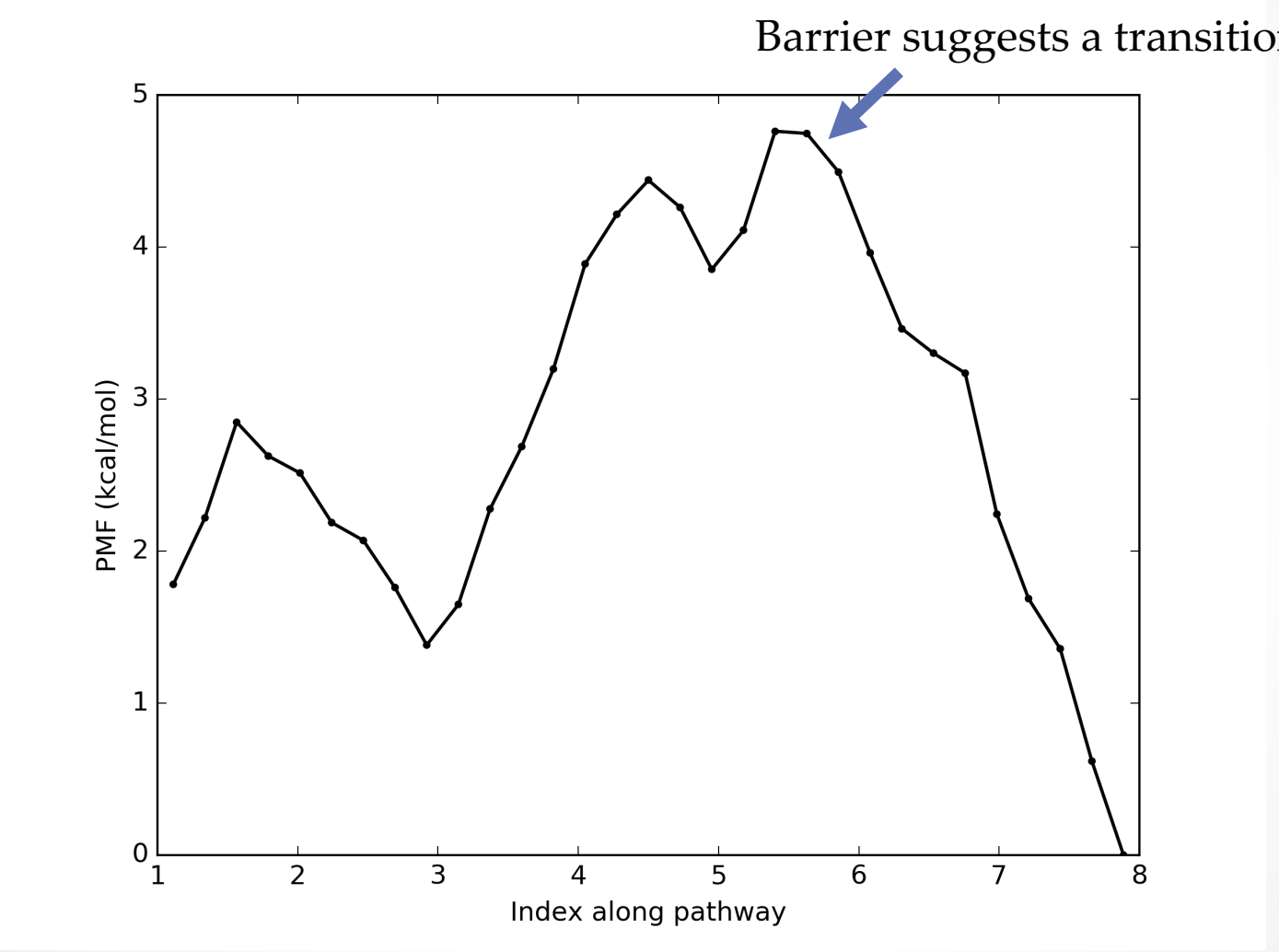
Step7. Analysis of free energy profile along the converged pathway

```
$ cd 7_analysis/
$ ls
mbar_analysis.inp   pathcv_analysis.inp   plot.pypmf_analysis.inp
remd_convert.inp    run.shtrj_analysis.inp

# run the analysis
$ ./run.sh

# display plot.png or please scp it to your local machine
$ display plot.png
```

What run.sh does

- Calls trj_analysis to calculate dihedral angles from umbrella sampling data
- Calls mbar_analysis to perform MBAR and obtain unbiasing weights from dihedral angle data
- Calls pathcv_analysis to get tangential and perpendicular coordinates to the pathway
- Calls pmf_analysis to evaluate free energy profile (pmf.dat) from the MBAR weights and tangential coordinates to the pathway

plot.png


Barrier suggests a transition state
PMF (kcal/mol)
Index along pathway
0
1
2
3
4
5
1
2
3
4
5
6
7
8

Advanced topics

- REUS instead of Umbrella sampling
- Cartesian coordinates instead of dihedral angles
- Targeted MD for generating an initial pathway
- Tool for preparing rst files along an initial pathway (rpath_generator)

Please check GENESIS website (some topics will be soo
illustrated in updated tutorials)

References

http://www.aics.riken.jp/labs/cbrt/tutorial/