

University of North Carolina at Chapel Hill

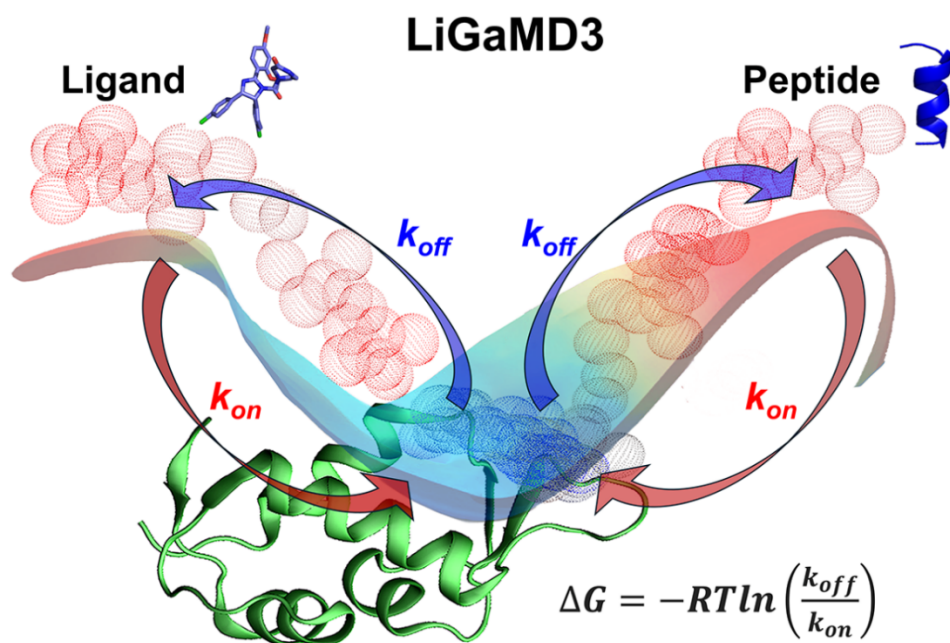
Department of Pharmacology and

Computational Medicine Program

GaMD Workshop 2024

## Ligand Gaussian accelerated Molecular Dynamics 3 (GaMD3) Tutorial

Jinan Wang, Amy Liu, Keya Joshi, Victor Adediwura and Yinglong Miao



## Abstract

This tutorial demonstrates how to use AMBER24+ and associated tools to perform and analyze LiGaMD3 simulations. The simulations can capture repetitive dissociation and binding of small molecules and flexible peptides, thereby exploring their binding mechanisms, thermodynamics and kinetics. Using host-guest binding of  $\beta$ -cyclodextrin and 1-butanol as a model system, we will demonstrate how to conduct LiGaMD3 simulations to capture ligand dissociation and rebinding. Additionally, we will demonstrate how to use the *PyReweighting* scripts to calculate 1D and 2D free energy landscapes and employ a 3D reweighting approach to predict ligand binding free energy. Prior knowledge of AMBER and basic Molecular Dynamics (MD) simulations is required.

## Table of Contents

<b>1. Introduction</b>	4
<b>2. RUNNING LiGAMD3 SIMULATIONS ON HOST-GUEST SYSTEM</b>	6
2.1 System Setup	6
2.2 LiGAMD3 Equilibration	7
2.3 LiGAMD3 Production	12
<b>3. SIMULATION ANALYSIS</b>	15
3.1 1D Potential of Mean Force Calculation	15
3.3 2D Potential of Mean Force Calculation	17
3.3 3D Potential of Mean Force Calculation	20
<b>4. QUESTIONS</b>	21
<b>5. REFERENCES</b>	21

## 1. Introduction

LiGaMD3 is an enhanced sampling method that applies triple boosts to three key energy terms involved in ligand dissociation, rebinding, and conformational changes in the system. This significantly enhances the sampling efficiency of interactions between small molecules or peptides and target proteins. In benchmark studies using the MDM2 system bound by a small molecule (Nutlin 3) and two highly flexible peptides (PMI and P53), LiGaMD3 efficiently captured repeated dissociation and binding events within 2-microsecond simulations. Therefore, LiGaMD3 offers a robust and efficient approach for modeling binding of small molecule and flexible peptide ligands. Details of the LiGaMD3 method and reweighting algorithms, are available in References 1 and 2.

The goal of this tutorial is to guide users through setting up, performing and analyzing LiGaMD3 simulations using AMBER24+. For simplicity and efficiency, we will use host-guest binding of  $\beta$ -cyclodextrin and 1-butanol as our model system. After running the LiGaMD3 simulations, the *PyReweighting* scripts will be employed to calculate 1D and 2D Potential of Mean Force (PMF) free energy profiles. Finally, 3D reweighting will be used to predict the ligand binding free energy.

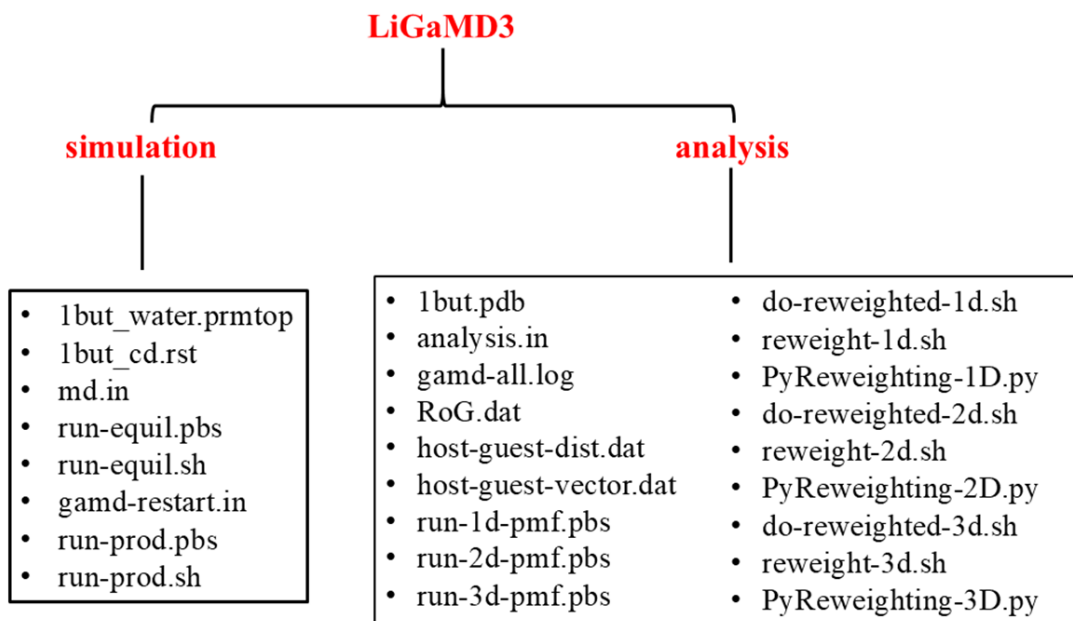
LiGaMD3 is implemented only in the GPU version of AMBER (*pmemd.cuda*) starting from AMBER 2024. Therefore, it is recommended to use AMBER24+ for these simulations. This tutorial assumes that the readers are familiar with using AMBER for conventional MD (cMD) simulations, such as energy minimization, heating, equilibration and cMD simulations. General documentations, tutorials, and configuration file templates are available in the Tutorial Section of the AMBER website (<https://ambermd.org/tutorials/>). The simulation lengths in this tutorial are intentionally reduced for demonstration; however, it is important to note that they are typically insufficient for practical applications. Additionally, sufficient equilibration of the system with cMD simulations before running LiGaMD3 simulations is highly recommended.

### Completion of this tutorial requires:

- AMBER24+ installed (<https://ambermd.org>).
- Python3 with NumPy and SciPy installed (<https://www.python.org/>).
- Files required for the LiGaMD3 tutorial can be download at [LiGaMD3.tar.bz2](#) or copied from folder `/shared/data/LiGaMD3/` at UNC cloud server. Extract files by command in Linux system:

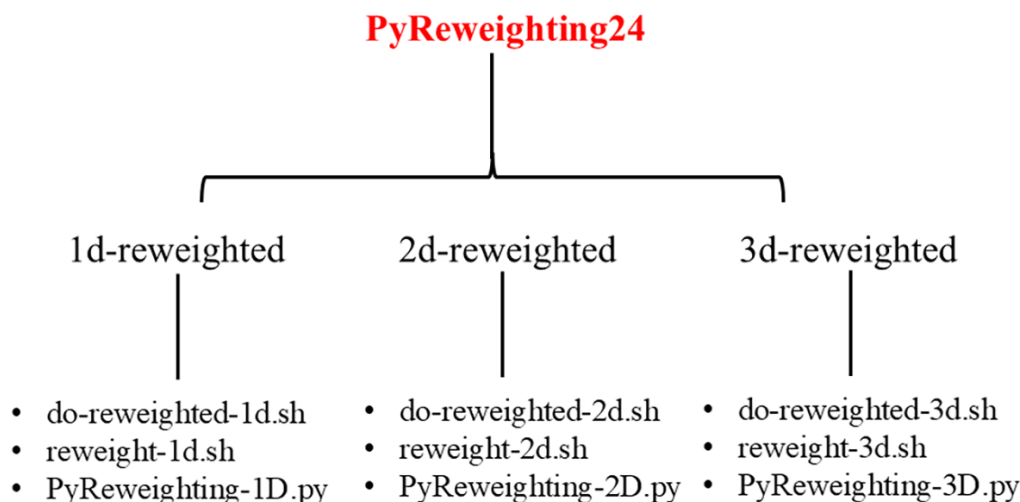
`tar -xvjf LiGaMD3.tar.bz2`. As shown in **Fig. 1A**, the LiGaMD3 folder contains two sub-folders: **simulation** and **analysis**.

- The **simulation** folder includes eight files: *lbut\_water.prmtop* and *lbut\_cd.rst*, representing the system's topology and coordinate files after 20ns cMD, respectively; *md.in* and *gamd-restart.in*, representing the input files for LiGaMD3 equilibration and production, respectively; *run-equil.pbs* and *run-prod.pbs*, the scripts used to submit jobs in the UNC cloud via the SLURM queue system; *run-equil.sh* and *run-prod.sh*, the bash scripts to run corresponding LiGaMD3 equilibration and production simulations using your own workstations.
- The **analysis** folder includes 18 files: *lbut.pdb* (initial PDB file for visualization), *analysis.in* (input file for *cpptraj*), *gamd-all.log* (the GaMD log file), *host-guest-dist.dat* (reaction coordinate file calculated with *cpptraj*), *RoG.dat* (reaction coordinate file calculated with *cpptraj*), *host-guest-vec.dat* (reaction coordinate file calculated with *cpptraj*), *run-[1-3]d-pmf.pbs* (SLURM submitting scripts for UNC cloud server), *run-[1-3]d-pmf.sh* (bash scripts for 1-3D PMF calculations using your own workstation), *reweight-[1-3]d.sh* (bash scripts for reweighting calculations), *PyReweighting-[1-3]D.py* (Python scripts for reweighting calculations).



**Figure 1:** Directory structure of LiGaMD3 tutorial files.

• PyReweighting scripts can be download at [PyReweighting24.tar.bz2](#) or copied from following folder /shared/data/PyReweighting24 on the UNC cloud server. As shown in **Fig. 2**, there are three folders: 1d-reweighted, 2d-reweighted, and 3d-reweighted, each corresponding to the calculation of 1D, 2D, and 3D PMFs. Each folder includes two bash scripts (*do-reweighted-[1-3]d.sh* and *reweight-[1-3]d.sh*), and a Python script (*PyReweighting-[1-3]D.py*) for performing the respective calculations.



**Figure 2:** Directory structure of PyReweighting24 files.

## 2. RUNNING LiGAMD3 SIMULATIONS ON HOST-GUEST SYSTEM

### 2.1 System Setup

LiGaMD3 is performed after well-equilibrated cMD simulations. Detailed instructions for system setup and cMD simulations can be found on the AMBER tutorial website (<https://ambermd.org/tutorials/>). Two recommended approaches for setting up the system are through *tLeap* in AMBER or the CHARMM-GUI website. Following tutorials are highly recommended:

- *tLeap*: <https://ambermd.org/tutorials/pengfei/index.php>
- CHARMM-GUI: <https://ambermd.org/tutorials/CHARMM-GUI.php>

For simplicity, we have prepared the system files after energy minimization, heating, equilibration with restraints, and 20 ns cMD simulations (under directory **simulation**). These files can be found in the “**simulation**” folder, including the topology (*lbut\_water.prmtop*) and coordinate (*lbut\_cd.rst*). Since cMD is not the focus of this tutorial, we use these files directly.

**Following are instructions for this part using UNC cloud server.** Make sure that the AMBER environment is properly configured by running appropriate *source* command. Copy the “LiGaMD3” folder located at /shared/data/LiGaMD3 to your home directory.

```
source /shared/apps/software/amber24/amber.sh
cd ~/
cp -r /shared/data/LiGaMD3 ./
cd ~/LiGaMD3/simulation
```

**Following are instructions for users using your own workstation.** Make sure that you have uploaded the LiGaMD3.tar.bz2 to your home directory. Extract the required files from tar file for the current tutorial.

```
source /shared/apps/software/amber24/amber.sh ## this one needs to be modified according to your own local system
cd ~/
tar -xvjf LiGaMD3.tar.bz2
cd ~/LiGaMD3/simulation
```

## 2.2 LiGaMD3 Equilibration

The *pmemd.cuda* in AMBER 24+ will be used to perform LiGaMD3 equilibration. Three files are required for LiGaMD3 simulation:

- 1but\_water.prmtop (the system topology file)
- 1but\_cd.rst (the coordinate file, after 20 ns of cMD simulations)
- md.in (the LiGaMD3 input file)

The command to run LiGaMD3 simulations is the same as conventional AMBER simulations:

```
pmemd.cuda -O -i md.in -p 1but_water.prmtop -c 1but_cd.rst -o md-1.out -x md-1.nc -r gamd-1.rst -gamd gamd-1.log &
```

The output files from these simulations are as follows:

- md-1.out: Contains the energy information from the simulation.
- md-1.nc: Trajectory file, storing the MD simulation frames.

- `gamd-1.rst`: Restart file, including both coordinate and velocity information.
- `gamd-1.log`: Logs the boost potential information used for reweighting.
- `gamd-restart.dat`: Stores the potential statistics used to apply boosts for the LiGaMD3 simulations.
- `mdinfo`: Updates the current state of the simulation, useful for monitoring progress and checking the simulation status.

**Following are instructions for running the LiGaMD3 equilibration using UNC cloud server.**

Make sure that you are in the `~/LiGaMD3/simulation` directory and that the required files, `run-equil.pbs`, `md.in`, `lbut_water.prmtop` and `lbut_cd.rst`, are available in this directory. A bash script (`run-equil.pbs`) has been prepared for running the job.

```
cd ~/LiGaMD3/simulation
sbatch run-equil.pbs
squeue
```

One copy of batch script `run-equil.pbs` to run LiGaMD3 equilibration on UNC cloud server.

```
#!/bin/sh
#SBATCH --job-name=equilbration
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=00:20:00
#SBATCH --output=log.%x.%j
#SBATCH --gres=gpu:1
#SBATCH --qos=gpu_access
source /shared/apps/software/amber24/amber.sh
which pmemd.cuda
# gamd equilibration
pmemd.cuda -O -i md.in -p lbut_water.prmtop -c lbut_cd.rst -o md-1.out -x md-1.nc -r
gamd-1.rst -gamd gamd-1.log
```

The current simulations take approximately 10 minutes to finish in the UNC cloud server.



Following are instructions for running the LiGaMD3 equilibration using your own workstation. Make sure that you are in the ~/LiGaMD3/simulation directory and all required files, *run-equil.sh*, *md.in*, *lbut\_water.prmtop* and *lbut\_cd.rst*, are present in this directory. A bash script (*run-equil.sh*) has been prepared for running the equilibration job.

```
cd ~/LiGaMD3/simulation
chmod +x run-equil.sh
./run-equil.sh &
```

A copy of script *run-equil.sh* is shown below:

```
#!/bin/sh
#make sure to modify the path of amber24 to the correct location
#source /shared/apps/software/amber24/amber.sh
which pmemd.cuda
# gamd equilibration
pmemd.cuda -O -i md.in -p lbut_water.prmtop -c lbut_cd.rst -o md-1.out -x md-1.nc -r gamd-1.rst -gamd gamd-1.log
```

Following is an example of *md.in* file for the LiGaMD3 equilibration simulations.

.....  
GaMD equilibration simulation

```
&cntrl
imin=0,      ! No minimization
irest=0,      ! This IS a new MD simulation
ntx=1,        ! read coordinates only
! Temperature control
ntt=3,        ! Langevin dynamics
gamma_ln=1.0, ! Friction coefficient (ps^-1)
tempi=300.0, ! Initial temperature
temp0=300.0,  ! Target temperature
ig=-1,        ! random seed
! Potential energy control
cut=9.0,      ! nonbonded cutoff, in angstroms
! MD settings
nstlim=10100000, ! simulation length
dt=0.002,     ! time step (ps)
! SHAKE
ntc=2,      ! Constrain bonds containing hydrogen
ntf=1,      ! Complete interaction is calculated
! Control how often information is printed
ntpr=1000,    ! Print energies every 1000 steps
ntwx=1000,    ! Print coordinates every 1000 steps to the trajectory
ntwr=5000,    ! Print a restart file every 5000 steps
ntxo=2,       ! Write NetCDF format
```

```

ioutfm=1,    ! Write NetCDF format
! Wrap coordinates when printing them to the same unit cell
iwrap=1,
ntwprt = 168,
! Constant pressure control. Note that ntp=3 requires barostat=1
barostat=1, ! Berendsen... change to 2 for MC barostat
ntp=1,       ! 1=isotropic, 2=anisotropic, 3=semi-isotropic w/ surften
pres0=1.0,   ! Target external pressure, in bar
taup=0.5,    ! Berendsen coupling constant (ps)
! LiGaMD3 parameters
igamd = 28, iE = 1, irest_gamd = 0,
ntcmd = 100000, nteb = 10000000,
ntave = 20000, ntcmdprep = 40000,
ntebprep = 40000, sigma0P = 0.1,
sigma0D = 6.0, sigma0B=6.0,
iEP=1,iED=1,iEB=1,
bgpro2atm=1,edpro2atm=147,
gti_cpu_output = 0, gti_add_sc = 1,
icfe = 1,
ifsc = 1,
timask1 = ':8',
scmask1 = ':8',
timask2 = ',',
scmask2='',
/

```

In this tutorial, we reduced *nteb* to 400,000 and *nstlim* to 500,000 to save simulation time.

**Below are some specific parameters for LiGaMD3 simulations.**

```

igamd          Flag to apply boost potential
                  = 28 triple boost on the ligand essential potential energy, the remaining non-bonded and
                      bonded potential energy of the entire system

iEP, iED, iEB  Flag to set the threshold energy  $E$ 
                  = 1 (default) set threshold energy to lower bound  $E=V_{\max}$ 
                  = 2 set threshold energy to upper bound  $E=V_{\min}+(V_{\max}-V_{\min})/k_0$ 

irest_gamd     Flag to restart GaMD simulation
                  = 0 (default) new simulation
                  = 1 restart simulation

ntcmd          Number of cMD steps to calculate  $V_{\max}$ ,  $V_{\min}$ ,  $V_{\text{avg}}$ ,  $\sigma_V$  (default 1,000,000)

nteb           Number of biasing equilibration steps (default 1,000,000)

sigma0P        Upper limit of the standard deviation of the first potential boost (default 6.0 kcal/mol).

```

***sigma0D*** Upper limit of the standard deviation of the second potential boost (default 6.0 kcal/mol).

***sigma0B*** Upper limit of the standard deviation of the third potential boost (default 6.0 kcal/mol).

***timask1*** Specifies atoms of the ligand. This variable has already been used in AMBER. The default is an empty string.

***scmask1*** Specifies atoms of the ligand that will be described using soft core in ambmask format. The default is an empty string.

***bgpro2atm*** Start atomic number of the protein region.

***edpro2atm*** End atomic number of the protein region.

***icfe*** The basic flag for free energy calculations.

***ifsc*** Flag for softcore potentials  
 = 0 SC potentials are not used (default)  
 = 1 SC potentials are used.

***gti\_cpu\_output*** = 1 (default): the softcore  $\lambda$ -derivative terms will be combined and output the same result for each TI region (to match the CPU-version output)  
 = 0: the softcore  $\lambda$ -derivative terms will be output for each TI region.

.....

Other parameters follow the cMD settings; Please refer to the AMBER manual for detailed information. Note that for LiGaMD3, you should set ***ntf=1*** to calculate all interactions and ***barostat=1*** to control pressure, as these settings have been widely tested. Below are some hints for setting parameters in the md.in input files, particularly for LiGaMD3 equilibration:

- **ntave**: Set to **4 x total number of atoms in the simulation**
- **ntcmd**: Set to **5 x ntave**
- **ntcmdprep**: Set to **2 x ntave**
- **ntebprep**: Set to **2 x ntave**
- **nteb**: Set to **tens of ns**, but this should require multiple times of ntave
- **nstlim**: Should be the sum of nteb and ntcmd

It is essential that the gamd-restart.dat file is generated in the working directory, as it will be used to calculate the parameters for applying the boost. In LiGaMD3, this file should contain three lines with the values for Vmin, Vmax, Vavg, and  $\sigma_V$  for each boost energy term.

3.1631393432617188	-24.502395661547780	-11.506001469540969	2.4883399570573745
-16003.105296557231	-16984.280963839512	-16093.121229529763	22.401105159053714
279.76529085077345	175.88247824925929	249.03960114756225	6.8265080264473452

## 2.3 LiGaMD3 Production

The command to perform LiGaMD3 production simulations is the same as that used for LiGaMD3 equilibration simulation, but with different input files (*gamd-restart.in* and *gamd-1.rst*) as following:

```
pmemd.cuda -O -i gamd-restart.in -p 1but_water.prmtop -c gamd-1.rst -o md-2.out -x md-2.nc -r gamd-2.rst -gamd gamd-2.log
```

All necessary files can be found in the '**simulation**' directory extracted from the LiGaMD3.tar.bz2 or under shared folder “/shared/data/LiGaMD3/simulation” in the UNC cloud server (**Fig. 1**).

**The following instructions are for running the LiGaMD3 production on the UNC cloud server.** Make sure that you are in the ~/LiGaMD3/simulation directory and all following necessary files are available in the simulation directory.

- *run-prod.pbs*
- *gamd-restart.in*
- *1but\_water.prmtop*
- *gamd-1.rst*
- *gamd-restart.dat*

We have prepared a script (*run-prod.pbs*) for using the SLURM to submit LiGaMD3 production simulation jobs.

```
cd ~/LiGaMD3/simulation  
sbatch run-prod.pbs  
squeue
```

A copy of *run-prod.pbs* is shown below:

```
#!/bin/sh
#SBATCH --job-name=prod
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=00:20:00
#SBATCH --output=log.%x.%j
#SBATCH --gres=gpu:1
#SBATCH --qos=gpu_access
source /shared/apps/software/amber24/amber.sh
which pmemd.cuda
#gamd production
pmemd.cuda -O -i gamd-restart.in -p 1but_water.prmtop -c gamd-1.rst -o md-2.out -x md-2.nc
-r gamd-2.rst -gamd gamd-2.log
```

This will generate the trajectory file (*md-2.nc*) and the log file (*gamd-2.log*). Both files will be used for post analysis.

**The following are instructions for running the LiGaMD3 production using your own workstation:** Make sure that you are in the *~/LiGaMD3/simulation* directory and all following required files are present in this directory.

- *run-prod.sh*
- *gamd-restart.in*
- *1but\_water.prmtop*
- *gamd-1.rst*
- *gamd-restart.dat*

A bash script (*run-prod.sh*) has been prepared for running the jobs.

```
cd ~/LiGaMD3/simulation
chmod +x run-prod.sh
./run-prod.sh &
```

A copy of *run-prod.sh* is shown below:

```
#!/bin/sh
#source /shared/apps/software/amber24/amber.sh
which pmemd.cuda
#gamd production
pmemd.cuda -O -i gamd-restart.in -p 1but_water.prmtop -c gamd-1.rst -o md-2.out -x md-2.nc
-r gamd-2.rst -gamd gamd-2.log
```

Following is an example of input file for the GaMD3 production simulations.

.....  
GaMD production simulation

```
&cntrl
  imin=0,      ! No minimization
  irest=0,     ! This IS a new MD simulation
  ntx=1,       ! read coordinates only
  ! Temperature control
  ntt=3,       ! Langevin dynamics
  gamma_ln=1.0, ! Friction coefficient (ps^-1)
  tempi=300.0,  ! Initial temperature
  temp0=300.0, ! Target temperature
  ig=-1,       ! random seed
  ! Potential energy control
  cut=9.0,     ! nonbonded cutoff, in angstroms
  !fswitch=10.0, ! Force-based switching
  ! MD settings
  nstlim=500000, ! simulation length
  dt=0.002,    ! time step (ps)
  ! SHAKE
  ntc=2,       ! Constrain bonds containing hydrogen
  ntf=1,       ! Complete interaction is calculated

  ! Control how often information is printed
  ntp=100,     ! Print energies every 1000 steps
  ntwx=100,    ! Print coordinates every 1000 steps to the trajectory
  ntwr=5000,   ! Print a restart file every 10K steps (can be less frequent)
  ntxo=2,      ! Write NetCDF format
  ioutfm=1,    ! Write NetCDF format (always do this!)
  ! Wrap coordinates when printing them to the same unit cell
  iwrap=1,
  ntwprt = 162,
  ! Constant pressure control. Note that ntp=3 requires barostat=1
  barostat=1,  ! Berendsen... change to 2 for MC barostat
  ntp=1,       ! 1=isotropic, 2=anisotropic, 3=semi-isotropic w/ surften
  pres0=1.0,   ! Target external pressure, in bar
  taup=0.5,    ! Berendsen coupling constant (ps)

  ! GaMD parameters
  igamd = 28, iE = 1, irest_gamd = 1,
  ntcmd = 0, nteb = 0, ntave = 20000,
  ntcmdprep = 0, ntebprep = 0,
  sigma0P = 0.1, sigma0D = 6.0, sigma0B=6.0, iEP=1, iED=1, iEB=1,
  bgpro2atm=1, edpro2atm=147,
  gti_cpu_output = 0, gti_add_sc = 1,
  icfe = 1,
  ifsc = 1,
```

```
timask1 = ':8',  
scmask1 = ':8',  
timask2 = "",  
scmask2 = "",
```

```
/
```

Note: To save simulation time, we set a smaller *nstlim*=500,000, corresponding to a brief 10 ns run, allowing the simulation to finish within approximately 10 minutes. However, for the actual simulations, hundreds of nanoseconds will be required. AMBER provides the *ntwprt* option, which is useful for saving a subset of atoms in the trajectory. This is particularly helpful for large systems where the solvent is not of primary interest. One important parameter is that for the LiGaMD3 production, we need to set *irest\_gamd*=1 as we don't want to update the potential statistic anymore.

*irest\_gamd*     Flag to restart GaMD simulation  
                 = 0 (default) new GaMD simulation  
                 = 1 restart GaMD simulation

### 3. SIMULATION ANALYSIS

#### 3.1 1D Potential of Mean Force Calculation

The most widely used reaction coordinates include distance, angle, and RMSD, which can be calculated using *cpptraj*. Since usage of *cpptraj* is not our focus, we have calculated the reaction coordinate files saved in the **analysis** folder, including *host-guest-dist.dat*, *RoG.dat*, and *host-guest-vector.dat*. You can navigate to the analysis folder and review these files by using following command:

```
cd ~/LiGaMD3/analysis && ls
```

In this section, we will use the distance between host and guest (i.e., *host-guest-dist.dat*) as the reaction coordinate to calculate the 1D Potential of Mean Force (PMF). The scripts *do-reweighted-1d.sh*, *reweighted-1d.sh*, and *Pyreweighting-1D.py* are required for performing these calculations.

**The following instructions outline how to run the 1D reweighting job on the UNC cloud server.** Make sure that following required files are in the `~/LiGaMD3/analysis` directory:

- *host-guest.dat*
- *gamd-all.log*
- *run-1d-pmf.pbs*
- *do-reweighted-1d.sh*
- *PyReweighting-1D.py*

We will use SLURM to submit the 1D reweighting jobs, rather than executing them on the login node. The script named *run-1d-pmf.pbs* has been prepared for job submission.

```
cd ~/ligamd3/analysis
sbatch run-1d-pmf.pbs
squeue
```

A copy of script *run-1d-pmf.pbs* is shown below.

```
#!/bin/sh
#SBATCH --job-name=pmf-1d
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=00:20:00
#SBATCH --output=log.%x.%j

source /shared/apps/software/amber24/amber.sh
cd ~/LiGaMD3/analysis
chmod +x do-reweighted-1d.sh
./do-reweighted-1d.sh host-guest.dat gamd-all.log
```

**The following instructions outline how to run the 1D reweighting job using your own workstation.** Make sure that the following required files are in the `~/LiGaMD3/analysis` directory

- *host-guest.dat*
- *gamd-all.log*
- *run-1d-pmf.sh*
- *do-reweighted-1d.sh*
- *PyReweighting-1D.py*

The script named *run-1d-pmf.sh* is used to do the 1D reweighting calculation.

```
cd ~/LiGaMD3/analysis
chmod +x run-1d-pmf.sh
./run-1d-pmf.sh &
```



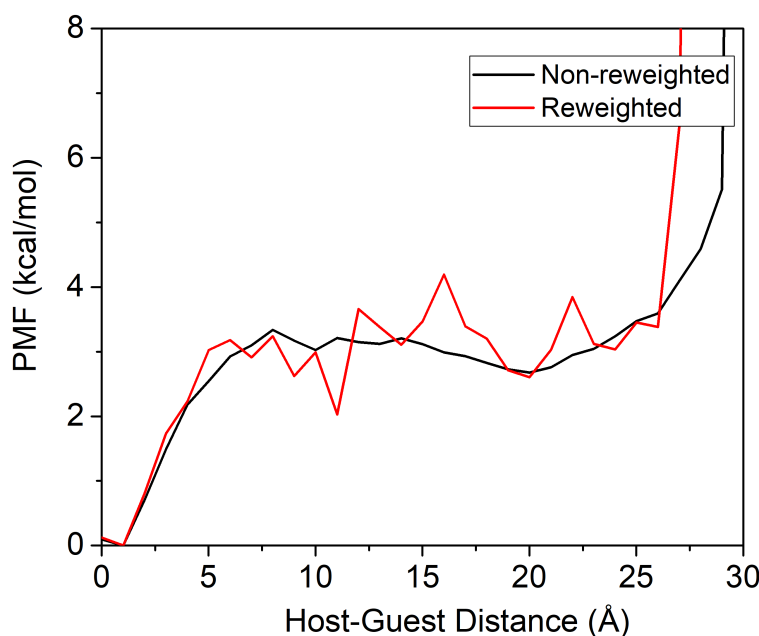
A copy of *run-1d-pmf.sh* is shown below:

```
#!/bin/sh
cd ~/LiGaMD3/analysis
chmod +x do-reweighted-1d.sh
./do-reweighted-1d.sh host-guest-dist.dat gamd-all.log
```

After completing the calculation, you will obtain corresponding PMF files, being saved in the *pmf-host-guest-dist-bin1.0-cutoff500* subfolder.

- *pmf-c2-Phi.dat-reweight-disc1.0.dat.xvg* corresponds to the reweighting PMF.
- *pmf-Phi.dat-noweight-disc1.0.dat.xvg* represents the non-reweighted PMF.

You can download these two files to your laptop and use **xmgrace**, **Excel**, or **OriginPro** to create time-course plot. Below is one generated from *OriginPro* (**Fig 3**).



**Figure 3.** 1D Potential of Mean Force (PMF) profile regarding the distance between 1-butanol and  $\beta$ -cyclodextrin from 100 ns LiGaMD3 simulation.

### 3.3 2D Potential of Mean Force Calculation

In this section, we will use the distance between host and guest (i.e., *host-guest-dist.dat*) and the radius of gyration (Rg) of the host (from *RoG.dat*). The scripts *do-reweighted-2d.sh*, *reweighted-2d.sh*, and *Pyreweighting-2D.py* are required for performing these calculations.

**Following are instructions to run the 2D reweighting job on the UNC cloud server.** Make sure that following required files are available in the ~/LiGaMD3/analysis directory:

- *host-guest-dist.dat*
- *RoG.dat*
- *gamd-all.log*
- *run-2d-pmf.pbs*
- *do-reweighted-2d.sh*
- *PyReweighting-2D.py*

We will use SLURM to submit the 2D reweighting jobs, rather than executing them on the login node. The script named *run-2d-pmf.pbs* has been prepared for job submission. Below is the command used to submit the job:

```
cd ~/ligamd3/analysis
sbatch run-2d-pmf.pbs
squeue
```

A copy of script *run-2d-pmf.pbs* is shown below.

```
#!/bin/sh
#SBATCH --job-name=pmf-2d
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=00:20:00
#SBATCH --output=log.%x.%j

source /shared/apps/software/amber24/amber.sh
cd ~/LiGaMD3/analysis
chmod +x do-reweighted-2d.sh
./do-reweighted-2d.sh host-guest.dat RoG.dat gamd-all.log
```

**The following instructions are to run the 2D reweighting job using your own workstation.** Make sure that the following required files are available in the ~/LiGaMD3/analysis directory:

- *host-guest-dist.dat*
- *RoG.dat*
- *gamd-all.log*
- *run-2d-pmf.sh*
- *do-reweighted-2d.sh*
- *PyReweighting-2D.py*

The script *do-reweighted-2d.sh* was used to perform the calculation.

```
cd ~/LiGaMD3/analysis
chmod +x run-2d-pmf.sh
./run-2d-pmf.sh
```

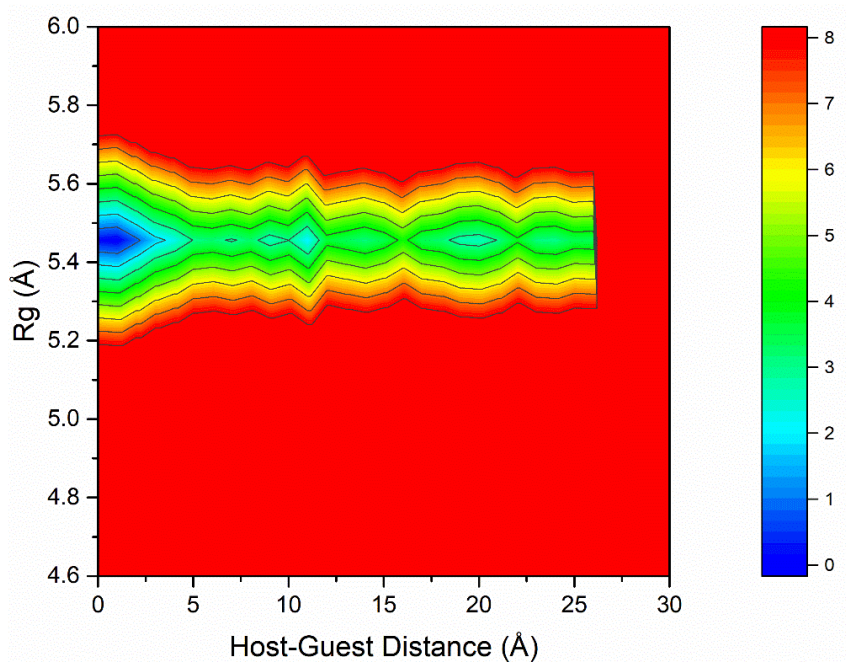
A copy of *run-2d-pmf.sh* is shown below:

```
#!/bin/sh
cd ~/LiGaMD3/analysis
chmod +x do-reweighted-2d.sh
./do-reweighted-2d.sh host-guest.dat RoG.dat gamd-all.log
```

All the PMF values are saved under the **pmf-2d-binx1.0-biny1.0-cutoff500** folder.

- pmf-2D-c2-Phi\_Psi-reweight-discx1.0-discy1.0.xvg corresponding to reweighted 2D PMF.
- pmf-2D-Phi\_Psi-noweight-discx1.0-discy1.0.xvg corresponding to non-reweighted 2D PMFs.

You can download the pmf-2D-c2-Phi\_Psi-reweight-discx1.0-discy1.0.xvg file to your laptop and use **OriginPro** to create plot, as shown below (**Fig 4**):



**Figure 4.** 2D PMF profile regarding the distance between the 1-butanol and  $\beta$ -cyclodextrin, and the radius of gyration ( $R_g$ ) of cyclodextrin from a 100 ns LiGaMD3 simulation.

### 3.3 3D Potential of Mean Force Calculation

The vector calculated with *cpptraj* was used for the 3D reweighting to predict the ligand binding free energy. Details about the 3D reweighting approach can be found in Ref. 3.

**The following instructions outline how to run the 3D reweighting job on the UNC cloud server.** Make sure that the following required files are available in the `~/LiGaMD3/analysis` directory:

- *host-guest-vector.dat*
- *gamd-all.log*
- *run-3d-pmf.pbs*,
- *do-reweighted-3d.sh*
- *PyReweighting-3D.py*

We will use SLURM to submit the 3D reweighting jobs, rather than executing them on the login node. The script named *run-3d-pmf.pbs* has been prepared for job submission.

```
cd ~/ligamd3/analysis
cp /shared/data/LiGaMD3/analysis/run-3d-pmf.pbs ./
sbatch run-3d-pmf.pbs
```

A copy of script *run-3d-pmf.pbs* is shown below:

```
#!/bin/sh
#SBATCH --job-name=pmf-3d
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --time=00:20:00
#SBATCH --output=log.%x.%j

source /shared/apps/software/amber24/amber.sh

cd ~/LiGaMD3/analysis
chmod +x do-reweighted-3d.sh
./do-reweighted-3d.sh host-guest-vector.dat gamd-all.log 9.5 10.0
```

**The following instructions outline how to run the 3D reweighting job using your own workstation.** Make sure that the following required files are available in the `~/LiGaMD3/analysis` directory:

- *host-guest-vector.dat*
- *gamd-all.log*
- *run-3d-pmf.sh*,

- *do-reweighted-3d.sh*
- *PyReweighting-3D.py*

The script named *run-3d-pmf.sh* has been prepared for job submission. Below is the command used for this purpose:

```
cd ~/ligamd3/analysis
chmod +x run-3d-pmf.sh
./run-3d-pmf.sh
```

A copy of *run-3d-pmf.sh* is shown below:

```
#!/bin/sh
cd ~/LiGaMD3/analysis
chmod +x do-reweighted-3d.sh
./do-reweighted-3d.sh host-guest-vector.dat gamd-all.log 9.5 10.0
```

The ligand binding free energy, calculated using the screening with  $dG = -25.785$  kcal/mol through AMD-weighted CE reweighting from the log file or the screen. However, note that the current simulations are not long enough to achieve fully converged predictions.

#### 4. QUESTIONS

If you encounter any issues or have further questions, please post them to GaMD mailing list: [gamd-discuss@lists.sourceforge.net](mailto:gamd-discuss@lists.sourceforge.net) after subscription at <https://sourceforge.net/projects/gamd/lists/gamd-discuss>. For updates and latest tutorials of GaMD, visit the GaMD website: <https://www.med.unc.edu/pharm/miaolab/resources/gamd/>.

#### 5. REFERENCES

- (1) Wang, J.; Miao, Y.\* Ligand Gaussian Accelerated Molecular Dynamics 3 (LiGaMD3): Improved Calculations of Binding Thermodynamics and Kinetics of Both Small Molecules and Flexible Peptides. *Journal of Chemical Theory and Computation* **2024**, 20 (14), 5829-5841.
- (2) Miao, Y.\*; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *Journal of Chemical Theory and Computation* 2014, 10 (7), 2677-2689.

- (3) Miao, Y.\*; Bhattarai, A.; Wang, J. Ligand Gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *Journal of Chemical Theory and Computation* 2020, 16 (9), 5526-5547.