# 1/13/2021 Week 1 Module 1
# Interactive Exercise for Gibbs sampling HMC

- In this module, you will learn how to run Robosample using its Python interface
- Run a short rigid body dynamics HMC simulation on your desktop machine

# Preparing a script to run Gibbs sampling HMC with Robosample

Github

# Robosample setup

- Robosample is a molecular simulation program that uses algorithms designed primarily for robot mechanics
  - ideal for implementing reduced coordinates constrained dynamics

- Initial setup:
  - conda install -c conda-forge mdtraj
  - cd ~/Apps/Robosample
  - git pull
  - add lines indicated on the right to ~/.bashrc
  - source ~/.bashrc

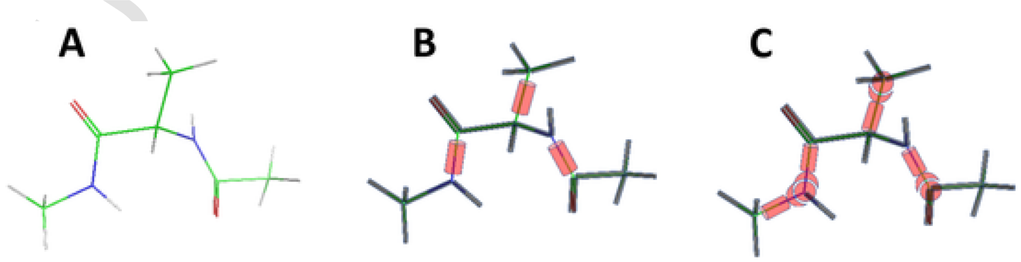
```
if [ -z "$PYTHONPATH" ]; then
            export PYTHONPATH=$PYTHONPATH:"/home/chemuser/Apps/Robosample/tools/"
else
            export PYTHONPATH="/home/chemuser/Apps/Robosample/tools/"
fi

export ROBOSAMPLEDIR="/home/chemuser/Apps/Robosample/"
```

# Why write a script?

- Molecular dynamics simulations require a lot of information about
  - Input data
    - coordinates
    - topology
      - which atoms are included in energy terms
      - parameters for functions in energy terms
  - System description
    - periodicity
    - constraints
  - Integrators
    - algorithms to propagate forward in time
    - adjust box size
    - adjust kinetic energy (temperature)
  - Simulation
    - how long to run
    - how much output data to store
- Robosample can be run from widely-used computer programming languages, python and C++, facilitating extension and combination with other code

# Robosample script

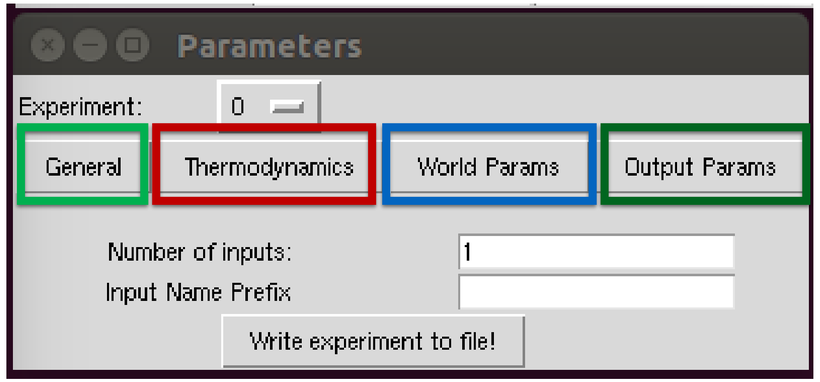
- Robosample simulations require even more information due to Gibbs sampling
  - How many sets of constraints (worlds or blocks)
  - Constraint specifications for every world:
    - what coordinates are frozen
    - what type of joints are used
- This is the script generated by Robosample GUI or Python interface:
- Let's go through what the parameters mean, and I'll give you values for a simple simulation of deca-alanine

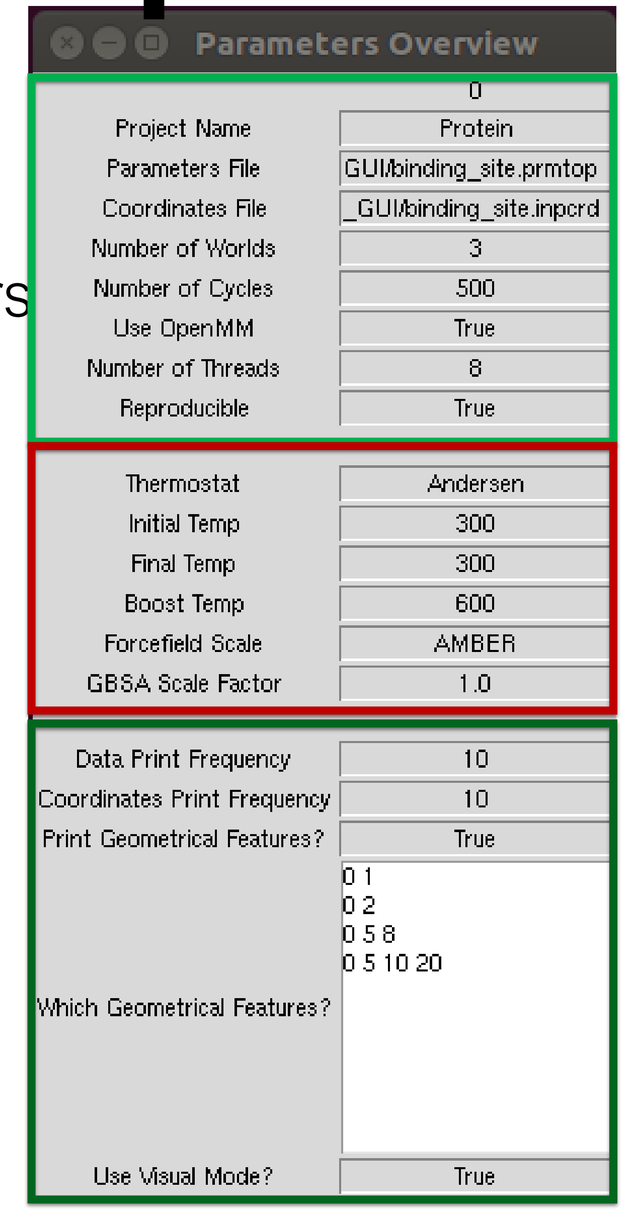


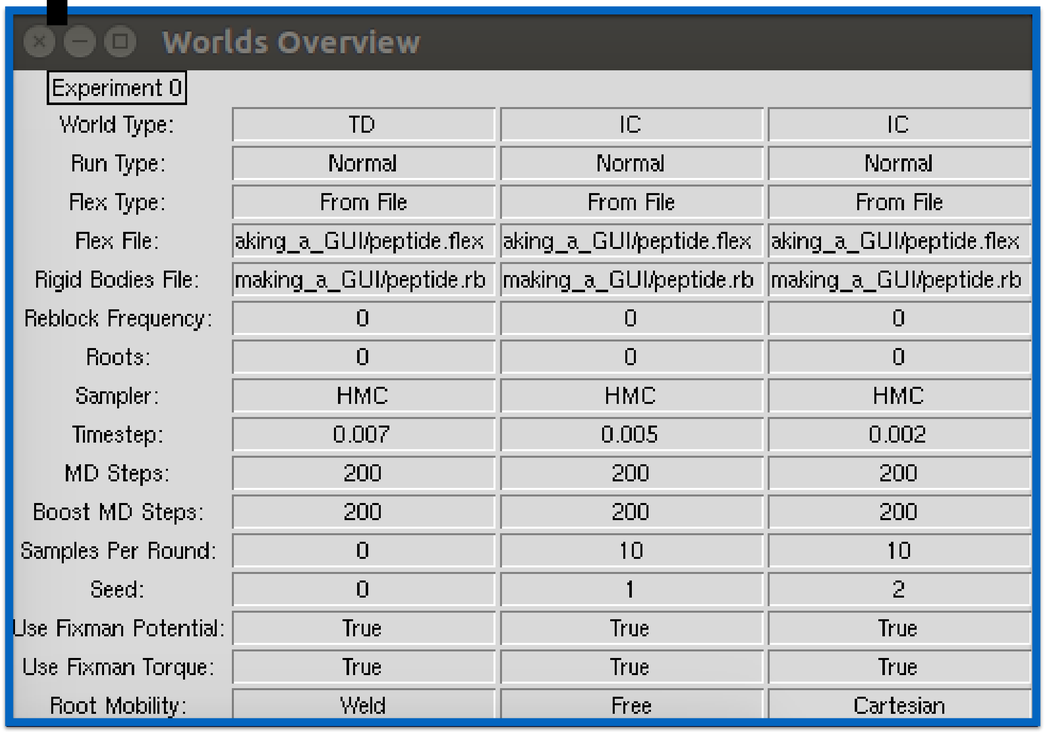

```
# Robosample input
MOLECULES robots/bot0
PRMTOP bot.prmtop bot.prmtop       # Parameter file
INPCRD bot.rst7 bot.rst7     # Coordinate / Restart file
RBFILE bot.rb bot.rb    # Rigid bodies definition file
FLEXFILE bot.all.flex bot.hinges.flex # Flexibility definition file
ROOT_MOBILITY Cartesian Free # Ground to Compound mobilizer
OUTPUT_DIR robots robots

# Simulation
RUN_TYPE Normal Normal # normal HMC or Non-Eq HMC
ROUNDS 50
ROUNDS_TILL_REBLOCK 10 10
RANDOM_WORLD_ORDER FALSE FALSE
WORLDS R0 R1                        # Regimen (IC, TD, MIX, RB, RBMIX)
ROOTS 0 0
SAMPLER HMC HMC
TIMESTEPS 0.001 0.009              # Timesteps to be used with regimens
MDSTEPS 30 30  # Number of MD trial steps
BOOST_MDSTEPS 1 1
SAMPLES_PER_ROUND 1 3  # Number of acc-rej steps within a mixing round
REPRODUCIBLE FALSE FALSE
SEED 999 999

# Thermodynamics
THERMOSTAT Andersen Andersen   # Thermostat
TEMPERATURE_INI  300 300
TEMPERATURE_FIN  300 300
BOOST_TEMPERATURE  1 1      # Boost temperature
FFSCALE AMBER AMBER     # Force field
GBSA 0 0         # GBSA scale factor

# Correction factors
FIXMAN_POTENTIAL TRUE TRUE # Use Fixman potential
FIXMAN_TORQUE TRUE TRUE    # Use Fixman torque

# Output
VISUAL TRUE TRUE             # Use the visualizer
PRINT_FREQ  1 1
WRITEPDBS 1 0    # Write pdbs
GEOMETRY FALSE FALSE         # Calculate geometric features

DISTANCE 1 2
DIHEDRAL 1 2 3 4 1 2 3 4

# Software specs
THREADS 0 0
OPENMM TRUE TRUE
~
~
~
"inp.test" 48L, 1435C
```

# Robosample script builder

- Robosample GUI allows the user to group the parameters by:
  - constraints (blue) and
  - further by type



# Robosample Python interface

- Let's go through what the parameters mean and I'll give you values for a simple simulation of deca-alanine
- Let's create a directory called "Robosample" on the Desktop.
- Copy "ala10" directory from "~/Apps/Robosample/tools/demo/" directory. The directory can also be found on github.
- Let's start editing a file, ex. "simulate.py".

```
from robosample import *

# Load Amber files
prmtop = AmberPrmtopFile("ala10/ligand.prmtop")
inpcrd = AmberInpcrdFile("ala10/ligand.rst7")

# Hardware platform
platform = Platform.getPlatformByName('CPU')

properties={'nofThreads': 2}

# Create a Robosample system by calling createSystem on prmtop
system = prmtop.createSystem(createDirs = False,
        nonbondedMethod = "CutoffPeriodic",
        nonbondedCutoff = 1.44*nanometer,
        constraints = None,
        rigidWater = True,
        implicitSolvent = True,
        soluteDielectric = 1.0,
        solventDielectric = 78.5,
        removeCMMotion = False
)

integrator = HMCIntegrator(300*kelvin,   # Temperature of head bath
                           0.006*picoseconds) # Time step


simulation = Simulation(prmtop.topology, system, integrator, platform, properties)
simulation.reporters.append(PDBReporter('robots/', 10))
simulation.context.setPositions(inpcrd.positions)

# run simulation
simulation.step(5)
```

- For input files enter “ala10/ligand.prmtop” and “ala10/ligand.inpcrd”
- “Platform” describes the version of the code and the hardware it will run on
  - “CUDA” and “OpenCL” are meant for GPUs, which make MD simulations much faster. “CUDA” only works with Nvidia GPUs and “OpenCL” on others
  - “CPU” is a faster version of Reference.
  - Since VirtualBox machine don’t usually have GPUs, let’s use “CPU”
- The properties dictionary indicates the number of threads to be used
- “Forcefield” is the set of parameters and functions that describes the energy of a system. Currently, Robosample only supports Amber functional form.

```
from robosample import *

# Load Amber files
prmtop = AmberPrmtopFile("ala10/ligand.prmtop")
inpcrd = AmberInpcrdFile("ala10/ligand.rst7")

# Hardware platform
platform = Platform.getPlatformByName('CPU')

properties={'nofThreads': 2}
```

- “Nonbonded method” describes how long-range interactions are treated. The more interactions that need to be computed, the slower an MD simulation will be.
  - Cutoffs don’t perform calculations if two particles are beyond a certain distance apart.
  - Ewald methods calculate long-range interaction energies between a system and its periodic images. It assumes the system repeats indefinitely.
  - Since we are using implicit solvent, we don’t need periodicity.
  - Let’s use “CutoffNonPeriodic”.
- “Constraints”
  - force a degree of freedom to be a certain value
- Let’s keep the other “System” parameters as is

```
# Create a Robosample system by calling createSystem on prmtop
system = prmtop.createSystem(createDirs = False,
        nonbondedMethod = "CutoffPeriodic",
        nonbondedCutoff = 1.44*nanometer,
        constraints = None,
        rigidWater = True,
        implicitSolvent = True,
        soluteDielectric = 1.0,
        solventDielectric = 78.5,
        removeCMMotion = False
)
```

```
integrator = HMCIntegrator(300*kelvin,   # Temperature of
                           0.006*picoseconds) # Time step
```

- “Integrator” is the algorithm that goes from one configuration to the next
  - Verlet is completely deterministic
  - Langevin adds some random noise to the motion. The level of noise maintains the system at a certain temperature.
  - Brownian is so random that there is no momentum
  - Variable methods use different time steps and depend on an error tolerance
  - Robosample Hybrid Monte Carlo uses Velocity Verlet

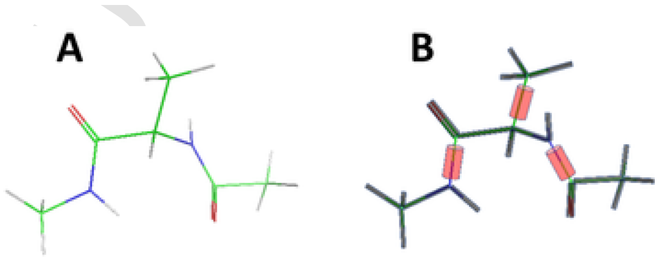
```
integrator = HMCIntegrator(300*kelvin,   # Temperature of
                           0.006*picoseconds) # Time step
```

- Thermostat
  - allow the kinetic energy of the system to change by modifying velocities
    - Andersen
    - Berendsen
    - Nose Hoover
  - Robosample uses Andersen
- Barostats keep the system at a certain pressure
  - allows the volume of the system to change
    - Hoover
    - Rahman
    - Parrinello
  - since we are using implicit water, let's not use a barostat

- “Reporters” store data about the simulation
  - “PDB” structure files are saved in the specified output directory
  - “DCD” is a binary file format for molecular dynamics trajectories
- “Report Interval” is how often the data are stored

```
simulation = Simulation(prmtop.topology, system, integrator, platform, properties)
simulation.reporters.append(PDBReporter('robots/', 10))
simulation.context.setPositions(inpcrd.positions)
```

- Let’s return to constraints, which are the heart of Robosample
  - the script generated by the Python interface consists of a series of KEYWORDS in uppercase each with a list of values corresponding to it
  - why a list of values: we can have however many worlds we want
- Our simulation has two worlds




```
# Robosample input
MOLECULES robots/bot0
PRMTOP bot.prmtop bot.prmtop       # Parameter file
INPCRD bot.rst7 bot.rst7     # Coordinate / Restart file
RBFILE bot.rb bot.rb   # Rigid bodies definition file
FLEXFILE bot.all.flex bot.hinges.flex # Flexibility definition file
ROOT_MOBILITY Cartesian Free # Ground to Compound mobilizer
OUTPUT_DIR robots robots

# Simulation
RUN_TYPE Normal Normal # normal HMC or Non-Eq HMC
ROUNDS 50
ROUNDS_TILL_REBLOCK 10 10
RANDOM_WORLD_ORDER FALSE FALSE
WORLDS R0 R1                       # Regimen (IC, TD, MIX, RB, RBMIX)
ROOTS 0 0
SAMPLER HMC HMC
TIMESTEPS 0.001 0.009              # Timesteps to be used with regimens
MDSTEPS 30 30  # Number of MD trial steps
BOOST_MDSTEPS 1 1
SAMPLES_PER_ROUND 1 3  # Number of acc-rej steps within a mixing round
REPRODUCIBLE FALSE FALSE
SEED 999 999

# Thermodynamics
THERMOSTAT Andersen Andersen   # Thermostat
TEMPERATURE_INI  300 300
TEMPERATURE_FIN  300 300
BOOST_TEMPERATURE  1 1      # Boost temperature
FFSCALE AMBER AMBER     # Force field
GBSA 0 0         # GBSA scale factor

# Correction factors
FIXMAN_POTENTIAL TRUE TRUE # Use Fixman potential
FIXMAN_TORQUE TRUE TRUE   # Use Fixman torque

# Output
VISUAL TRUE TRUE            # Use the visualizer
PRINT_FREQ  1 1
WRITEPDBS 1 0    # Write pdbs
GEOMETRY FALSE FALSE         # Calculate geometric features

DISTANCE 1 2
DIHEDRAL 1 2 3 4 1 2 3 4

# Software specs
THREADS 0 0
OPENMM TRUE TRUE
```

- Flexibility files contain information about the bonds to be considered flexible
- Flexibility file can be modified manually to change:
  - bonds allowed to move
  - type of joints between them
- Let’s try to allow less bonds to move
- Let’s try to change joint type from Pin to BallM

```
6 8 Pin # N N 1 ALA CA C 1 ALA 1.02834 1.02834
16 18 Pin # N N 2 ALA CA C 2 ALA 0.97698 0.97698
26 28 Pin # N N 3 ALA CA C 3 ALA 0.97272 0.97272
36 38 Pin # N N 4 ALA CA C 4 ALA 0.97698 0.97698
46 48 Pin # N N 5 ALA CA C 5 ALA 0.97272 0.97272
56 58 Pin # N N 6 ALA CA C 6 ALA 0.97698 0.97698
66 68 Pin # N N 7 ALA CA C 7 ALA 0.97272 0.97272
76 78 Pin # N N 8 ALA CA C 8 ALA 0.97698 0.97698
86 88 Pin # N N 9 ALA CA C 9 ALA 0.97272 0.97272
96 98 Pin # N N 10 ALA CA C 10 ALA 1.04308 1.04308
8 14 Pin # CA C 1 ALA C C 1 ALA 1.02834 1.02834
18 24 Pin # CA C 2 ALA C C 2 ALA 0.97698 0.97698
28 34 Pin # CA C 3 ALA C C 3 ALA 0.97272 0.97272
38 44 Pin # CA C 4 ALA C C 4 ALA 0.97698 0.97698
48 54 Pin # CA C 5 ALA C C 5 ALA 0.97272 0.97272
58 64 Pin # CA C 6 ALA C C 6 ALA 0.97698 0.97698
68 74 Pin # CA C 7 ALA C C 7 ALA 0.97272 0.97272
78 84 Pin # CA C 8 ALA C C 8 ALA 0.97698 0.97698
88 94 Pin # CA C 9 ALA C C 9 ALA 0.97272 0.97272
98 104 Pin # CA C 10 ALA C C 10 ALA 1.04308 1.04308
8 10 Pin # CA C 1 ALA CB C 1 ALA 1.02834 1.02834
18 20 Pin # CA C 2 ALA CB C 2 ALA 0.97698 0.97698
28 30 Pin # CA C 3 ALA CB C 3 ALA 0.97272 0.97272
38 40 Pin # CA C 4 ALA CB C 4 ALA 0.97698 0.97698
48 50 Pin # CA C 5 ALA CB C 5 ALA 0.97272 0.97272
58 60 Pin # CA C 6 ALA CB C 6 ALA 0.97698 0.97698
68 70 Pin # CA C 7 ALA CB C 7 ALA 0.97272 0.97272
78 80 Pin # CA C 8 ALA CB C 8 ALA 0.97698 0.97698
88 90 Pin # CA C 9 ALA CB C 9 ALA 0.97272 0.97272
98 100 Pin # CA C 10 ALA CB C 10 ALA 1.04308 1.04308
```

```
#6 8 Pin # N N 1 ALA CA C 1 ALA 1.02834 1.02834
#16 18 Pin # N N 2 ALA CA C 2 ALA 0.97698 0.97698
#26 28 Pin # N N 3 ALA CA C 3 ALA 0.97272 0.97272
#36 38 Pin # N N 4 ALA CA C 4 ALA 0.97698 0.97698
#46 48 Pin # N N 5 ALA CA C 5 ALA 0.97272 0.97272
#56 58 Pin # N N 6 ALA CA C 6 ALA 0.97698 0.97698
#66 68 Pin # N N 7 ALA CA C 7 ALA 0.97272 0.97272
#76 78 Pin # N N 8 ALA CA C 8 ALA 0.97698 0.97698
#86 88 Pin # N N 9 ALA CA C 9 ALA 0.97272 0.97272
#96 98 Pin # N N 10 ALA CA C 10 ALA 1.04308 1.04308
#8 14 Pin # CA C 1 ALA C C 1 ALA 1.02834 1.02834
#18 24 Pin # CA C 2 ALA C C 2 ALA 0.97698 0.97698
#28 34 Pin # CA C 3 ALA C C 3 ALA 0.97272 0.97272
38 44 Pin # CA C 4 ALA C C 4 ALA 0.97698 0.97698
#48 54 Pin # CA C 5 ALA C C 5 ALA 0.97272 0.97272
#58 64 Pin # CA C 6 ALA C C 6 ALA 0.97698 0.97698
#68 74 Pin # CA C 7 ALA C C 7 ALA 0.97272 0.97272
#78 84 Pin # CA C 8 ALA C C 8 ALA 0.97698 0.97698
#88 94 Pin # CA C 9 ALA C C 9 ALA 0.97272 0.97272
#98 104 Pin # CA C 10 ALA C C 10 ALA 1.04308 1.04308
#8 10 Pin # CA C 1 ALA CB C 1 ALA 1.02834 1.02834
#18 20 Pin # CA C 2 ALA CB C 2 ALA 0.97698 0.97698
#28 30 Pin # CA C 3 ALA CB C 3 ALA 0.97272 0.97272
#38 40 Pin # CA C 4 ALA CB C 4 ALA 0.97698 0.97698
#48 50 Pin # CA C 5 ALA CB C 5 ALA 0.97272 0.97272
#58 60 Pin # CA C 6 ALA CB C 6 ALA 0.97698 0.97698
#68 70 Pin # CA C 7 ALA CB C 7 ALA 0.97272 0.97272
#78 80 Pin # CA C 8 ALA CB C 8 ALA 0.97698 0.97698
#88 90 Pin # CA C 9 ALA CB C 9 ALA 0.97272 0.97272
#98 100 Pin # CA C 10 ALA CB C 10 ALA 1.04308 1.04308
```

# Additional Resources

- In preparation