Protocol for Basic Protein Molecular Dynamics Simulation Using Gromacs

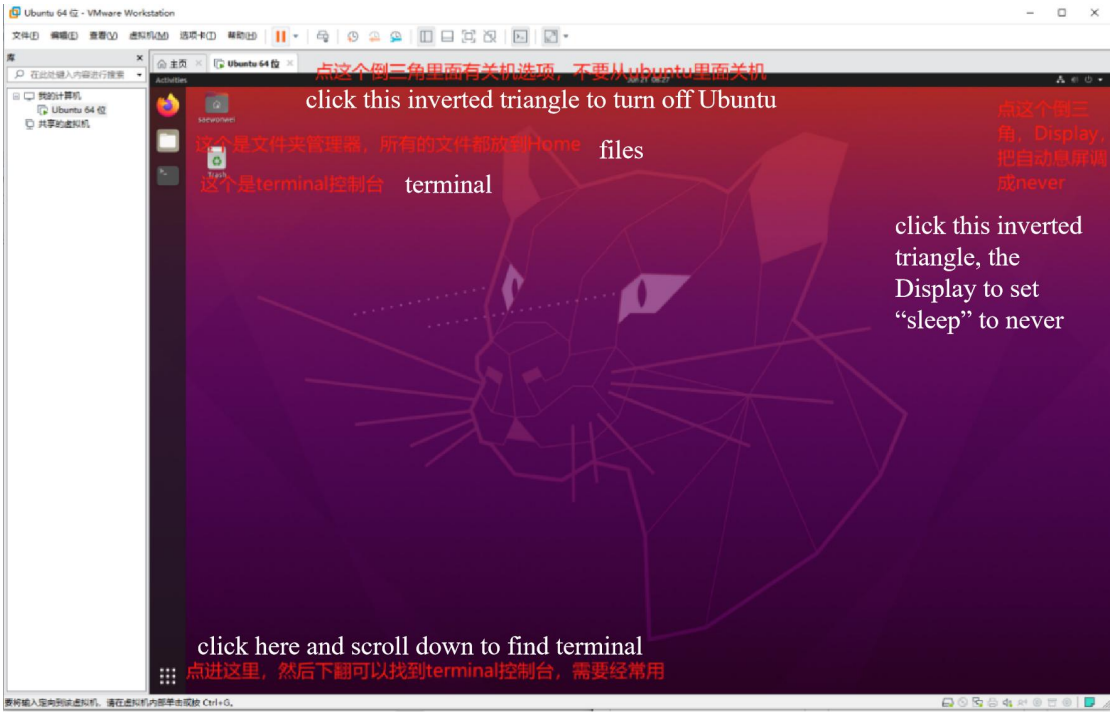
Correspondence to: Shiyuan Wei, School of Bioscience and Bioengineering, South China University of Technology, Guangzhou, 510006, Guangdong, China
E-mail:201864881365@mail.scut.edu.cn

1. Installation of Ubuntu

Visit https://ubuntu.com to download the latest Ubuntu installation file and install it in VMware. It is recommended to separate around 30Gb storage for Ubuntu virtual disk or even more, so as to make sure there will be enough storage for molecular modeling in Gromacs.

2. Introduction to Ubuntu

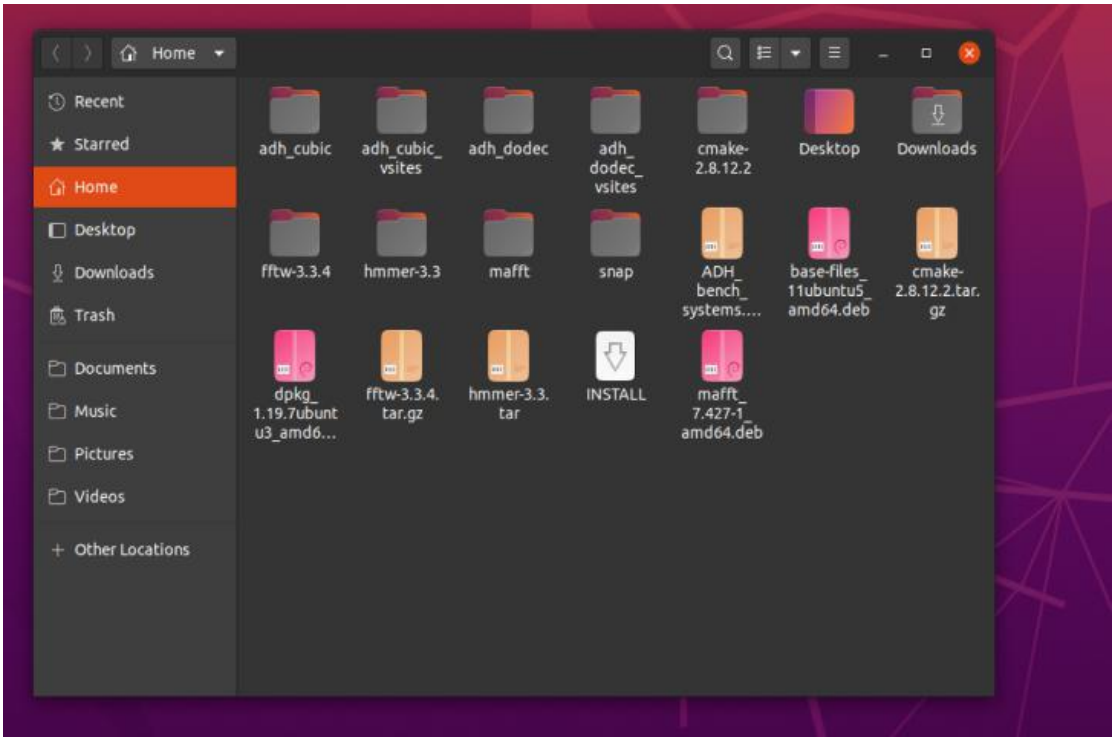
Ubuntu is an operation system based on Linux, a different platform and operation system compared with Windows, the most common one we are familiar with, which means, there exists many differences when using Ubuntu.



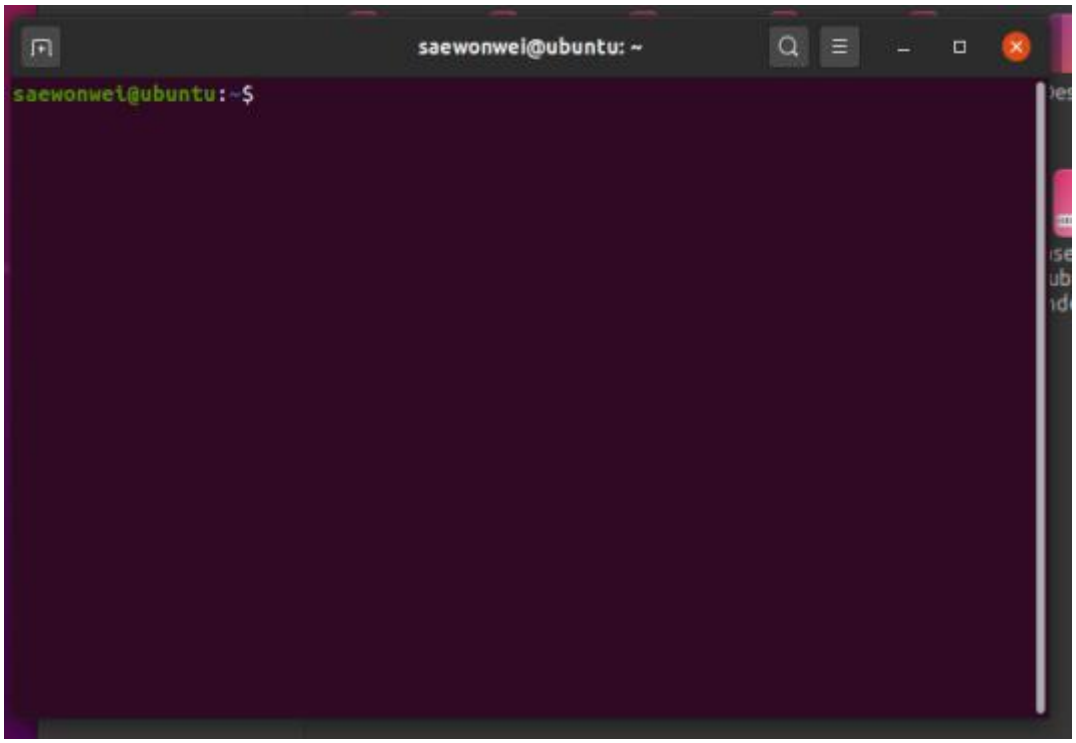
Considering that iGEM is an international competition with a lot of the contestants coming from China and Chinese-speaking countries, we specially prepare both English(white) and Chinese(red) annotations so as to help more iGEMers to easily get the hang of our protocol.

As the screenshot shows, this is the homepage of Ubuntu in VMware Workstation. There are three icons on the upper left site, Firefox browser, file manager, and terminal. To accomplish basic molecular modeling via Gromacs, we need to install and run it in terminal, which is different from the way we used to use software in Windows. If you can’t find the terminal on the upper left site, search for it in the bottom left menu. To make sure the Gromacs runs constantly, click the upper right sit triangle, then click the option menu, find the display menu, and block Ubuntu’s auto sleep function. Speaking from our own experience, we recommend turning off the Ubuntu through button in VMware, rather than directly powering off it inside the

Ubuntu, as an approach to prevent its unnecessary crash.



This screenshot shows the interface of file manager in Ubuntu, almost the same as we saw in Windows system. However, the extension name of compressed file in Ubuntu is tar.gz or .tar, rather than .zip or .7z in Windows.



This screenshot shows the interface of terminal in Ubuntu. Terminal is a place where we can input our order, like installing a software, or executing a software such as Gromacs.

In some cases, administrator's password is required to do specific command. Input your password and it is fine whether it is shown on the screen or not. Remember

to press ENTER after your input.

Each command you input needs to be executed, and only when the steps are all done, a green word username@ubuntu will be shown on the screen, as the screenshot shows.

3. Installation of Gromacs

Apart from Gromacs itself, some other software are needed to be installed previous to Gromacs.

Installation of gcc

Open terminal, input sudo apt install gcc. Wait until it is installed.

During the next few steps, some errors may be reported, which is very usual, so don't worry. Google it and the answer is easy to approach. Some common errors have been listed below.

Could not open lock file /var/lib/dpkg/lock-frontend-open(13: Permission denied),E: flAbsPath on /var/lib/dpkg/status failed - realpath (2: No such file or directory),etc..

Installation of g++

Open terminal, input sudo apt-get upgrade, then input sudo apt-get install g++.

Installation of fftw

Use a browser in Windows system, open the URL http://fftw.org/fftw-3.3.4.tar.gz, download the file fftw-3.3.4.tar.gz in Windows. Drag the file into the Ubuntu file manager interface, open it and extract it into HOME.

Open the terminal and input tar xvf fftw-3.3.4.tar.gz.

Input cd fftw-3.3.4.

Input ./configure.

Input make.

Input sudo make install. Input cd.

Installation of cmake

Use a browser in Windows system, open the URL http://www.cmake.org/files/v2.8/cmake-2.8.12.2.tar.gz , and download the file cmake-2.8.12.2.tar.gz in Windows. Drag the file into the Ubuntu file manager interface, open it and extract it into HOME.

Open the terminal and input tar xvf cmake-2.8.12.2.tar.gz.

Input cd cmake-2.8.12.2.

Input cd cmake-2.8.12.2.

Input ./configure. Input make.

Input suto make install.

Input cd. Input export PATH=/opt/cmake-2.8.12.2/bin/:$PATH.

Turn off your Ubuntu and restart it.

Installation of Gromacs

Open the terminal, input apt install gromacs, wait until it is finish, then close the terminal and reopen it. Input gmx to test if the Gromacs is installed successfully.

4. Basic molecular modeling in Gromacs

This protocol only introduces some basic functions of Gromacs, and some basic commands in Gromacs. If more specific information is needed, please check the official guiding documents from Gromacs.

1. Preparation

To perform protein molecular modeling, a pdb file from NCBI Blast or other protein databases is required, which is quite easy to access. Naming the pdb file with the protein's name is also needed, especially when you need to process multiple proteins at the same time. Drag the pdb file into Ubuntu system as the way showed previously, and create a file folder for it is all the preparation. Here, we use RelB-RelE to name the pdb file in order to demonstrate clearly.

2. Create some file

Open the terminal and input gmx pdb2gmx -ignh -ff amber99sb-ildn -f RelB-RelE.pdb -o RelB-RelE.gro -p RelB-RelE.top -water tip3p. Input gmx editconf -f RelB-RelE.gro -o RelB-RelE-PBC.gro -bt cubic -d 1.0.

3. Pass some definitions to the preprocessor

Input vi em-vac-pme.mdp to create a .mdp file. Press insert in your keyboard and input these following informations in the terminal.

define =-DFLEXIBLE (choose the flexible water model to use the steepest descent method to minimize the energy)

integrator = steep (choose the steepest descent method to minimize the energy)

emtol = 500.0 (kJ mol^-1^ nm^-1, if the maximum value of the force is less than this value, the energy minimization is considered to converge)

Emstep = 0.01 (nm, set the initial step length)

Nsteps = 1000 (set the maximal number of iterations to minimize the energy)

nstenergy = 1 (set the energy output frequency)

Energygrps = system (the energy group which will be output)

Ntlist = 1 (update the near list frequency, 1 means update it every step)

Ns type = grid (set the near list as grid or simple)

Coulombtype = PME (the calculation method of long-range electrostatic, PME for Ewald method, which can be replaced by cut-off as well)

rlist = 1.0.

rcoulomb = 1.0.

vdwtype = cut-off.
rvdw = 1.0.
constraints = none.
pbc = xyz.

After inputting all these information, press ESC and input :wq to quit.

Input gmx grompp -f em-vac-pme.mdp -c RelB-RelE-PBC.gro -p RelB-RelE.top -o em-vac.tpr -maxwarn 1.

Input gmx mdrun -v -deffnm em-vac 1000steps.

Input gmx solvate -cp em-vac.gro -cs spc216.gro -p RelB-RelE.top -o RelB-RelE-b4ion.gro.

Input vi em-sol-pme.mdp, to create a .mdp file. Press insert in your keyboard and input these following information in the terminal.

define = -DFLEXIBLE
integrator = steep
emtol = 250.0
nsteps = 10000
nstenergy = 1
energygrps = System
nstlist = 1
ns_type = grid
coulombtype = PME
rlist = 1.0
rcoulomb = 1.0
rvdw = 1.0
constraints = none
pbc = xyz

Press ESC in your keyboard and input :wq to quit.

Input gmx grompp -f em-sol-pme.mdp -c RelB-RelE-b4ion.gro -p RelB-RelE.top -o ion.tpr -maxwarn 1.

Input gmx genion -s ion.tpr -o RelB-RelE-b4em.gro -neutral -conc 0.15 -p RelB-RelE.top, and press 13 to choose the solution system (SOL).

Input gmx grompp -f em-sol-pme.mdp -c RelB-RelE-b4em.gro -p RelB-RelE.top -o em-sol.tpr.

Input gmx mdrun -v -deffnm em-sol.

Input vi em-sol-pme.mdp, to create a .mdp file. Press insert in your keyboard and input these following informations in the terminal.

define = -DFLEXIBLE
integrator = cg
emtol = 250.0
nsteps = 10000
nstenergy = 1
energygrps = System
nstlist = 1
ns_type = grid
coulombtype = PME
rlist = 1.0
rcoulomb = 1.0
rvdw = 1.0
constraints = none
pbc = xyz

Press ESC in your keyboard and input :wq to quit.

Input vi nvt-pr-md.mdp, to create a .mdp file. Press insert in your keyboard and input these following information in the terminal.

define = -DPOSRES
integrator = md
dt = 0.002
nsteps = 50000
nstxout = 500
nstvout = 500
nstenergy = 500
nstlog = 500
energygrps = Protein Non-Protein
nstlist = 5
ns_type = grid
pbc = xyz
rlist = 1.0
coulombtype = PME
pme_order = 4
fourierspacing = 0.16
rcoulomb = 1.0
vdw-type = Cut-off
rvdw = 1.0
tcoupl = v-rescale
tc-grps = Protein Non-Protein
tau_t = 0.1
ref_t = 300
DispCorr = EnerPres
pcoupl = no
gen_vel = yes
gen_temp = 300

gen_seed = -1
constraints = all-bonds
continuation = no
constraint_algorithm = lincs
lincs_iter = 1
lincs_order = 4

Press ESC in your keyboard and input :wq to quit.

Input gmx grompp -f nvt-pr-md.mdp -c em-sol.gro -p RelB-RelE.top -o nvt-pr.tpr -r em-sol.gro.

Input gmx mdrun -deffnm nvt-pr -v.

Input vi npt-pr-md.mdp, to create a .mdp file. Press insert in your keyboard and input these following information in the terminal.

define = -DPOSRES
integrator = md
dt = 0.002
nsteps = 50000
nstxout = 500
nstvout = 500
nstfout = 500
nstenergy = 500
nstlog = 500
energygrps = Protein Non-Protein
nstlist = 5
ns-type = Grid
pbc = xyz
rlist = 1.0
coulombtype = PME
pme_order = 4
fourierspacing = 0.16
rcoulomb = 1.0
vdw-type = Cut-off
rvdw = 1.0
Tcoupl = v-rescale
tc-grps = Protein Non-Protein
tau_t = 0.1 0.1
ref_t = 300 300
DispCorr = EnerPres
Pcoupl = Parrinello-Rahman
Pcoupltype = Isotropic
tau_p = 0.5
compressibility = 4.5e-5
ref_p = 1.0
refcoord_scaling = com
gen_vel = no

constraints = all-bonds
continuation = yes
constraint_algorithm = lincs
lincs_iter = 1
lincs_order = 4

Press ESC in your keyboard and input :wq to quit.

Input gmx grompp -f npt-pr-md.mdp -c nvt-pr.gro -p RelB-RelE.top -o npt-pr.tpr -r nvt-pr.gro.

Input gmx mdrun -deffnm npt-pr -v.

Input vi npt-pr-md.mdp, to create a .mdp file. Press insert in your keyboard and input these following information in the terminal.

integrator = md
dt = 0.002
nsteps = 50000000; 1 ns
nstxout = 500
nstvout = 500
nstfout = 500
nstenergy = 500
nstlog = 500
energygrps = Protein Non-Protein
nstlist = 5
ns-type = Grid
pbc = xyz
rlist = 1.0
coulombtype = PME
pme_order = 4
fourierspacing = 0.16
rcoulomb = 1.0
vdw-type = Cut-off
rvdw = 1.0
Tcoupl = v-rescale
tc-grps = Protein Non-Protein
tau_t = 0.1 0.1
ref_t = 300 300
DispCorr = EnerPres
Pcoupl = Parrinello-Rahman
Pcoupltype = Isotropic
tau_p = 2.0
compressibility = 4.5e-5
ref_p = 1.0
gen_vel = no
constraints = all-bonds
continuation = yes
constraint_algorithm = lincs

lincs_iter = 1

lincs_order = 4

Press ESC in your keyboard and input :wq to quit.

Input gmx grompp -f npt-nopr-md.mdp -c npt-pr.gro -p RelB-RelE.top -o npt-nopr.tpr

Acknowledgement

This protocol was inspired by our team advisor Jerome Rumdon Lon.