GROMACSを使用した
温度レプリカ交換分子動力学法(T-REMD)
実習

独立行政法人理化学研究所
HPCI計算生命科学推進プログラム

チーム員　波内 良樹

yoshiki.namiuchi@riken.jp

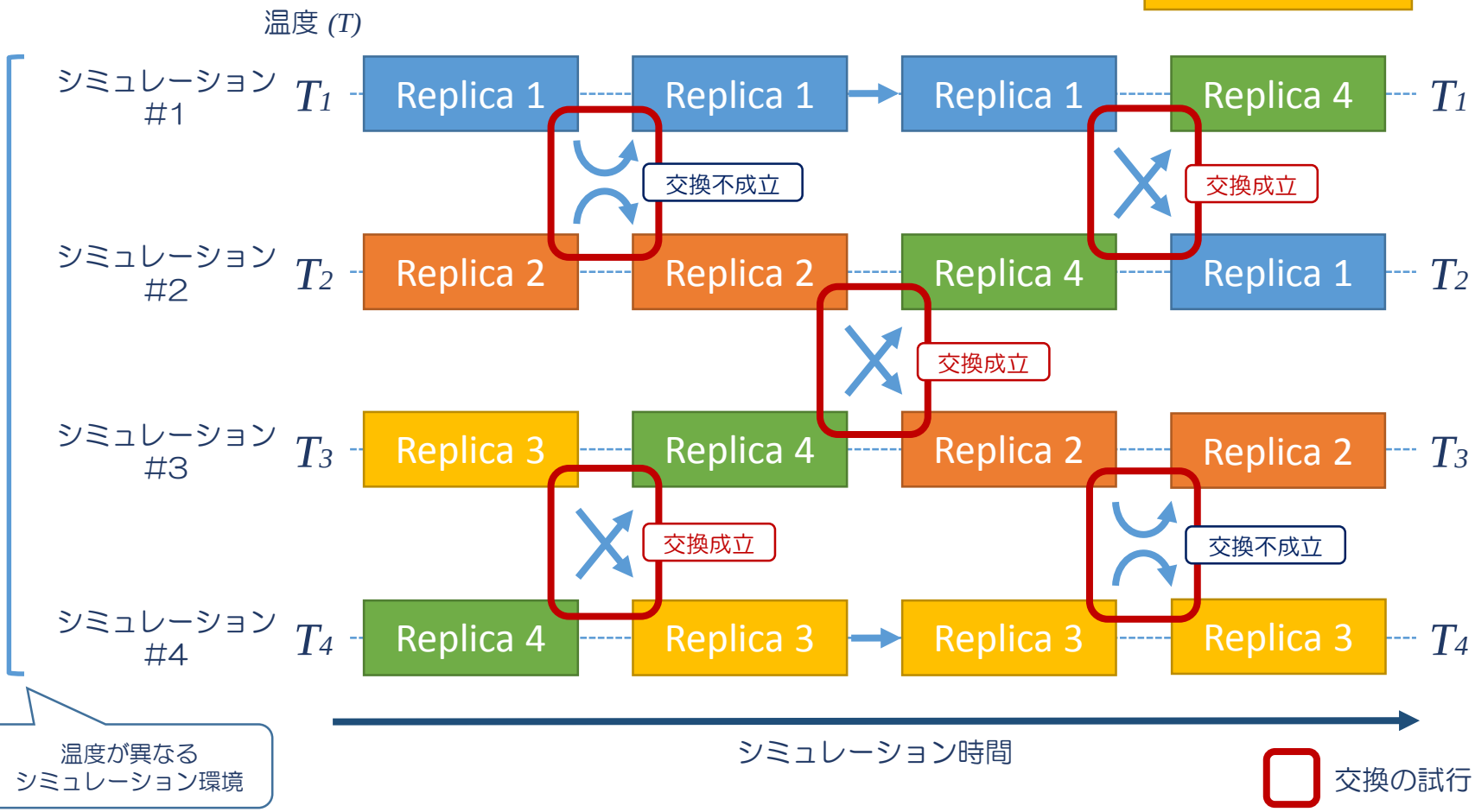
GROMACSによるT-REMDシミュレーション

GROMACSの温度レプリカ交換分子動力学法
(T-REMD)シミュレーション

温度レプリカ交換分子動力学法の実装方法

- 温度を交換
- 構造を交換

GROMACSはこちら
(構造を交換)



レプリカ毎の温度の決定

期待する交換確率(0.2~0.3)になるように、短いT-REMDシミュレーションを繰り返して温度を調整する。

Temperature generator for REMD-simulations

http://folding.bmc.uu.se/remd/

温度の目安に便利

入力

Exchange probability:	0.2
Lower temperature limit:	298
Upper temperature limit:	302
Number of water molecules:	10824
Number of protein atoms:	1960

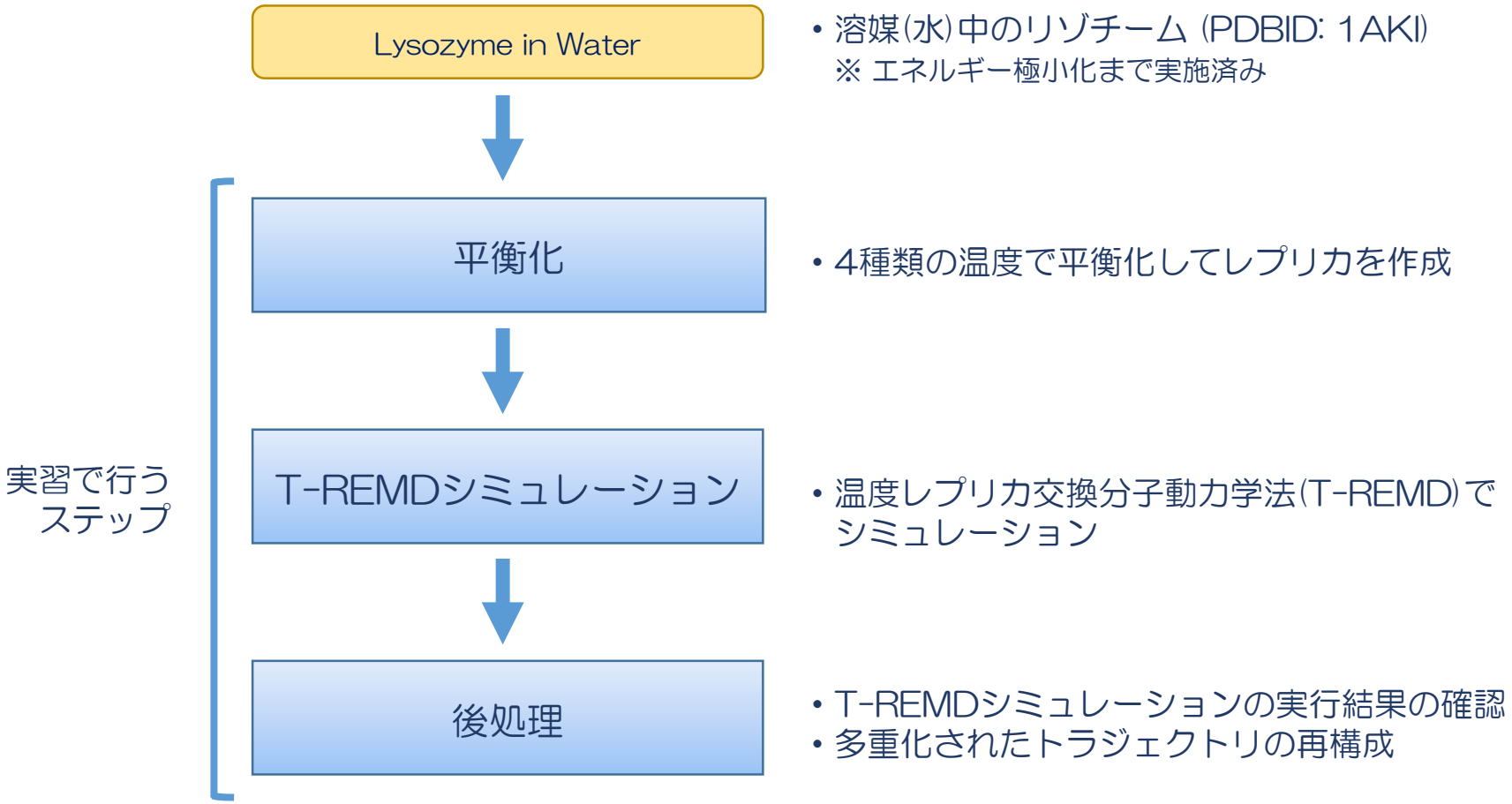
タンパク質の原子数、水分子の数、温度範囲、交換確率を入力すると、レプリカ毎の温度を生成してくれる。

出力

For your scripting pleasures we also give the temperatures below as one long comma-separated string. Enjoy.

298.00, 299.57, 301.14, 302.72 ← 4個の温度を生成

実習のながれ



※ 本実習では、平衡化およびT-REMDシミュレーションはNVTサンプリングで行っています。

SCLS計算機システムのGROMACS

実習で使用するGROMACS

GROMACS 4.6.2	倍精度、MPI、OpenMP ※ GROMACSのコマンド名に、サフィックス「_mpi_d」が付きます。 ※ すべてのコマンドで並列処理が可能な訳ではありません。 （grompppは逐次、mdrunはMPI並列、等々）

GROMACSを使用するにあたって

ジョブスクリプトに以下の記述を追加（環境変数の設定）

```
module load Gromacs/4.6.2

export FLIB_FASTOMP=FALSE
```

※ GROMACSコマンド実行時に、以下のメッセージが表示されます。

```
jwe1050i-w The hardware barrier couldn't be used and continues processing using the software barrier.
```

SCLS計算機システム(FX10)に実装されているハードウェアバリア機能が利用できない場合に表示されます。
(FLIB_FASTOMP=FALSEに設定しているため)　動作に問題はありません。

実習の準備

```
[user1@scls ~]$ cp -r ~namiuchi/lec/lec_gmx_remd .

[user1@scls ~]$ cd lec_gmx_remd

[user1@scls lec_gmx_remd]$ ls -lF
合計 24
drwxr-xr-x 2 user1 group1 4096  6月 24 16:47 2014 pre/
drwxr-xr-x 2 user1 group1 4096  6月 24 16:47 2014 script/
drwxr-xr-x 3 user1 group1 4096  6月 24 16:47 2014 sim0/
drwxr-xr-x 3 user1 group1 4096  6月 24 16:47 2014 sim1/
drwxr-xr-x 3 user1 group1 4096  6月 24 16:47 2014 sim2/
drwxr-xr-x 3 user1 group1 4096  6月 24 16:47 2014 sim3/
[user1@scls lec_gmx_remd]$
```

実習用環境を格納したディレクトリlec_gmx_remdを自身のホームディレクトリ配下にコピー

カレントディレクトリを、コピーしたlec_gromacs_remdディレクトリに変更

lec_gromacs_remdディレクトリの内容を表示

トポロジーファイルやエネルギー極小化済み構造ファイルを格納しているディレクトリ

ジョブスクリプトを格納しているディレクトリ

4種類の温度のシミュレーション環境を格納するディレクトリ

平衡化

各シミュレーション環境のディレクトリに、それぞれの温度で平衡化したレプリカを作成

平衡化のパラメータファイル

```
[user1@scls lec_gmx_remd]$ less sim0/equil.mdp

; Temperature coupling
tcoupl          = V-rescale     ; modified Berendsen thermostat
; pressure coupling
pcoupl          = no            ; no pressure coupling in NVT
gen-vel         = yes

[user1@scls lec_gmx_remd]$
```

シミュレーション#0ディレクトリ(sim0)の
平衡化用パラメータファイル(equil.mdp) を参照

一部抜粋

レプリカ毎の温度設定

```
[user1@scls lec_gmx_remd]$ grep ref-t sim*/equil.mdp

sim0/equil.mdp:ref-t   = 298.00  ; reference temperature, one for each group, in K
sim1/equil.mdp:ref-t   = 299.57  ; reference temperature, one for each group, in K
sim2/equil.mdp:ref-t   = 301.14  ; reference temperature, one for each group, in K
sim3/equil.mdp:ref-t   = 302.72  ; reference temperature, one for each group, in K

[user1@scls lec_gmx_remd]$
```

各シミュレーションディレクトリ(sim*)の
平衡化用パラメータファイル(equil.mdp)の
温度設定パラメータを参照

それぞれの温度設定パラメータの値

平衡化 [grompp] (1)

ジョブスクリプト

```
[user1@scls lec_gmx_remd]$ cat script/grompp_equil.sh
#!/bin/sh
#------ pjsub options ------#
#PJM -L "rscgrp=small"
#PJM -L "node=1"
#PJM -L "elapse=10:00"
#PJM -j
#------ Program Execution --------#
module load Gromacs/4.6.2

export FLIB_FASTOMP=FALSE

for dir in sim[0123]
do
  cd $dir
  grompp_mpi_d -f equil.mdp -c ../pre/em.gro -p ../pre/topol.top -o equil.tpr
  cd ..
done

[user1@scls lec_gmx_remd]$
```

リソースグループ「small」
使用ノード数「1」
最大経過時間「10分」
標準エラー出力を標準出力に向ける

GROMACS環境変数の設定

各シミュレーションディレクトリで
順にgromppを実行

平衡化 [grompp] (2)

ジョブの投入と状態確認

```
[user1@scls lec_gmx_remd]$ pjsub script/grompp_equil.sh
[INFO] PJM 0000 pjsub Job 15151 submitted.

[user1@scls lec_gmx_remd]$ pjstat

  ACCEPT QUEUED  STGIN  READY RUNING RUNOUT STGOUT   HOLD  ERROR   TOTAL
       0      0      0      0      1      0      0      0      0       1
s      0      0      0      0      1      0      0      0      0       1

JOB_ID     JOB_NAME   MD ST  USER     START_DATE      ELAPSE_LIM NODE_REQUIRE
15151      grompp_equ NM RUN user1    04/22 10:33:54  0000:10:00 1

[user1@scls lec_gmx_remd]$
```

ジョブ実行結果

```
[user1@scls lec_gmx_remd]$ less grompp_equil.sh.oXXXX
             :
[user1@scls lec_gmx_remd]$ ls -lF sim*/equil.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 10:33 2014 sim0/equil.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 10:33 2014 sim1/equil.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 10:34 2014 sim2/equil.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 10:34 2014 sim3/equil.tpr

[user1@scls lec_gmx_remd]$
```

ジョブの実行結果の確認

生成された各レプリカの
平衡化用実行入力ファイル

平衡化 [mdrun] (1)

ジョブスクリプト

```
[user1@scls lec_gromacs]$ cat script/mdrun_equil.sh
#!/bin/sh
#------ pjsub options ------#
#PJM -L "rscgrp=small"
#PJM -L "node=12"
#PJM --mpi "proc=192"
#PJM --mpi "rank-map-bychip"
#PJM -L "elapse=10:00"
#PJM -j
#------ Program Execution ------#
module load Gromacs/4.6.2

export FLIB_FASTOMP=FALSE

mpiexec mdrun_mpi_d -v -multidir sim[0123] -deffnm equil
```

リソースグループ「small」
使用ノード数「8」
プロセス数「128」
各ノードへのプロセスの割り振りの指定
最大経過時間「10分」
標準エラー出力を標準出力に向ける

GROMACS環境変数の設定

mdrunの実行
※講習会向けにシミュレーション時間を短く設定(10ps)

【 -multidir オプション 】... 複数のディレクトリに格納された系を並列処理

sim0, sim1, sim2, sim3の4個のディレクトリに格納された実行入力ファイルを、
それぞれ3ノード(48プロセス)使用して並列に処理

平衡化 [mdrun] (2)

ジョブの投入と状態確認

```
[user1@scls lec_gmx_remd]$ pjsub script/mdrun_equil.sh
[INFO] PJM 0000 pjsub Job 15152 submitted.

[user1@scls lec_gmx_remd]$ pjstat

  ACCEPT QUEUED  STGIN  READY RUNING RUNOUT STGOUT   HOLD  ERROR   TOTAL
       0      0      0      0      1      0      0      0      0       1
s      0      0      0      0      1      0      0      0      0       1

JOB_ID     JOB_NAME   MD ST  USER     START_DATE      ELAPSE_LIM NODE_REQUIRE
15152      mdrun_equl NM RUN user1    04/22 11:00:16  0000:10:00 8

[user1@scls lec_gmx_remd]$
```

ジョブ実行結果

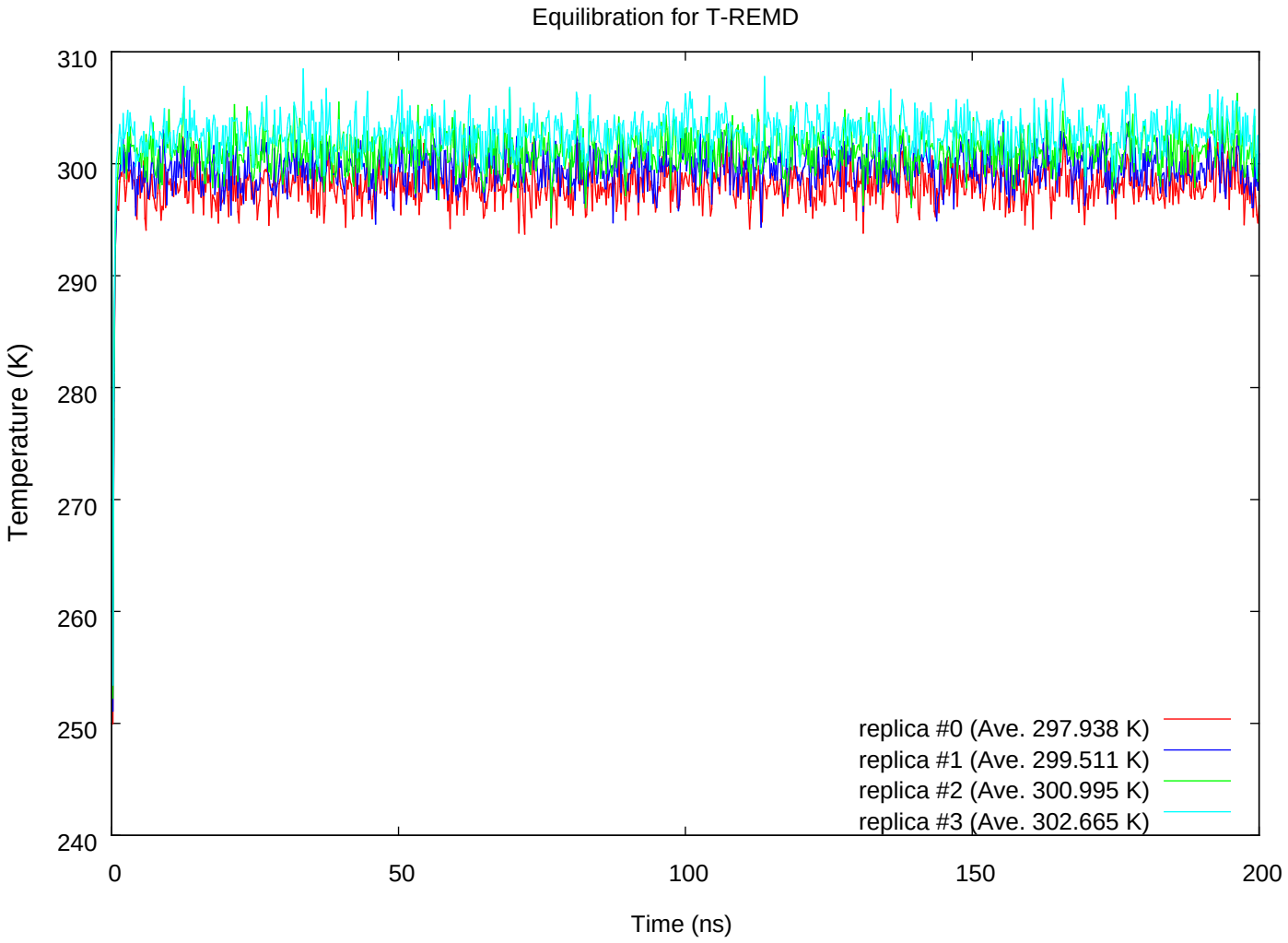
```
[user1@scls lec_gmx_remd]$ less mdrun_equil.sh.oXXXX
              :
[user1@scls lec_gmx_remd]$ less sim0/equil.log
              :
[user1@scls lec_gmx_remd]$ less sim3/equil.log
              :
[user1@scls lec_gmx_remd]$
```

ジョブの実行結果の確認

レプリカ0の平衡化処理のログファイル

レプリカ3の平衡化処理のログファイル

平衡化［結果］



※ 実習と同じパラメータで平衡化を200ps実行した結果

T-REMD シミュレーション

平衡化とのパラメータの違い

```
[user1@scls lec_gmx_remd]$ diff sim0/equil.mdp sim0/sim.mdp
54c54
< gen-vel           = yes
---
> gen-vel           = no
63c63
< continuation      = no
---
> continuation      = yes
[user1@scls lec_gmx_remd]$
```

速度は生成しない

実行環境を引き継ぐ

シミュレーション間のパラメータの違い

```
[user1@scls lec_gmx_remd]$ diff sim0/sim.mdp sim1/sim.mdp
49c49
< ref-t             = 298.00  ; reference temperature, one for each group, in K
---
> ref-t             = 299.57  ; reference temperature, one for each group, in K
55c55
< gen-temp          = 298.00  ; temperature for Maxwell distribution
---
> gen-temp          = 299.57  ; temperature for Maxwell distribution
[user1@scls lec_gmx_remd]$
```

違いは温度のみ

T-REMD シミュレーション [grompp] (1)

ジョブスクリプト

```
[user1@scls lec_gmx_remd]$ cat script/grompp_sim.sh
#!/bin/sh
#------ pjsub options ------#
#PJM -L "rscgrp=small"
#PJM -L "node=1"
#PJM -L "elapse=10:00"
#PJM -j
#------ Program Execution --------#
module load Gromacs/4.6.2

export FLIB_FASTOMP=FALSE

for dir in sim[0123]
do
  cd $dir
  grompp_mpi_d -f sim.mdp -c _equil_200ps/equil.gro -t _equil_200ps/equil.cpt ¥
               -p ../pre/topol.top -o sim.tpr
  cd ..
done

[user1@scls lec_gmx_remd]$
```

リソースグループ「small」
使用ノード数「1」
最大経過時間「10分」
標準エラー出力を標準出力に向ける

GROMACS環境変数の設定

※平衡化の結果は、
あらかじめ用意したもの(200ps実行)を使用

各シミュレーションディレクトリで、
順にgromppを実行

T-REMD シミュレーション [grompp] (2)

ジョブの投入と状態確認

```
[user1@scls lec_gmx_remd]$ pjsub script/grompp_sim.sh
[INFO] PJM 0000 pjsub Job 15153 submitted.

[user1@scls lec_gmx_remd]$ pjstat

  ACCEPT QUEUED  STGIN  READY RUNING RUNOUT STGOUT   HOLD  ERROR   TOTAL
       0      0      0      0      1      0      0      0      0       1
s      0      0      0      0      1      0      0      0      0       1

JOB_ID     JOB_NAME   MD ST  USER     START_DATE      ELAPSE_LIM NODE_REQUIRE
15153      grompp_sim NM RUN user1    04/22 11:43:28  0000:10:00 1

[user1@scls lec_gmx_remd]$
```

ジョブ実行結果

```
[user1@scls lec_gmx_remd]$ less grompp_sim.sh.oXXXX
             :
[user1@scls lec_gmx_remd]$ ls -lF sim*/sim.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 11:43 2014 sim0/sim.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 11:43 2014 sim1/sim.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 11:43 2014 sim2/sim.tpr
-rw-r--r-- 1 user1 group1 2157740  4月 22 11:43 2014 sim3/sim.tpr

[user1@scls lec_gmx_remd]$
```

ジョブの実行結果の確認

生成された
各シミュレーションの
実行入力ファイル

T-REMD シミュレーション [mdrun] (1)

ジョブスクリプト

```
[user1@scls lec_gromacs]$ cat script/mdrun_sim.sh
#!/bin/sh
#------ pjsub options ------#
#PJM -L "rscgrp=small"
#PJM -L "node=12"
#PJM --mpi "proc=192"
#PJM --mpi "rank-map-bychip"
#PJM -L "elapse=10:00"
#PJM -j
#------ Program Execution ------#
module load Gromacs/4.6.2

export FLIB_FASTOMP=FALSE

mpiexec mdrun_mpi_d -v -multidir sim[0123] -replex 100 -deffnm md
```

リソースグループ「small」
使用ノード数「8」
プロセス数「128」
各ノードへのプロセスの割り振りの指定
最大経過時間「10分」
標準エラー出力を標準出力に向ける

GROMACS環境変数の設定

mdrunの実行
※講習会向けにシミュレーション時間を短く(10ps)設定

【 -replex オプション 】… T-REMD実行

-multidirで指定されたsim0, sim1, sim2, sim3のシミュレーション環境に対して、100ステップ毎にレプリカ交換を試行する。1ステップの時間はパラメータファイルで設定 (この実習では2fs)

(100ステップは講習会向け、通常は500～1000ステップ程度)

T-REMD シミュレーション [mdrun] (2)

ジョブの投入と状態確認

```
[user1@scls lec_gmx_remd]$ pjsub script/mdrun_sim.sh
[INFO] PJM 0000 pjsub Job 15154 submitted.

[user1@scls lec_gmx_remd]$ pjstat

  ACCEPT QUEUED  STGIN  READY RUNING RUNOUT STGOUT   HOLD  ERROR   TOTAL
       0      0      0      0      1      0      0      0      0       1
s      0      0      0      0      1      0      0      0      0       1

JOB_ID     JOB_NAME   MD ST  USER     START_DATE      ELAPSE_LIM NODE_REQUIRE
15154      mdrun_sim. NM RUN user1    04/22 11:53:46  0000:10:00 8

[user1@scls lec_gmx_remd]$
```

ジョブ実行結果

```
[user1@scls lec_gmx_remd]$ less mdrun_sim.sh.oXXXX
              :
[user1@scls lec_gmx_remd]$
```

ジョブの実行結果の確認

T-REMD シミュレーションの実行結果の確認 (1)

各シミュレーション環境のディレクトリに、それぞれの実行ログファイルが出力されている。

```
[user1@scls lec_gmx_remd]$ less sim1/sim.log

             ⋮

Replica exchange in temperature
 298.0 299.6 301.1 302.7

Replica exchange interval: 100

Replica random seed: 25511
Replica exchange information below: x=exchange, pr=probability
Initial temperature: 300.779 K

Started mdrun on node 0 Thu Jul  3 16:12:23 2014

             ⋮

Replica exchange at step 100 time 0.2
Repl 0 <-> 1  dE_term =  3.926e+00 (kT)
Repl ex  0    1    2 x  3
Repl pr   .02       1.0


Replica exchange at step 200 time 0.4
Repl 1 <-> 2  dE_term =  2.432e+00 (kT)
Repl ex  0    1    2    3
Repl pr        .09

             ⋮
```

シミュレーション#1の実行ログファイルを表示

REMDシミュレーションの情報

交換を試行する対象となる隣接シミュレーション環境 (このステップではシミュレーション#0) とのポテンシャルエネルギーの比較

このステップ全体では、#0と#1、#2と#3の間で交換を試行し、#2と#3の間で交換成立 (x印)

このステップの交換試行の対象はシミュレーション#2

このステップ全体では、#1と#2の間で交換を試行し、交換は不成立

T-REMD シミュレーションの実行結果の確認 (2)

```
             ⋮
Replica exchange at step 300 time 0.6
Repl 0 <-> 1  dE_term =  4.279e+00 (kT)
Repl ex  0    1    2    3
Repl pr   .01       .04

             ⋮

Replica exchange at step 2600 time 5.2
Repl 1 <-> 2  dE_term =  1.796e+00 (kT)
Repl ex  0    1 x  2    3
Repl pr        .17

             ⋮

Replica exchange at step 4800 time 9.6
Repl 1 <-> 2  dE_term = -2.206e+00 (kT)
Repl ex  0    1 x  2    3
Repl pr        1.0

             ⋮

Replica exchange statistics
Repl  49 attempts, 25 odd, 24 even
Repl  average probabilities:
Repl     0    1    2    3
Repl      .33  .23  .21
Repl  number of exchanges:
Repl     0    1    2    3
Repl        9    6    6
Repl  average number of exchanges:
Repl     0    1    2    3
Repl      .36  .25  .24

             ⋮
```

ポテンシャルエネルギーは、#0より1#の方が約4.3 kT高い
交換確率は1%で交換は不成立

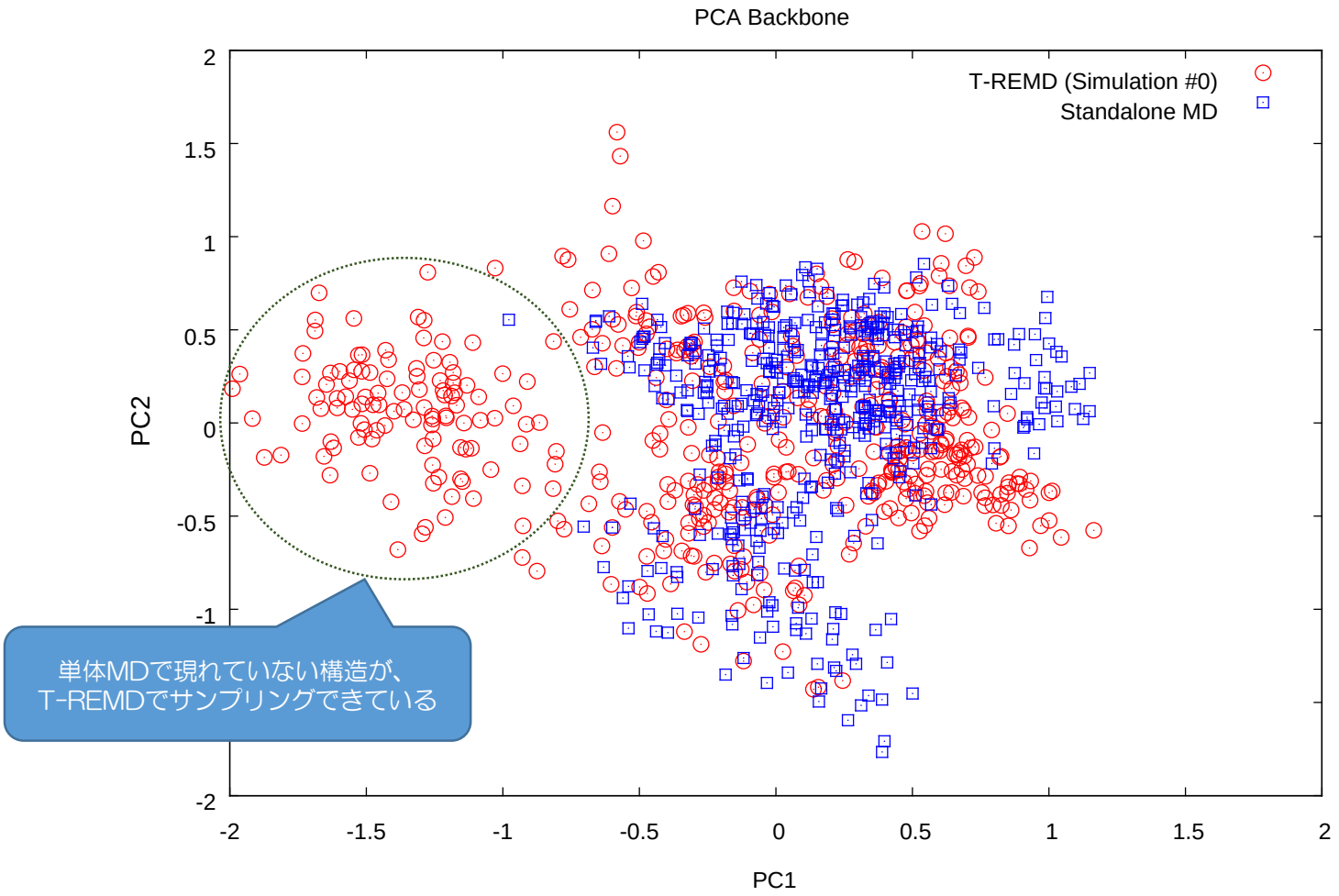
ポテンシャルエネルギーは、#1より#2の方が約1.8 kT高い
交換確率は17%で交換は成立

ポテンシャルエネルギーは、#1より#2の方が約2.2 kT低い
交換確率は100%で交換は成立

REMDシミュレーションのサマリ
(平均交換確率、交換回数、平均交換回数)

T-REMDシミュレーションの実行結果の確認 (3)

T-REMDのシミュレーション#0と、同じ初期構造(温度)を単体のMDで実行した結果のPCA



※ 実習と同じパラメータで、交換試行ステップを500でT-REMDを1ns実行したトラジェクトリを使用

多重化されたトラジェクトリの再構成 (1)

GROMACSのT-REMDシミュレーションは、
構造を交換

↓

各シミュレーションで出力されたトラジェクトリは、
温度について連続的

↓

構造について連続的なトラジェクトリは、
トラジェクトリファイルを再構成して生成

多重化されたトラジェクトリの再構成 (2)

交換情報ファイルの生成

いずれか一つのシミュレーション環境のログファイルから、交換情報を抽出してファイルに出力

```
[user1@scls lec_gmx_remd]$ pjsub --interact
[INFO] PJM 0000 pjsub Job 15155 submitted.
[INFO] PJM 0081 .connected.
[INFO] PJM 0082 pjsub Interactive job 15155 started.
[user1@a01-008 lec_gmx_remd]$ module load Gromacs/4.6.2


[user1@a01-008 lec_gmx_remd]$ demux.pl sim0/sim.log
-------------------------------------------------------------
Going to read a file containing the exchange information from
your mdrun log file (sim0/sim.log).
This will produce a file (replica_index.xvg) suitable for
demultiplexing your trajectories using trjcat,
as well as a replica temperature file (replica_temp.xvg).
Each entry in the log file will be copied 0 times.
-------------------------------------------------------------
There are 4 replicas.
Finished writing replica_index.xvg and replica_temp.xvg with 50 lines

[user1@a01-008 lec_gmx_remd]$
```

会話型ジョブの投入

GROMACS環境変数の設定

交換情報を抽出するPerlスクリプト
(ここではシミュレーション#0のログを使用)

ログファイルから、4レプリカの交換情報を50行(初期配置 + 交換試行49回)分
抽出して、replica_index.xvgとreplica_temp.xvgの2つのファイルに出力

多重化されたトラジェクトリの再構成 (3)

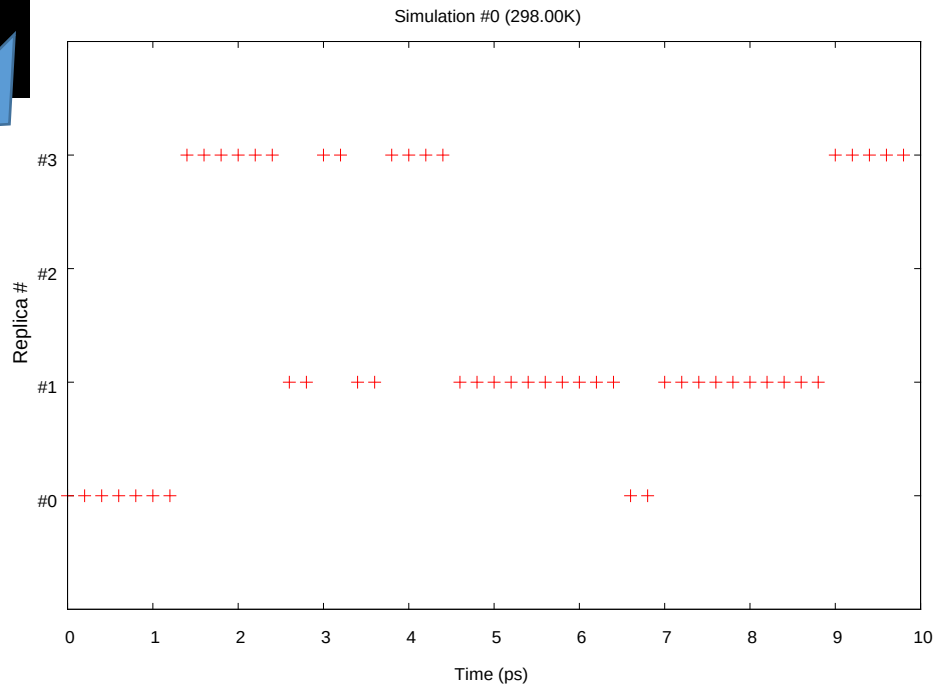
demux.plが生成するファイル(1) - replica_index.xvg

```
[user1@scls lec_gmx_remd]$ less replica_index.xvg
0                       0    1    2    3
0.2                     0    1    3    2
0.4                     0    1    3    2
0.6                     0    1    3    2
0.8                     0    3    1    2
1                       0    3    1    2
1.2                     0    3    1    2
1.4                     3    0    1    2
1.6                     3    0    1    2
                  ⋮

[user1@scls lec_gmx_remd]$
```

カラムはシミュレーション環境 (左から#0, #1, ...)
値はレプリカのインデックス

グラフ化すると...



特定のシミュレーション環境(右図はシミュレーション#0)に遷移したレプリカの様子

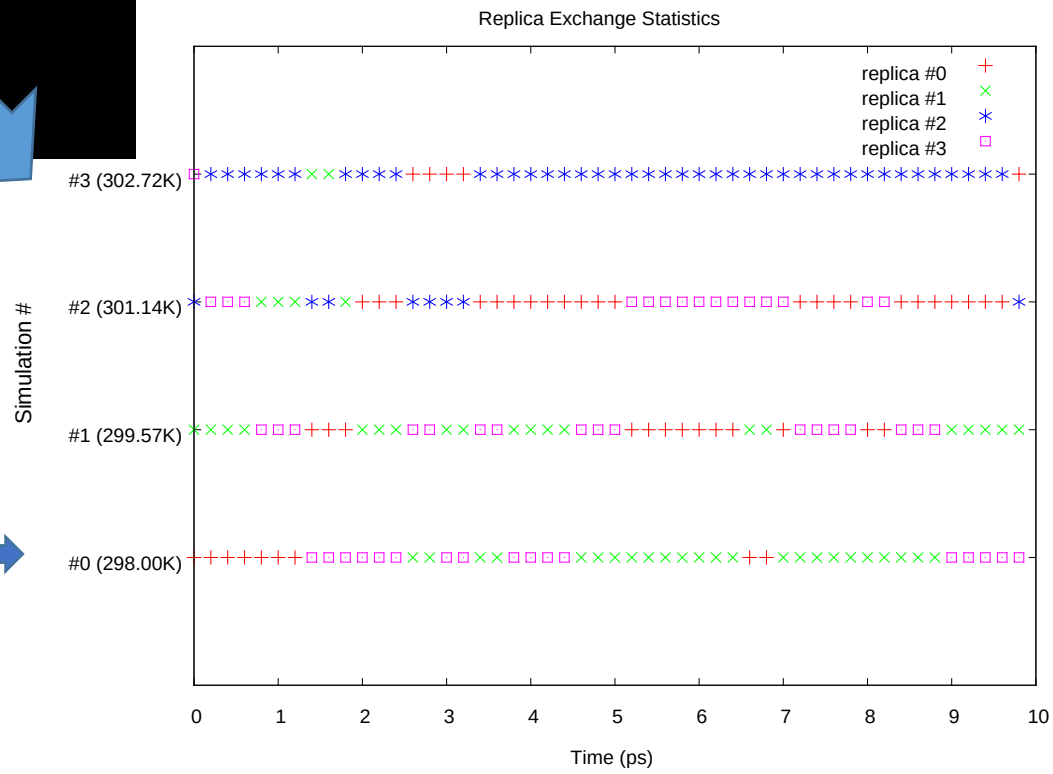
多重化されたトラジェクトリの再構成 (4)

demux.plが生成するファイル(2) - replica_temp.xvg

```
[user1@scls lec_gmx_remd]$ less replica_temp.xvg
0                       0    1    2    3
0.2                     0    1    3    2
0.4                     0    1    3    2
0.6                     0    1    3    2
0.8                     0    2    3    1
1                       0    2    3    1
1.2                     0    2    3    1
1.4                     1    2    3    0
1.6                     1    2    3    0
                  ⋮
[user1@scls lec_gmx_remd]$
```

カラムはレプリカのインデックス
値はシミュレーション環境

グラフ化すると...



シミュレーション環境間を遷移するレプリカの様子

多重化されたトラジェクトリの再構成 (5)

トラジェクトリの再構成

demul.plが生成したreplica_index.xvgを使用して、各シミュレーションのトラジェクトリを再構成

```
[user1@a01-008 lec_gmx_remd]$ trjcat_mpi_d -demux replica_index.xvg -f sim0/sim.xtc
sim1/sim.xtc sim2/sim.xtc sim3/sim.xtc
 :-)  G  R  O  M  A  C  S  (-:

           :
Read 4 sets of 50 points, dt = 0.2

           :
[user1@a01-008 lec_gmx_remd]$ exit
[INFO] PJM 0083 pjsub Interactive job 15155 completed.
[user1@scls lec_gmx_remd]$
```

多重化されたトラジェクトリを再構成するコマンド

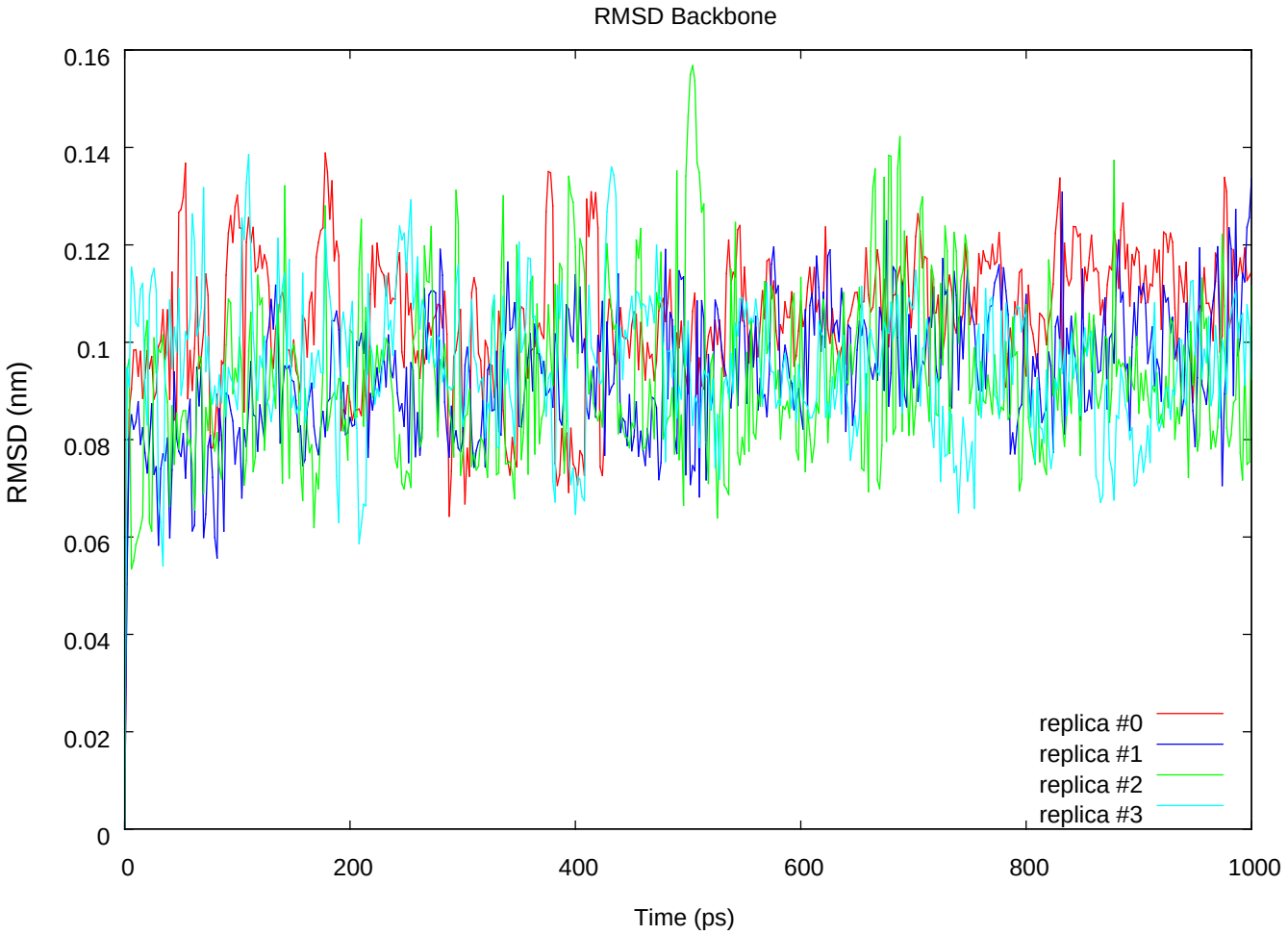
会話型ジョブの終了

再構成されたトラジェクトリファイル

```
[user1@scls lec_gmx_remd]$ ls -lF *.xtc
total 280
-rw-r--r-- 1 user1 group1 743536 Apr 22 14:20 0_trajout.xtc
-rw-r--r-- 1 user1 group1 743736 Apr 22 14:20 1_trajout.xtc
-rw-r--r-- 1 user1 group1 743560 Apr 22 14:20 2_trajout.xtc
-rw-r--r-- 1 user1 group1 743584 Apr 22 14:20 3_trajout.xtc
[user1@scls lec_gmx_remd]$
```

多重化されたトラジェクトリの再構成 (6)

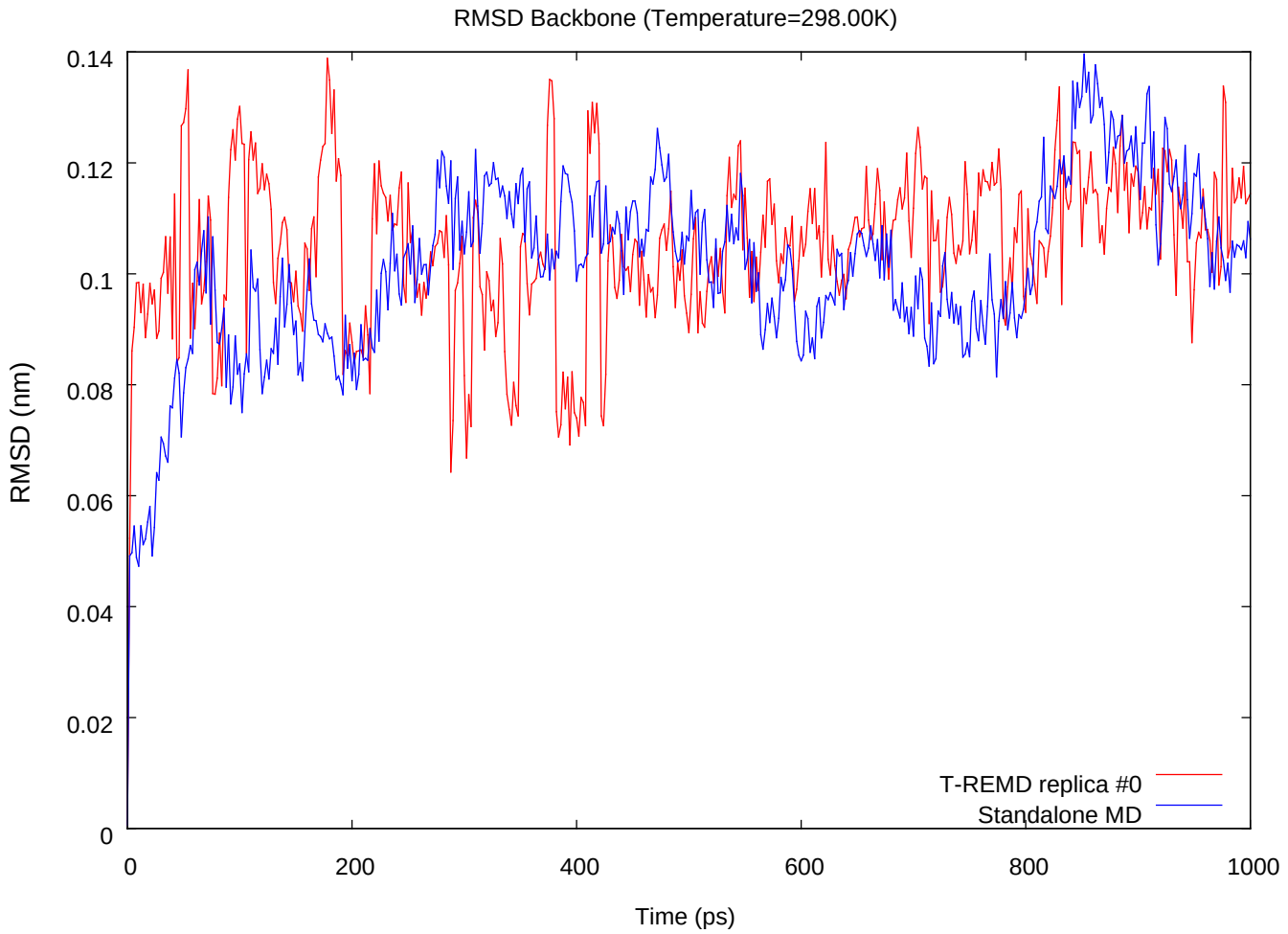
再構成したトラジェクトリから計算した各レプリカのRMSD



※ 実習と同じパラメータで、交換試行ステップを500でT-REMDを1ns実行したトラジェクトリを使用

多重化されたトラジェクトリの再構成 (7)

T-REMDのレプリカ#0と、同じ初期構造(温度)を単体のMDで実行した結果(RMSD)との比較



※ 実習と同じパラメータで、交換試行ステップを500でT-REMDを1ns実行したトラジェクトリを使用

おわりに

○ GROMACSに関する情報、マニュアル、ソースコード入手など

```
http://www.gromacs.org/
```

○ ご意見、ご質問などありましたらお知らせください。

```
mailto: Yoshiki.namiuchi@riken.jp
```

2014年9月30日 (第1.1版)

独立行政法人理化学研究所
HPCI計算生命科学推進プログラム