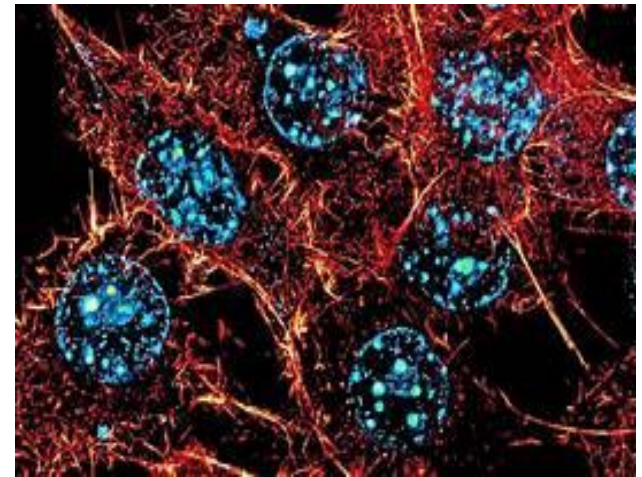
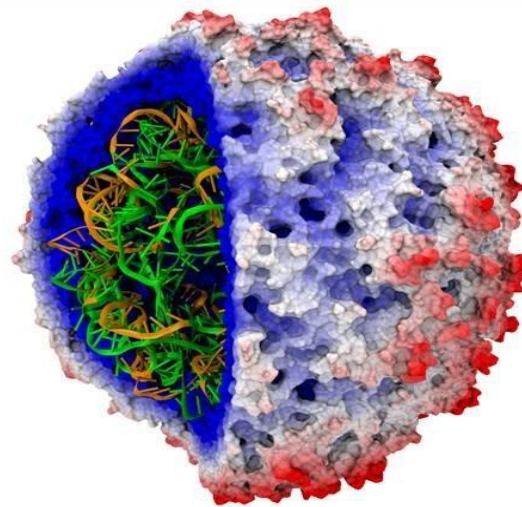
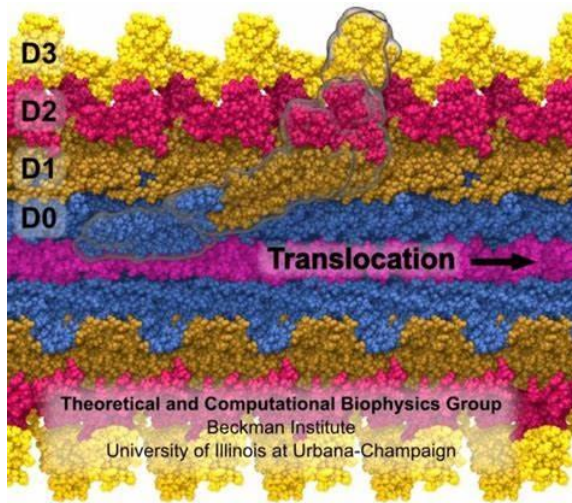


Vật lý Y sinh

Trần Thị Minh Thu

Email: ttmthu@hcmus.edu.vn



TÀI LIỆU THAM KHẢO

1. Vật lý thống kê, Nguyễn Nhật Khanh, Tủ sách Trường ĐHKHTN, 1999.
2. Mô phỏng trong Vật lý, Võ Văn Hoàng (Chủ biên) NXB Đại học Quốc gia TP HCM, 2016.
3. Understanding Molecular Simulation: From algorithms to application, Daan Frenkel and Berend Smit, Second edition, 2005, Academic press.

HÌNH THỨC THI VÀ CHẤM ĐIỂM

- Bài tập cá nhân và điểm danh tại lớp: 20%
- Báo cáo đồ án: 40%
- Điểm cuối kỳ: 40%

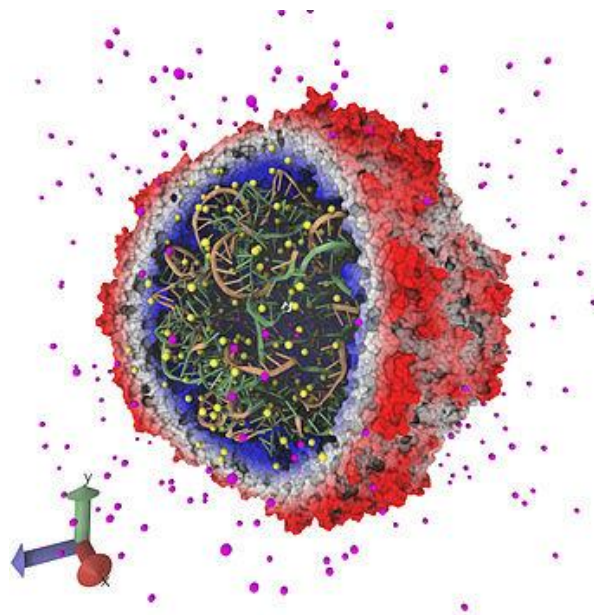


03 TC lý thuyết: 12 buổi; 04 tiết / 01 buổi

- Buổi 1+2: Giới thiệu về protein và khoa học tính toán
- Buổi 3: Phương pháp mô phỏng MD.
- Buổi 4: Kỹ thuật mô phỏng MD: thuật giải, trường lực, các phần mềm.
- Buổi 5: Kỹ thuật mô phỏng MD: thuật giải, trường lực, các phần mềm (tiếp theo).
- Buổi 6+7+8: Giới thiệu và hướng dẫn sử dụng phần mềm VMD, XMGRACE, GROMACS.
- Buổi 9+10: Giao đồ án môn học và hướng dẫn làm đồ án
- Buổi 11+12: Báo cáo đồ án



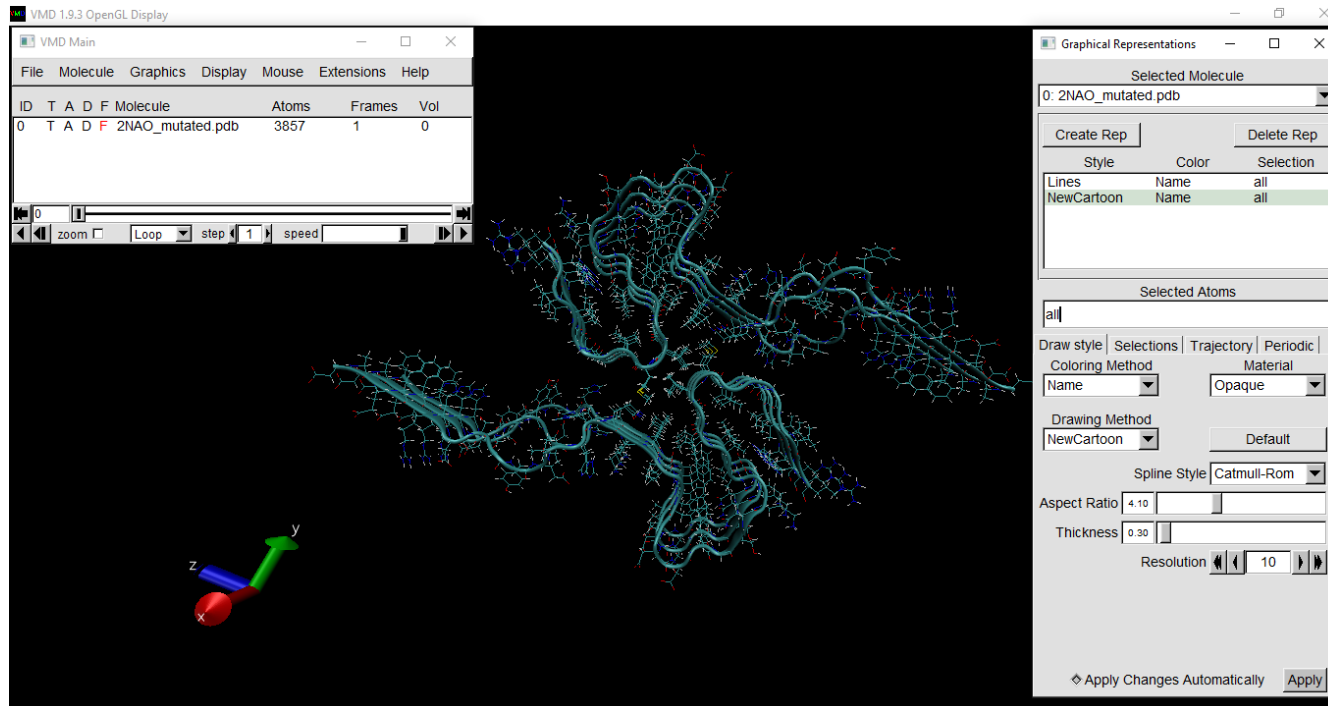
Buổi 5: Giới thiệu và hướng dẫn sử dụng phần mềm VMD, XMGRACE, GROMACS



VMD - Visual Molecular Dynamics

<http://www.ks.uiuc.edu/Research/vmd/>

- Giao diện đồ họa
- Miễn phí cho cộng đồng học thuật
- Dữ liệu đầu vào cho NAMD, CHARMM, Amber, Gromacs



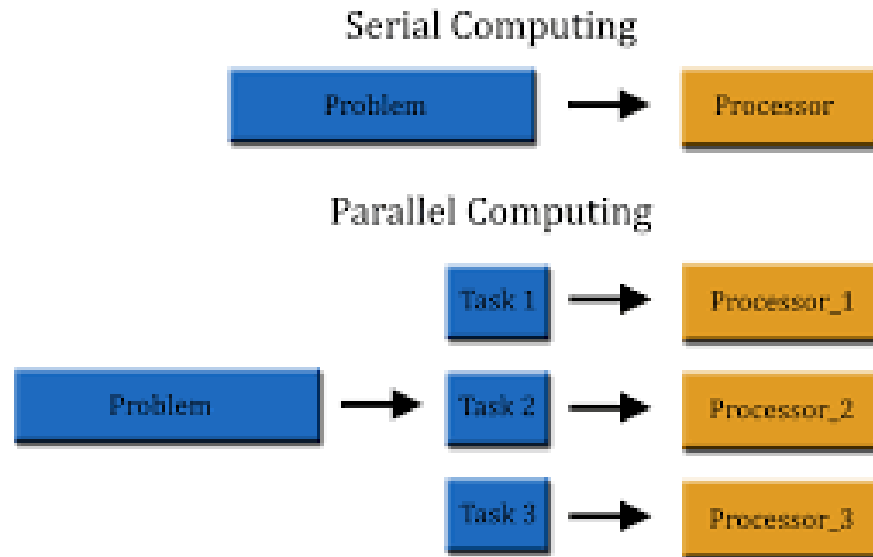
GROMACS

GROningen MAchine for Chemical Simulations

<http://www.gromacs.org/>

Why Gromacs?

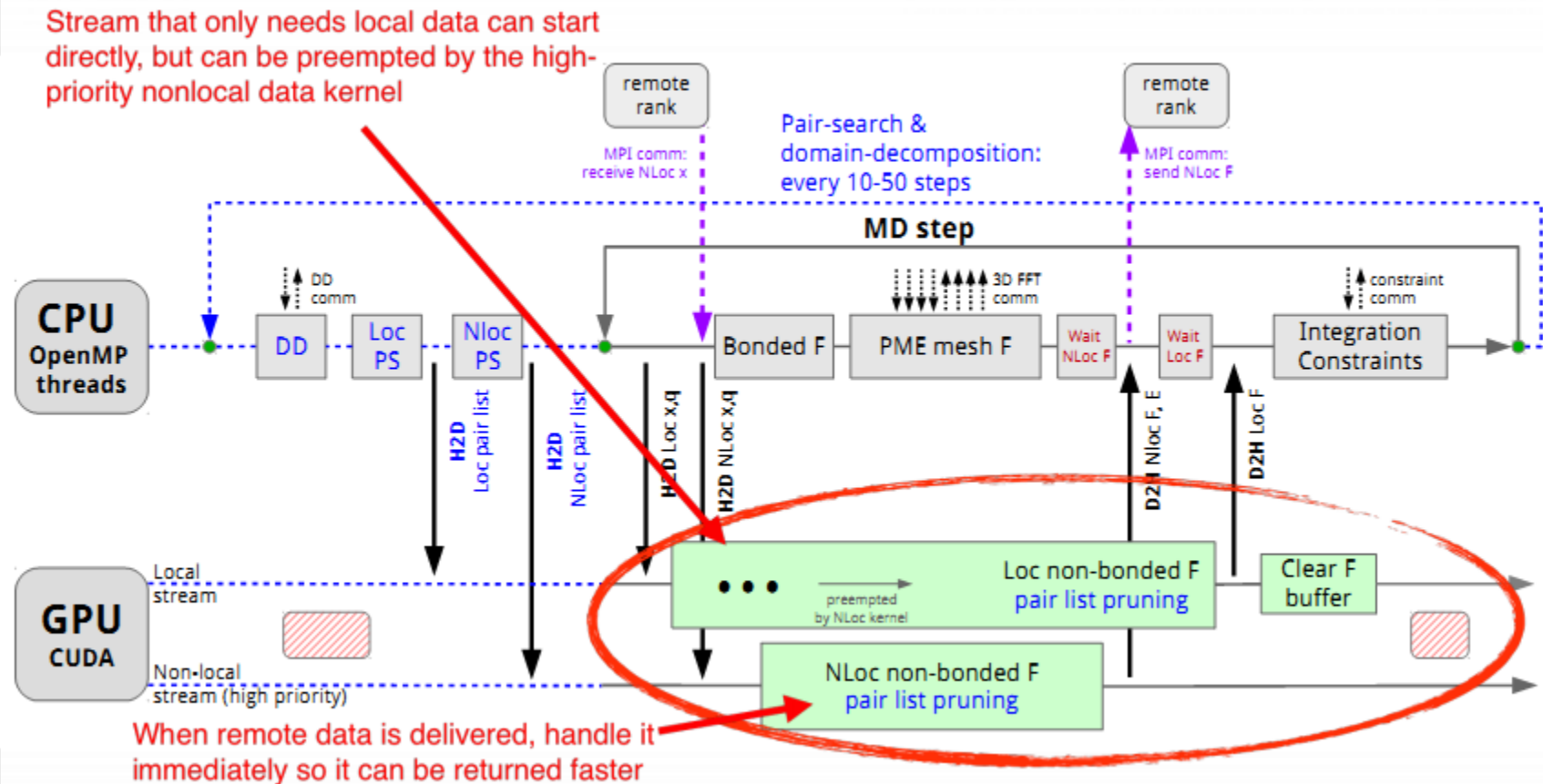
- Hỗ trợ mạnh tính toán song song



- Miễn phí
- Chỉ tốn máy tính và ổ cứng
- Mã nguồn mở

Why Gromacs?

- Hỗ trợ GPU



Why Gromacs?



Table of Contents

Welcome to the GROMACS documentation!

- Latest releases
- SYCL enabled development versions
- Current development branch
- Releases no longer being maintained
- Branch maintenance policy

Quick search

Go

Welcome to the GROMACS documentation!

Latest releases

- GROMACS 2021 series

[Documentation](#) for the current version

- 2021.1 released March 8th, 2021
 - [Download](#)
 - [Release Notes](#)
- 2021 released January 28th, 2021
 - [Download](#)
 - [Release Notes](#)

- GROMACS 2020 series

[Documentation](#) for the current version

- 2020.6 released March 4th, 2021
 - [Download](#)
 - [Release Notes](#)
- 2020.5 released January 6th, 2021
 - [Download](#)
 - [Release Notes](#)

◦ 2020.4 released October 6th, 2020

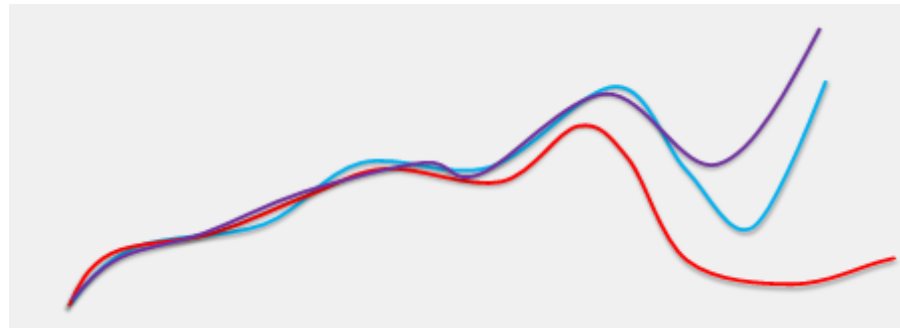
Lịch sử hình thành

1991-2000: Dự án GROMACS được thực hiện tại Bộ môn Hóa Lý, [Đại học Groningen](#), Hà Lan.

- Mục đích là tạo ra một hệ thống tính toán song song cho mô phỏng phân tử sinh hóa học.
- Được viết đầu tiên bằng [Fortran 77](#) (**GRO**ningen **MO**lecular **S**imulation - GROMOS), sau đó viết lại bằng C bởi cùng một nhóm tác giả.
- Từ 2001, GROMACS được phát triển bởi team Gromacs tại [Royal Institute of Technology](#) và [Uppsala University](#), [Sweden](#).

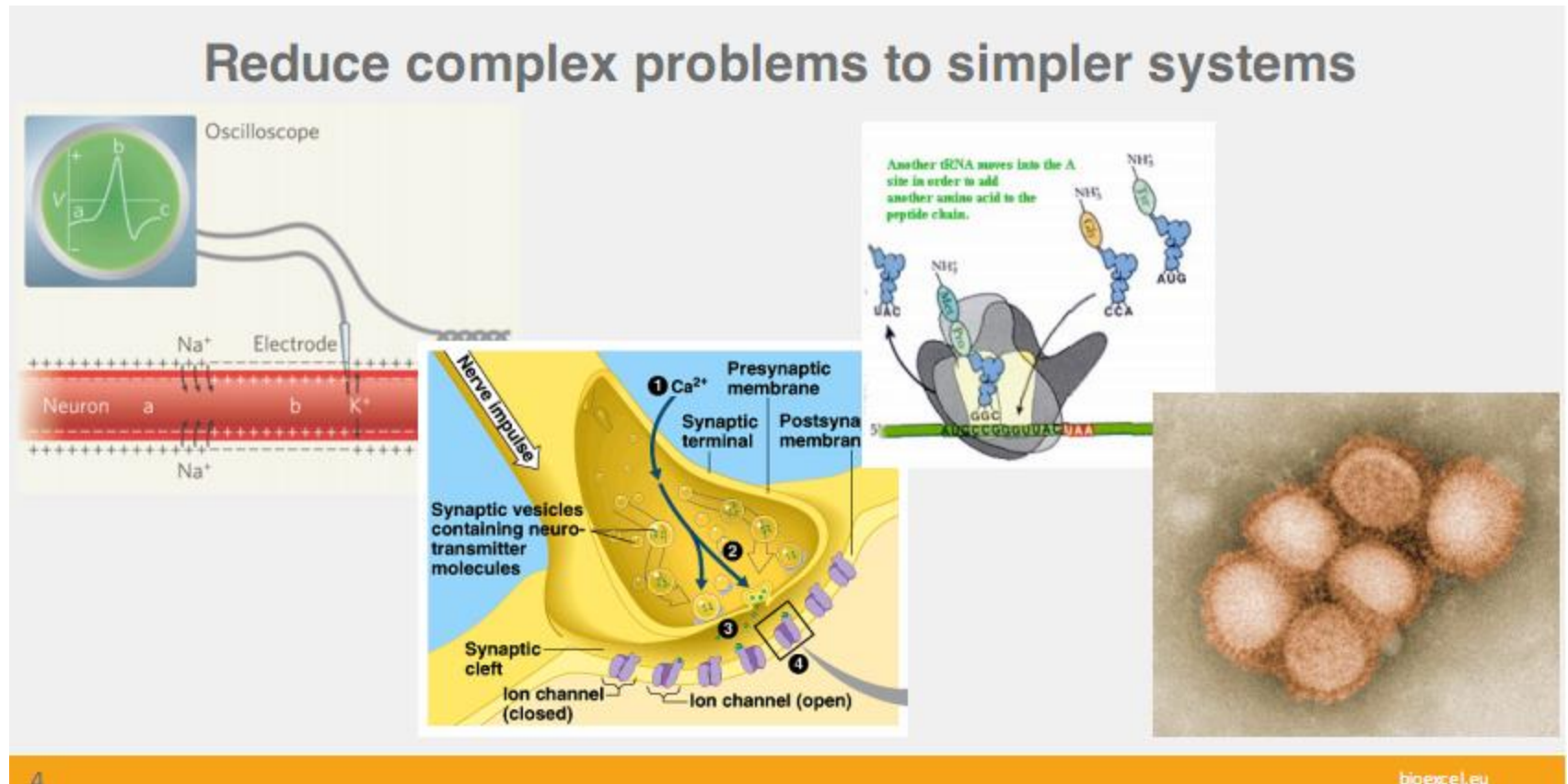
Tính năng

- very fast due to [algorithmic](#) and processor-specific optimizing, typically running 3-10 times faster than many simulation programs.
- operated via the [command-line interface](#), and can use files for input and output.
- provides calculation progress and [estimated time of arrival](#) (ETA) feedback, a trajectory viewer, and an extensive library for trajectory analysis.



Tính năng

- support for different [force fields](#) makes GROMACS very flexible. It can be executed in parallel, using [Message Passing Interface](#) (MPI) or [threads](#).
- contains a script to convert molecular coordinates from [Protein Data Bank](#) (PDB) files into the formats it uses internally.
- produces a trajectory file, describing the movements of the atoms over time.



Cài đặt

- C & C++ Compiler which comes built-in with Ubuntu.
- CMake – A linux software to make binaries
- BuildEssential – It is a reference for all the packages needed to compile a package.
- FFTW Library: a library used by Gromacs to compute discrete Fourier transform
- Gromacs 2020.4 Package

Cài đặt

- Getting Started

sudo apt-get update

sudo apt-get upgrade

- First step in installing Gromacs is to get cmake, In the terminal, type:

sudo apt-get install cmake

- Check the version of cmake by following command

cmake --version

- install build essential

sudo apt-get install build-essential

- To know the path of our current working directory, in terminal, type:

pwd

Cài đặt

- Install Fourier Transform

sudo apt-get install libfftw3-dev

- Install Gromacs

- Wget: <http://ftp.gromacs.org/pub/gromacs/gromacs-2020.4.tar.gz> Now extract GROMACS archive

tar xvzf gromacs-2020.4.tar.gz

cd gromacs-2020.4/

- Create a directory called “Build” where we will keep our compiled binaries

mkdir build

cd build

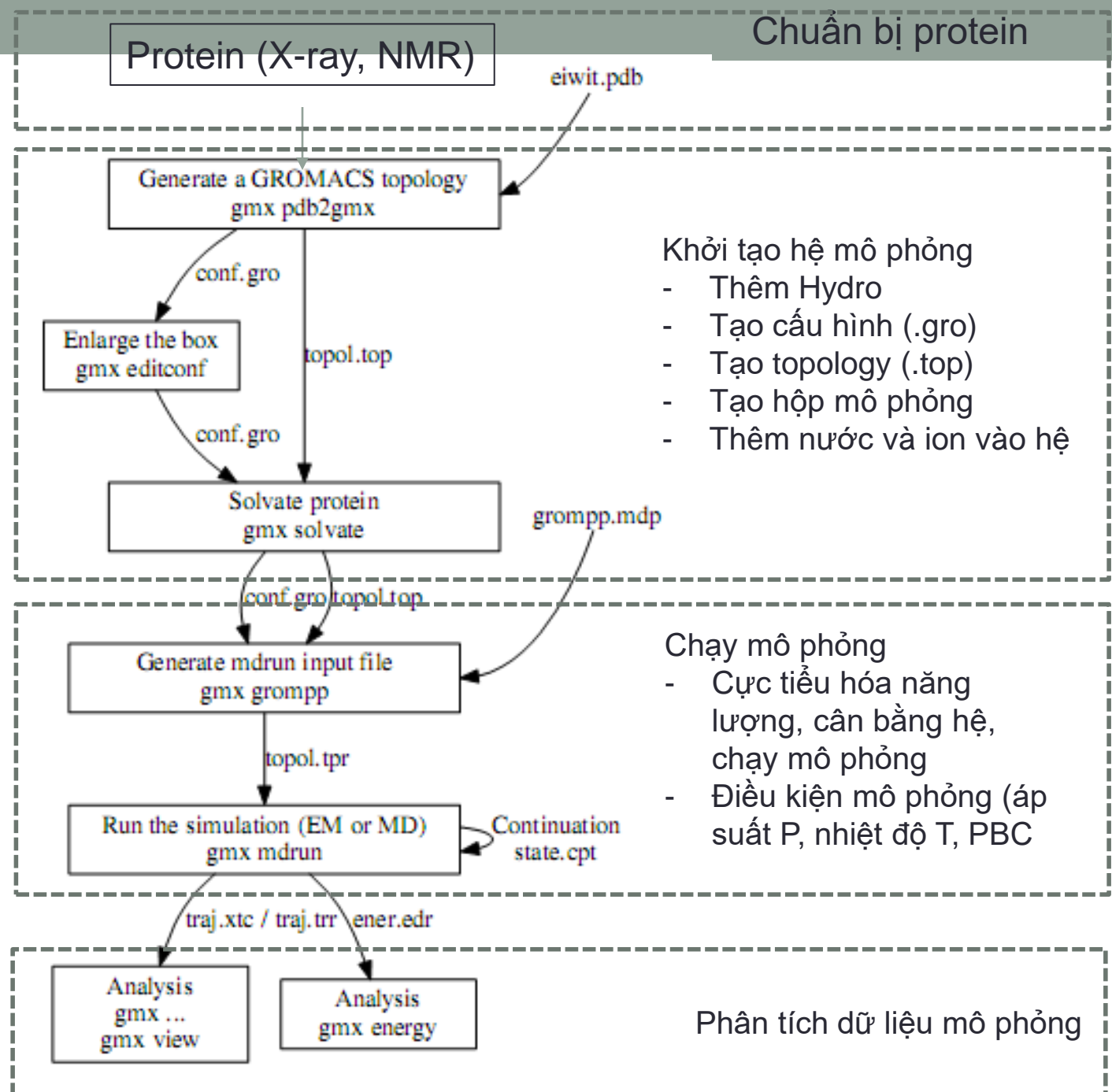
Cài đặt

- **Replace “pwdpath” with the path of working directory** that you have noted earlier in following command:
- `sudo cmake .. -DGMX_BUILD_OWN_FFTW=OFF -
DREGRESSIONTEST_DOWNLOAD=OFF -
DCMAKE_C_COMPILER=gcc -
DREGRESSIONTEST_PATH=pwdpath</str
ong>/Downloads/regressiontests-2020.4`

Cài đặt

- If everything goes well, the message in your terminal will say “Generating Done. Build files written... “. If not, make sure you have replaced the pwd path in command with the path of your home directory. If you have forgotten it, just open another terminal and type pwd.
- Now let's first check and make the real thing..
make checksudo
make install
- Now, It may take some time depending on your configuration. After completion, execute it:
source /usr/local/gromacs/bin/GMXRC
- After the successful installation, you may check the version of your Gromacs with a command to make sure installation finished as expected.
gmx pdb2gmx –version

Flowchart



File topology (.gro)

- ***gmx pdb2gmx***: chuyển thông tin từ file pdb sang file topology.
- Các thông tin của hệ được khai báo trong file topology
- Điều hiểu hết lệnh: ***gmx pdb2gmx -h*** (*h : help*)

```
gmx pdb2gmx -f 1AKI_clean.pdb -o 1AKI_processed.gro -water spce
```

Select the Force Field:

From '/usr/local/gromacs/share/gromacs/top':

- 1: AMBER03 protein, nucleic AMBER94 (Duan et al., J. Comp. Chem. 24, 1999-2012, 2003)
- 2: AMBER94 force field (Cornell et al., JACS 117, 5179-5197, 1995)
- 3: AMBER96 protein, nucleic AMBER94 (Kollman et al., Acc. Chem. Res. 29, 461-469, 1996)
- 4: AMBER99 protein, nucleic AMBER94 (Wang et al., J. Comp. Chem. 21, 1049-1074, 2000)
- 5: AMBER99SB protein, nucleic AMBER94 (Hornak et al., Proteins 65, 712-725, 2006)
- 6: AMBER99SB-ILDN protein, nucleic AMBER94 (Lindorff-Larsen et al., Proteins 78, 1950-58, 2010)
- 7: AMBERGS force field (Garcia & Sanbonmatsu, PNAS 99, 2782-2787, 2002)
- 8: CHARMM27 all-atom force field (CHARM22 plus CMAP for proteins)
- 9: GROMOS96 43a1 force field
- 10: GROMOS96 43a2 force field (improved alkane dihedrals)
- 11: GROMOS96 45a3 force field (Schuler JCC 2001 22 1205)
- 12: GROMOS96 53a5 force field (JCC 2004 vol 25 pag 1656)
- 13: GROMOS96 53a6 force field (JCC 2004 vol 25 pag 1656)
- 14: GROMOS96 54a7 force field (Eur. Biophys. J. (2011), 40,, 843-856, DOI: 10.1007/s00249-011-0700-9)
- 15: OPLS-AA/L all-atom force field (2001 aminoacid dihedrals)

File Topology (.top)

Khai báo thư viện trường lực opls-aa để gọi ra các tham số từ trường lực

```
#include "oplsaa.ff/forcefield.itp"
```

“Name”: Tên “Protein_A” khai báo tên của phân tử, dựa trên việc protein được ký hiệu là “chain A” trong tập tin cấu trúc PDB.

```
; Name      nrexcl  
Protein_A    3
```

Các thông tin về nguyên tử

STT của nhóm điện tích (charge group number)

[atoms]										
; nr	type	resnr	residue	atom	cgnr	charge	mass	typeB	chargeB	massB
; residue	1	LYS	rtp LYSH q +2.0							
1	opls_287	1	LYS	N	1	-0.3	14.0067	; qtot	-0.3	
2	opls_290	1	LYS	H1	1	0.33	1.008	; qtot	0.03	
3	opls_290	1	LYS	H2	1	0.33	1.008	; qtot	0.36	
4	opls_290	1	LYS	H3	1	0.33	1.008	; qtot	0.69	
5	opls_293B	1	LYS	CA	1	0.25	12.011	; qtot	0.94	
6	opls_140	1	LYS	HA	1	0.06	1.008	; qtot	1	

File Topology (.top)

- Khai báo các nguyên tử được giới hạn chuyển động (position restraints) trong quá trình cân bằng hệ.
- Tập tin “posre.itp” được tạo tự động bởi lệnh pdb2gmx (định nghĩa các hằng số lực, tương ứng hệ số đàn hồi của lò xo) của thế đàn hồi tác dụng lên nguyên tử.

```
; Include Position restraint file
#ifdef POSRES
#include "posre.itp"
#endif
```


File Topology (.top)

- Khai báo nước: sử dụng mô hình nước SPC/E

```
; Include water topology
#include "oplsaa.ff/spce.itp"

#ifdef POSRES_WATER
; Position restraint for each water oxygen
[ position_restraints ]
; i funct      fcx      fcy      fcz
  1    1      1000      1000      1000
#endif
```

- Tham số của ion được khai báo trong trường lực

```
; Include generic topology for ions
#include "oplsaa.ff/ions.itp"
```

File Topology (.top)

Khai báo toàn bộ thông tin về hệ mô phỏng:

[system]: Khai báo tên của hệ và tên này được ghi ra tập tin kết quả trong quá trình mô phỏng

[molecules]: Liệt kê các loại phân tử có trong hệ và số lượng loại phân tử đó.

```
[ system ]  
; Name  
LYSOZYME  
  
[ molecules ]  
; Compound      #mols  
Protein_A       1
```

File cấu trúc (.gro)

File .gro chứa các thông tin về vận tốc.

```
MD of 2 waters, t= 0.0
6
1WATER OW1 1 0.126 1.624 1.679 0.1227 -0.0580 0.0434
1WATER HW2 2 0.190 1.661 1.747 0.8085 0.3191 -0.7791
1WATER HW3 3 0.177 1.568 1.613 -0.9045 -2.6469 1.3180
2WATER OW1 4 1.275 0.053 0.622 0.2519 0.3140 -0.1734
2WATER HW2 5 1.337 0.002 0.680 -1.0641 -1.1349 0.0257
2WATER HW3 6 1.326 0.120 0.568 1.9427 -0.8216 -0.0244
1.82060 1.82060 1.82060
```

File tham số (.mdp)

File tham số (.mdp) chứa thông tin để mô phỏng MD: bước thời gian, số bước, nhiệt độ, áp suất.

Ví dụ: ions.mdp

```
; minim.mdp - used as input into grompp to generate em.tpr
; Parameters describing what to do, when to stop and what to save
integrator = steep      ; Algorithm (steep = steepest descent minimization)
emtol      = 1000.0     ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep     = 0.01       ; Minimization step size
nsteps     = 50000      ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the
interactions
nstlist     = 1         ; Frequency to update the neighbor list and long range forces
cutoff-scheme = Verlet   ; Buffered neighbor searching
ns_type     = grid      ; Method to determine neighbor list (simple, grid)
coulombtype  = PME       ; Treatment of long range electrostatic interactions
rcoulomb     = 1.0       ; Short-range electrostatic cut-off
rvdw         = 1.0       ; Short-range Van der Waals cut-off
pbc         = xyz       ; Periodic Boundary Conditions in all 3 dimensions
```

File đầu vào (.tpr)

- Kết nối các thông tin từ file cấu trúc (gro), file cấu hình (top), file tham số (mdp) và các lựa chọn khác -> file
- File tpr chứa tất cả các thông tin cần thiết cho mô phỏng bằng GROMACS.
- Sử dụng **gmx grompp**

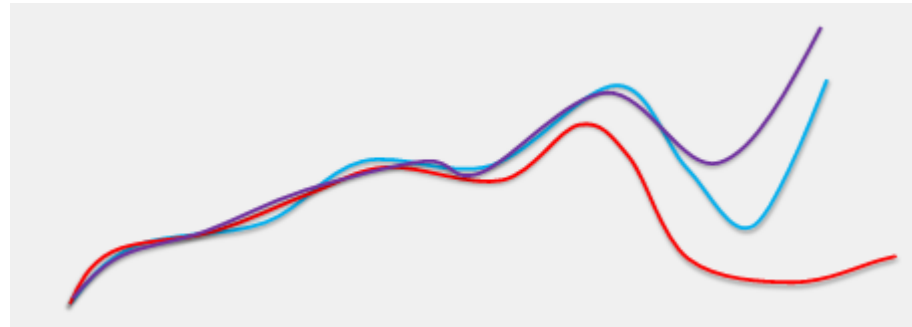
```
gmx grompp -f minim.mdp -c 1AKI_solv_ions.gro -p topol.top -o em.tpr
```

```
gmx mdrun -v -deffnm em
```

File kết quả (.trr, .tng, .xtc)

gmx mdrun
mdrun
mdrun_mpi.

trajectory file *.trr
logfile *.log,
checkpoint file *.cpt



Running Gromacs

Input files: *.pdb, *.gro, *.itp, *.top, *.mdp, *.tpr

Output files: *.trr, *.xtc, *.edr, *.log

Input:

- pdb : Protein data bank format
- gro: Gromacs format (atom co-ordinates)
- itp: atom topologies (charges, mass, radii, etc)
- top: forcefields, number of molecules, water, etc
- mdp: molecular dynamics simulation parameters
- tpr: all of the above

Output:

- trr: trajectory file (co-ordinates and velocity)
- xtc: trajectory file (co-ordinates only)
- edr: trajectory file (energies)
- log: CPU time, MFLOP, etc.

Converting pdb to gro:

```
pdb2gmx -f input.pdb -o output.gro -o protein.top
```

Solvation of the system:

```
editconf -f output.gro -box lx ly lz -o presol.gro
```

To check box size is OK

```
vmd presol.gro
```

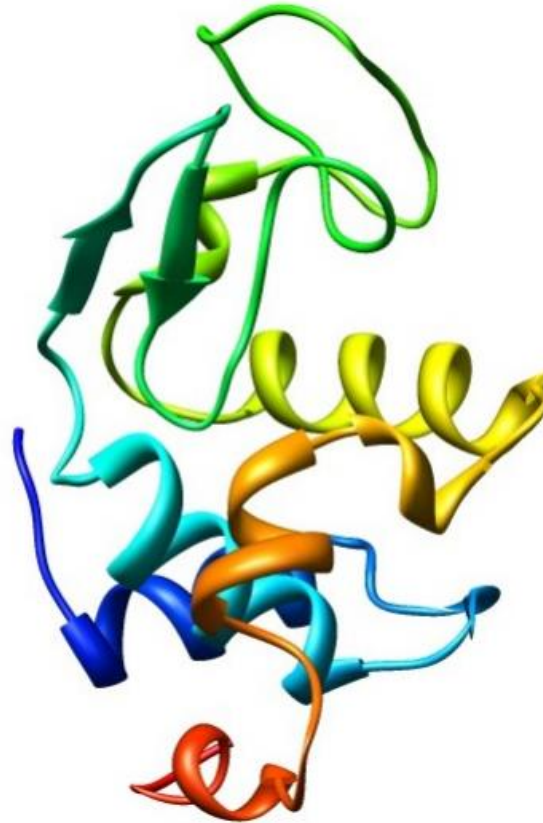
Set unitcell on (check that protein is in the center of the box and the edge is about 3 nm away from the protein)

Tutorial: Lysozyme in Water

Justin A. Lemkul, Ph.D.

Department of Pharmaceutical Sciences

University of Maryland, Baltimore



<http://www.mdtutorials.com/gmx/>