Tutorial

March 5, 2019

Aim

The aim of this tutorial is to run a Temperature Accelerated Molecular Dynamics/ driven-Adiabatic Free Energy Dynamics (TAMD/d-AFED) and Parallel Bias Metadynamics (PBMetaD) simulation using the PLUMED-AMBER interface. We will set up a simple simulation of alanine dipeptide in vacuum, analyze the output, estimate free energies and its convergence from the simulation.

Theory

The probability of visiting a point is given by

$$P(s) = \frac{e^{-\beta F(s)}}{Z} \tag{1}$$

where s is defined as a course-grained coordinate which will be called as collective coordinate throughout this tutorial, $\beta = 1/k_B T,\ k_B$ is the Boltzman constant, T is temperature of system and Z is the configuration partition function given by

$$Z = \int ds e^{-\beta F(s)} \tag{2}$$

From eq.(1), It is clear that there are two ways of accelerating the sampling along the s coordinate,

1. modify the β factor, or
2. modify the $F(s)$ factor

This tutorial will help you to touch upon the above discussed two ways of accelerating the sampling of s space. It will help you to perform the simulation of a model system using two well known methods named as:

1. Temperature Accelerated Molecular Dynamics/ driven-Adiabatic free energy Dynamics (TAMD/d-AFED) which modifies the β factor, and
2. Parallel Bias Metadynamics (PBMetaD) which modifies the $F(s)$ factor.

1 Temperature Accelerated Molecular Dynamics/ driven-Adiabatic Free Energy Dynamics (TAMD/d-AFED)

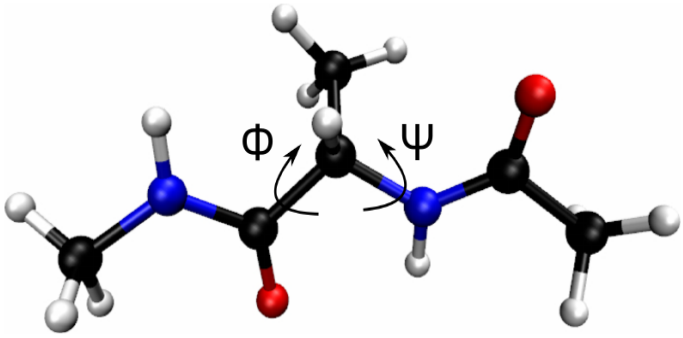
This method is proposed by J.B. Abrams and M. E. Tuckerman in 2008 where the sampling of collective variables (CVs) space in this method is accelerated by modifying the Boltzmann factor such that the temperature of CVs is kept high. It is done in extended space (non-physical) where the auxiliary variables are introduced which coupled with the physical/real CVs through a harmonic potential. To maintain the adiabaticity among the real CVs and auxiliary variables, the mass of the auxiliary variable are taken to be very high in comparison to the real CVs. The harmonic constant is taken to be high enough such that the motion of auxiliary variable should be exactly followed by the real variables. The free energy at normal temperature is given by

$$F(\mathbf{s}) = -\frac{1}{\tilde{\beta}} \ln \tilde{P}(\mathbf{s}) + f \tag{3}$$

where $\tilde{\beta} = 1/k_B\tilde{T}$ and $\tilde{T}$ is the temperature and $\tilde{P}(\mathbf{s})$ is the biased probability density of the accelerated CVs.
The method is applied on the model system to showcase the advantage of this method.

Alanine dipeptide system

The alanine dipeptide system will be used in vacuum as a model system. It is modelled using the ff14SB all-atom AMBER force field and the simulations were carried out using PLUMED-AMBER interface. The Collective variables used for this simulation are phi and psi torsion angles as shown in the figure as follows.



Setup and running a simulation
In order to run this simulation we need to prepare the required inputs for AMBER and PLUMED both seperately. The PLUMED input file is as follows

```
TORSION ATOMS=5,7,9,15 LABEL=phi
TORSION ATOMS=7,9,15,17 LABEL=psi

ex: EXTENDED_LAGRANGIAN ARG=phi,psi KAPPA=5260,5260 TAU=0.125,0.125
FRICTION=100,100 TEMP=1500

PRINT STRIDE=100 ARG=ex.phi_fict,ex.psi_fict FILE=COLVAR

PRINT ARG=ex.phi_vfict,ex.psi_vfict FILE=TEMP

PRINT STRIDE=100 ARG=phi,psi FILE=COLVAR_cc
```

The Label phi and psi represents the two torsional angels as shown in the figure and the TORSION ATOMS shows their serial numbers for their respective atoms involved in the torsion angle. KAPPA is the harmonic constant used to connect the real and auxiliary variables and TAU defines the mass of these auxiliary variables to maintain the adiabaticity. The plumed generates the outputs as FILES, COLVAR and COLVAR_cc files. The COLVAR_cc file prints the information of real variables and the COLVAR file prints the information about the auxiliary variables. The TEMP file prints the velocity of the auxiliary variables to check the temperature of the auxiliary variables.

The amber input files is as follows:

```
Alanine dipeptide
 &cntrl
  irest=0, ntx=1,
  imin=0, ntb=0,
  igb=0, ntpr=1000, ntwx=1000,
  ntt=3, gamma_ln=100.0,
  tempi=300.0, temp0=300.0
  nstlim=15000000, dt = 0.001,
  nscm=500,
  cut=999.0, rgbmax=999.0,
plumed=1,
plumedfile='plumed.dat'
&end
```

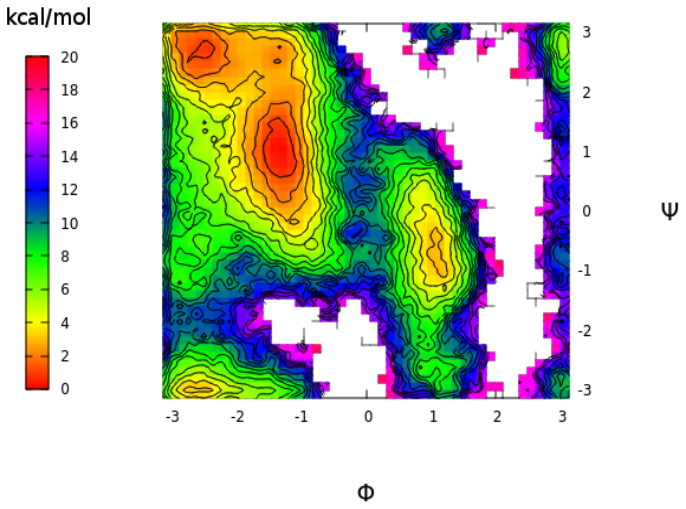
Please refer AMBER manual to understand the keywords used.

Once the topology, input coordinates, amber input, plumed input like necessary files are ready, the following sander command will run the simulation. The command for the sander is:

```
sander -O -i mds.in -o pbmtd.out -p prmtop -c diala.rst -r pbmtd.rst -x
pbmtd.mdcrd
```

Analysis of results

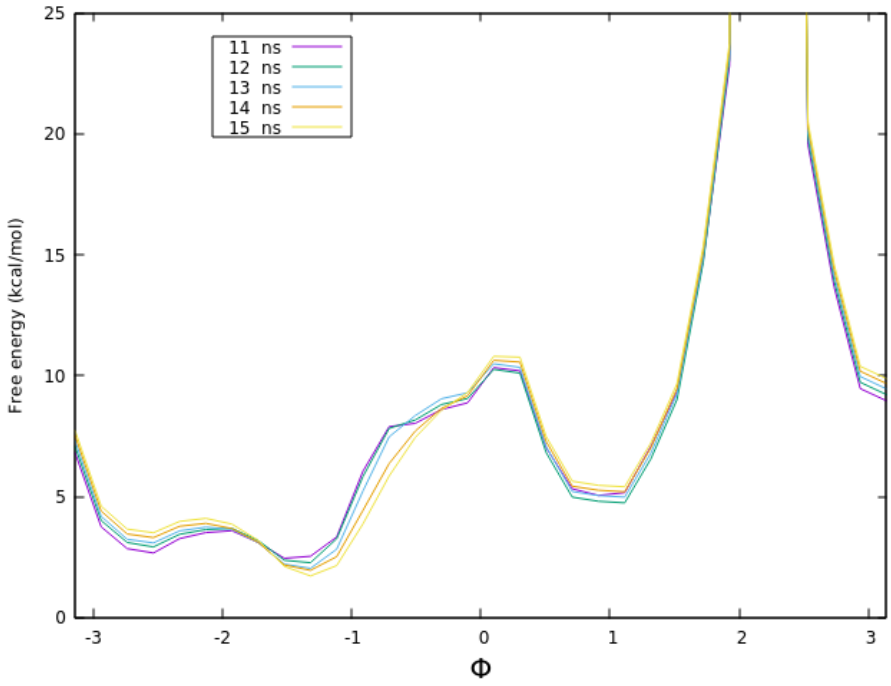
1. **Calculating free energy landscape along CVs**
   The simulation is ran at high temperature but our aim is to calculate the free energy at normal temperature. The reweighting code will assist you to make a room temperature two dimensional free energy landscape.

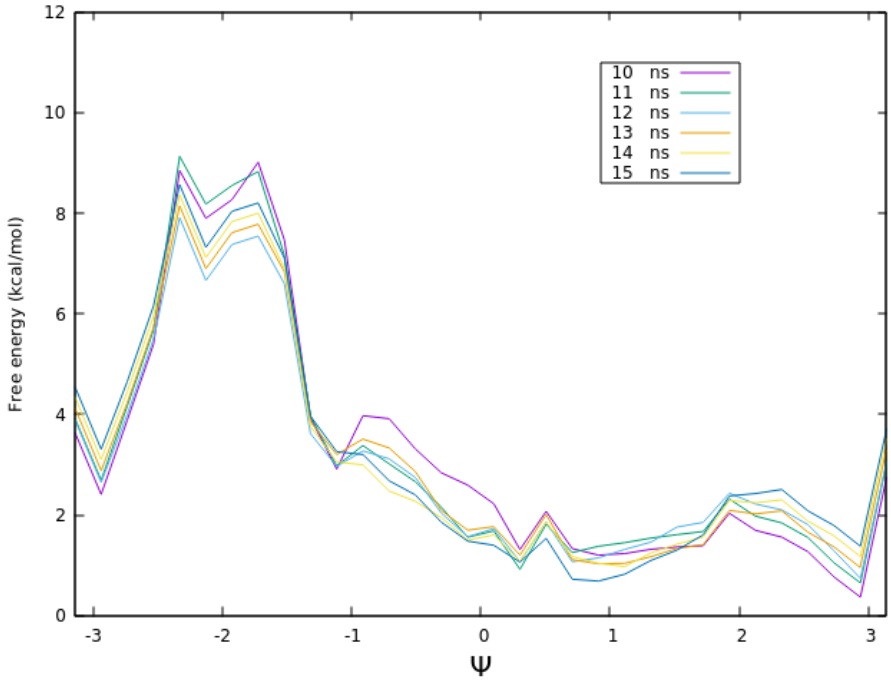


   The roughness on the surface clearly shows the effect of high temperature. It can be clearly seen that the temperature used was not enough to sample the whole free energy surface.

2. **Calculating one dimensional projection and checking convergence**

   The one dimensional projection is plotted from two dimensional free energy landscape as a function of time to check the convergence. The projection are found to be converged after 10 ns of simulation with an error of 0.5 kcal/mol.


11 ns
12 ns
13 ns
14 ns
15 ns
Free energy (kcal/mol)
Φ


10 ns
11 ns
12 ns
13 ns
14 ns
15 ns
Free energy (kcal/mol)
Ψ

2 Parallel Bias Metadynamics (PBMeTaD)

It is one of the recent enhanced sampling method proposed by Jim Pfaendtner and Massimiliano Bonomi in 2015 where low dimensional metadynamics (MTD) bias is applied on more that one CVs simultaneously. The probability of updating bias is determined by a factor on the fly to remove the correlation effect on CVs so that the convergence will not be affected.This way PBMetaD help us constructing multidimensional free energy landscape. The bias added at any instant of time(t) is given by

$$V_{PB}(S_1, ..., S_n, t) = -\frac{1}{\beta} \ln \left[\sum_{\alpha=1}^{n} \exp(-\beta V_\alpha(S_\alpha, t)) \right] \tag{4}$$

The probability of adding bias along any k^{th} CV out of total n CVs is given by

$$P_k(S_k) = \frac{\exp\{(-\beta V_k(S_k, t))\}}{\sum_{\alpha=1}^{n} \exp\{(-\beta V_\alpha(S_\alpha, t))\}} \tag{5}$$

This way the height of MTD bias is further scaled according to PBMetaD along each CV in the following fashion

$$h_k(t) = h_k(t) P_k(S_k) \tag{6}$$

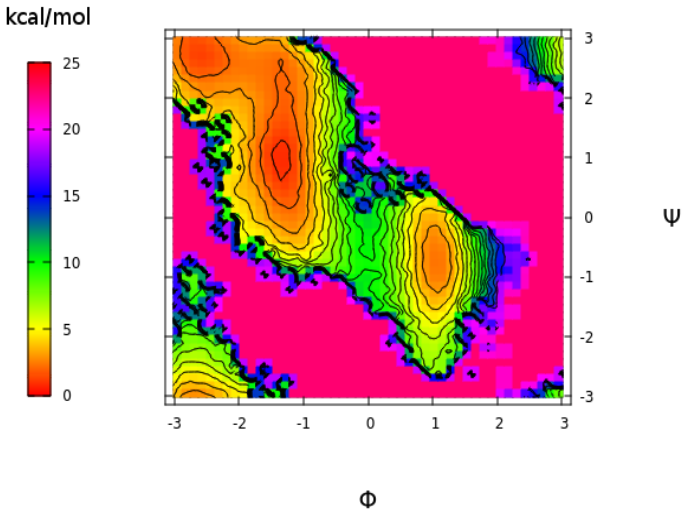
This way PBMetaD is able to add bias along many CVs simultaneously unlike MTD which is restricted to two or three CVs. The method is applied on alanine dipeptide using the same parameters as in TAMD/d-AFED tutorial part. The plumed input is changed as follows:

```
TORSION ATOMS=5,7,9,15 LABEL=phi
TORSION ATOMS=7,9,15,17 LABEL=psi

PBMETAD ARG=phi,psi PACE=500 HEIGHT=2.4 SIGMA=0.05,0.05 FILE=HILLS_phi,HILLS_psi
BIASFACTOR=10.0 LABEL=pb

PRINT STRIDE=50 ARG=phi,psi,pb.bias FILE=COLVAR
```

The keyword PACE in the plumed input determines the stride of Gaussian deposition in number of time steps, while the keyword HEIGHT specifies the height of the initial Gaussian in kJ/mol along every CV. For each CVs, one has to specified the width of the Gaussian by using the keyword SIGMA. Bias information will be written to the file indicated by the keyword FILE.

Analysis of results

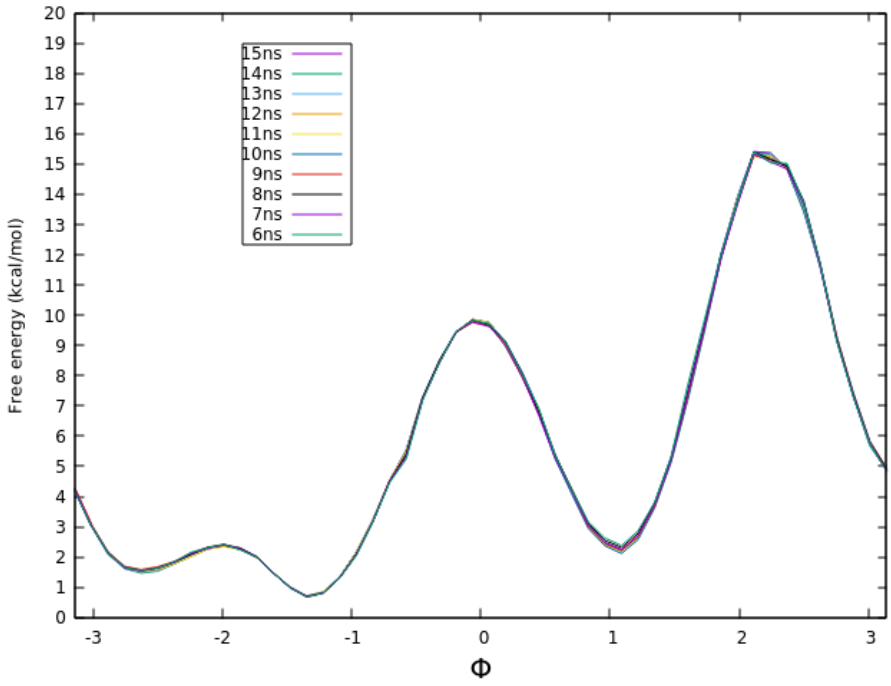
1. **Calculating free energy landscape along CVs**
   We accelerate the sampling by adding external bias into the system but our aim is to get the probability when there is no bias in the system. The reweighting code will assist you to make a two dimensional free energy landscape when there is no external bias added into the system.

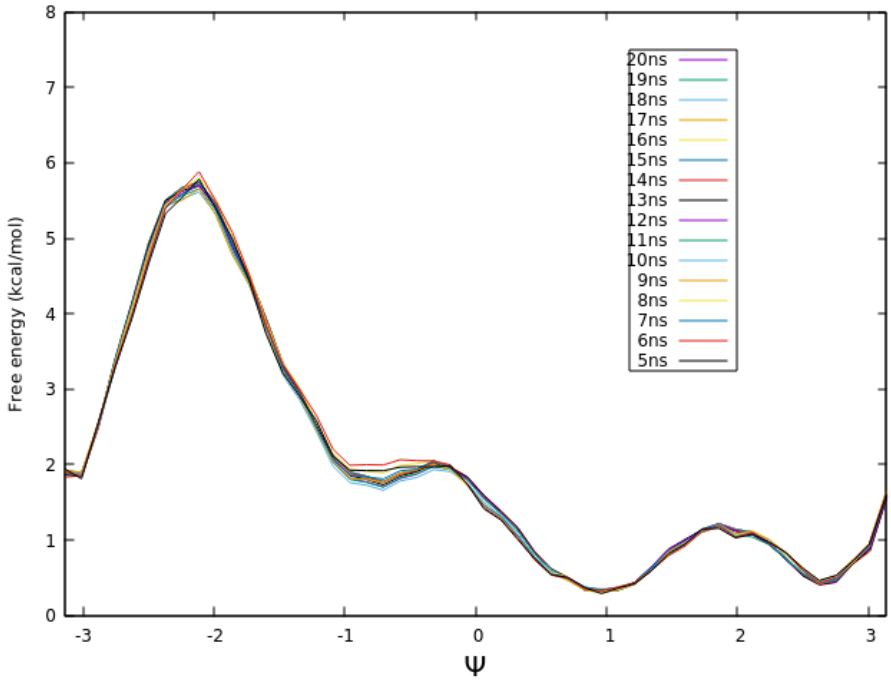


   It is seen that PBMetaD shows noisy behaviour as it was also mentioned in the real PBMeatD paper.

2. **Calculating one dimensional projection and checking convergence**

   The one dimensional projection is projected form two dimensional free energy landscape and the convergence is checked.
   The projection are found to be converged after 7 ns of simulation with less that 0.1 kcal/mol difference.


15ns
14ns
13ns
12ns
11ns
10ns
9ns
8ns
7ns
6ns
Free energy (kcal/mol)
Φ


20ns
19ns
18ns
17ns
16ns
15ns
14ns
13ns
12ns
11ns
10ns
9ns
8ns
7ns
6ns
5ns
Free energy (kcal/mol)
Ψ

(ABHINAV GUPTA)