

# MD Analysis

Naf Guo  
2025

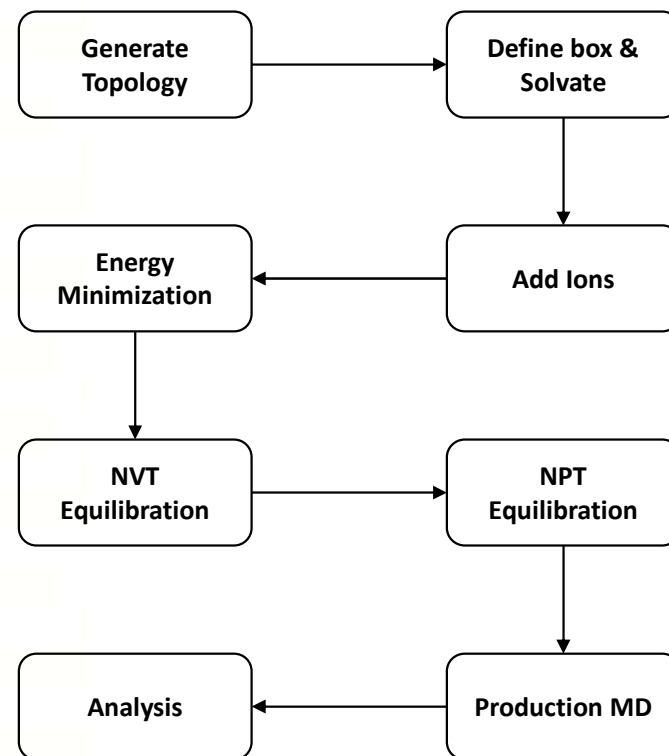
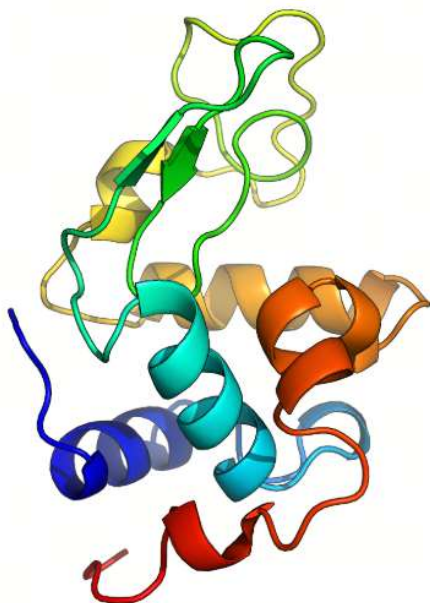
# Lysozyme in Water

<http://www.mdtutorials.com/gmx/lysozyme/index.html>

## GROMACS Tutorial

### Lysozyme in Water

Justin A. Lemkul, Ph.D.  
*Virginia Tech Department of Biochemistry*



This example will guide a new user through the process of setting up a simulation system containing a protein (lysozyme) in a box of water, with ions. Each step will contain an explanation of input and output, using typical settings for general use.

## 主要内容一：轨迹后处理

```
gmx trjconv -f md.xtc -s md.tpr -o md_centered_nopbc.xtc -pbc mol -ur compact -center
```

- -pbc mol, 整条链会被平移回主盒子, 保持分子结构的连续性
- -ur compact: 平移后分子质心尽可能靠近盒子中心 (默认推荐)

执行命令以后需要做2次选择

需要选择把谁居中, 且保证分子完整性 (不跨越周期性边界)

```
Select group for centering
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

需要选择输出哪部分原子, 这里我只输出了蛋白, 意味着水和离子都会从轨迹里被删除

```
Select group for output
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

```
gmx trjconv -f md_centered_nopbc.xtc -s md.tpr -o fit.xtc -fit rot+trans
```

- -fit rot+trans: 对齐到参考结构, 消除整体平动/转动

执行命令以后需要做2次选择

选择谁作为参考结构来消除平动和转动

```
Select group for least squares fit
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

选择输出哪部分原子

```
Select group for output
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

## 主要内容二: gromacs 轨迹分析命令

<https://manual.gromacs.org/documentation/current/reference-manual/analysis/analysis.html>

Gromacs 文档内提供了大量用于轨迹分析的命令, 包括rmsd计算, rmsf计算, 氢键分析, 二面角分析等

### gmx hbond

#### Synopsis

```
gmx hbond [-f [<.xtc/.trr/...>]] [-s [<.tpr/.gro/...>]] [-n [<.ndx>]]  
  [-o [<.ndx>]] [-num [<.xvg>]] [-dist [<.xvg>]]  
  [-ang [<.xvg>]] [-dan [<.xvg>]] [-b <time>] [-e <time>]  
  [-dt <time>] [-tu <enum>] [-fgroup <selection>]  
  [-xvg <enum>] [-[no]rmpbc] [-[no]pbc] [-sf <file>]  
  [-selrpos <enum>] [-seltype <enum>] [-r <selection>]  
  [-t <selection>] [-[no]m] [-[no]pf] [-cutoff <real>]  
  [-hbr <real>] [-hba <real>] [-de <string>] [-ae <string>]
```

<https://manual.gromacs.org/documentation/current/onlinehelp/gmx-hbond.html#gmx-hbond>

### Hydrogen bonds

`gmx hbond`

The program `gmx hbond` analyzes the *hydrogen bonds* (H-bonds) between all possible donors D and acceptors A. To determine if an H-bond exists, a geometrical criterion is used, see also Fig. 58:

$$\begin{aligned} r &\leq r_{HB} = 0.35 \text{ nm} \\ \alpha &\leq \alpha_{HB} = 30^\circ \end{aligned} \quad (471)$$

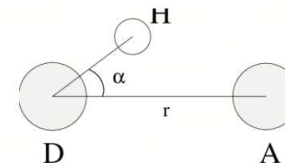
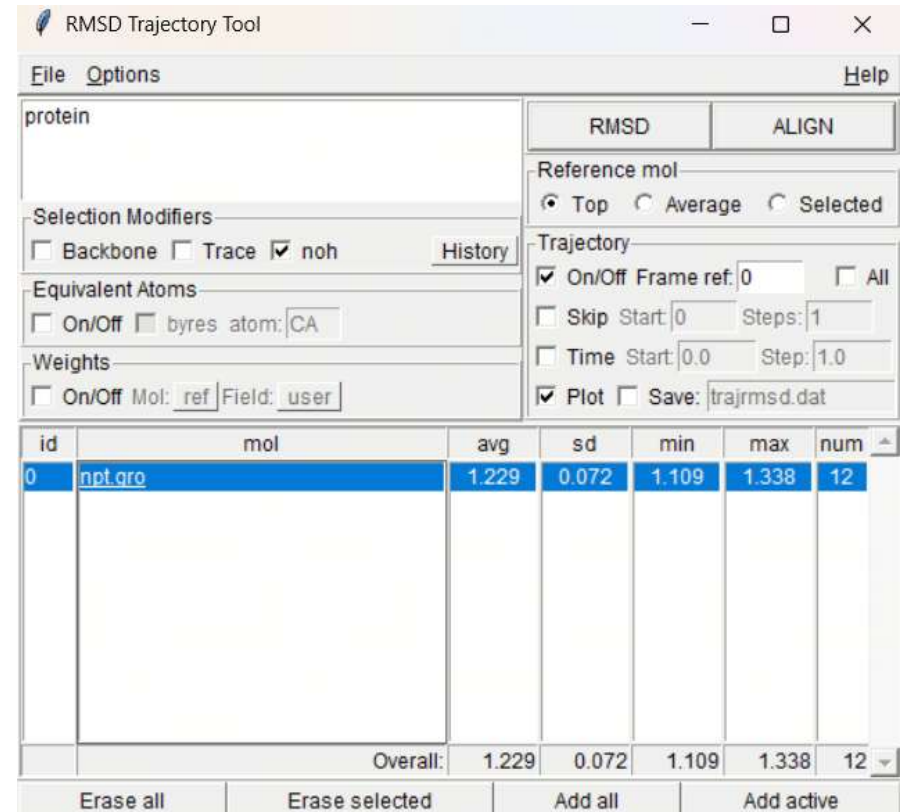
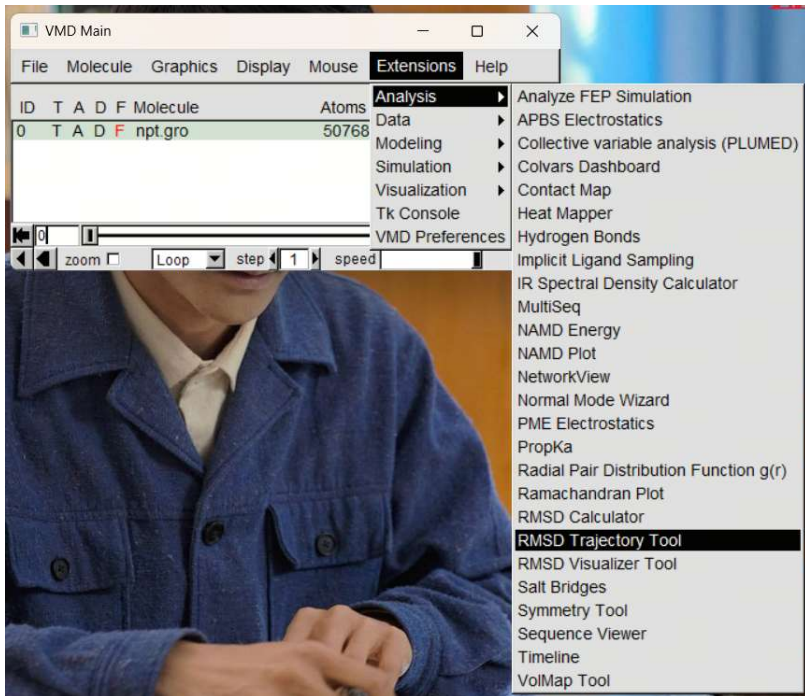


Fig. 58 Geometrical Hydrogen bond criterion.

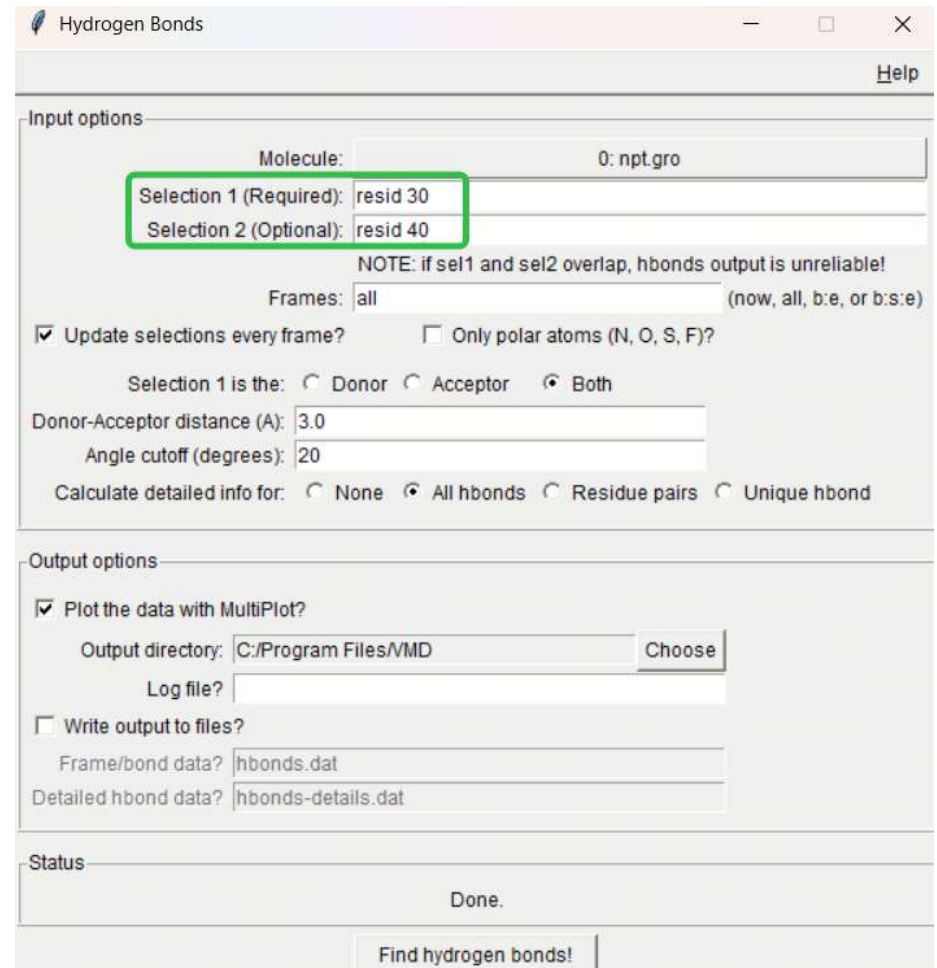
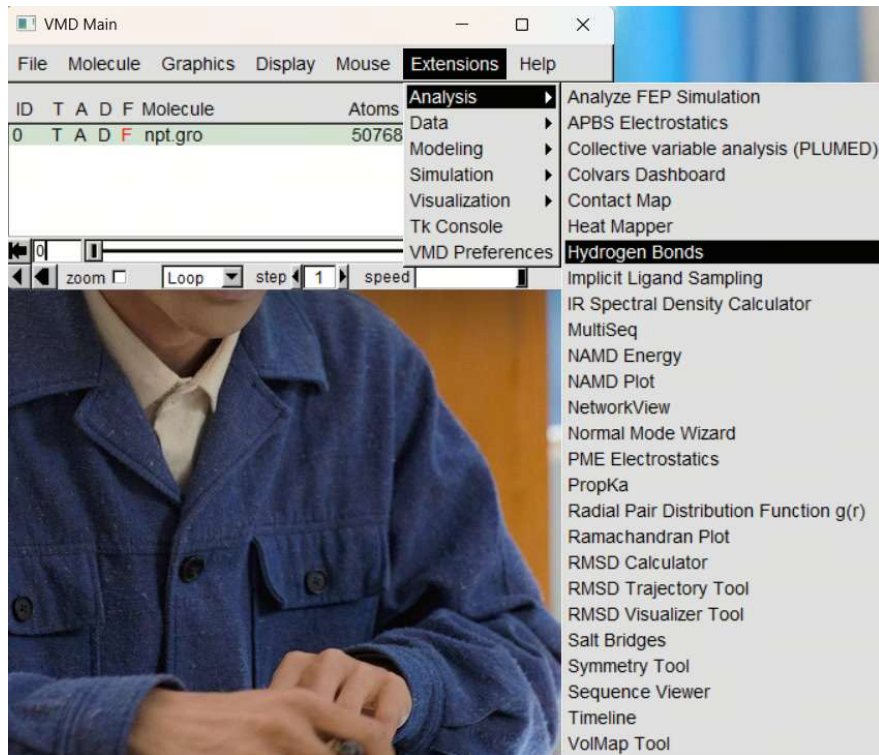
<https://manual.gromacs.org/documentation/current/reference-manual/analysis/hydrogen-bonds.html>

## 主要内容三：使用VMD计算rmsd





## 主要内容四：使用VMD计算氢键



## 主要内容五：VMD选择语言

<http://sobereva.com/504>

有一些关键词可以直接选择特定原子，以下举例一部分：

all: 所有原子

none: 不选择任何原子

noh: 氢以外的原子（即重原子）

ion: 离子

water: 水

backbone: 生物大分子骨架

sidechain: 生物大分子侧链

protein: 蛋白

nucleic: 核酸

helix: 螺旋

alpha\_helix: alpha螺旋（是helix中的子集，较长一段螺旋才算）

sheet: 折叠

turn: 转角

coil: 盘绕

alpha: 蛋白质的alpha碳

acidic: PH=7时带负电氨基酸

basic: PH=7时带正电氨基酸

charged: acidic和basic的并集

neutral: 电中性氨基酸

polar: 极性残基

hydrophobic: 疏水性残基

bonded: 成键的原子

hetero: 非蛋白质和核酸的部分

carbon、hydrogen、oxygen、nitrogen、sulfur: 相应元素。对于其它元素没法这么输入元素名来选择

用下面这些关键词可以通过属性选取原子，都是后面要接参数的

name: 原子名。例：name OW选择原子名叫OW的原子

index: 原子序号（从0开始！）。例：index 4

serial: 原子序号（从1开始）

type: 原子类型。例：type CA选择CA原则类型

element: 元素名。例：element P选择磷原子

resname: 残基名。例：resname ALA代表选择丙氨酸

residue: 残基编号，从0开始。例resid 372代表选择372号残基

resid: 残基编号，从1开始。若结构文件里有残基号则与之一致

chain: 链名。例：chain B代表选择B链

fragment: 片段编号。VMD对每个键连的片段自动设定一个编号。例：fragment 4代表选择片段4

numbonds: 成键数目。例：numbonds=2或numbonds 2代表选形成了两个键的原子

structure: 二级结构。例：structure H代表选择螺旋(helix)区域

x,y,z: X/Y/Z笛卡尔坐标

vx,vy,vz: X/Y/Z方向速度

beta: pdb文件中的beta值

occupancy: pdb文件中的原子占有率

mass: 原子质量

charge: 原子电荷

phi、psi: 蛋白质骨架角度

radius: 原子半径

...等等

within 6 of protein: 距离蛋白质6埃以内的原子

not within 5 of resname ADP: 距离名为ADP的分子5埃以外的原子

water within 5 of residue 8 to 44: 距离8~44号残基5埃以内的水

withinbonds 2 of index 31: 距离编号为31原子的两个键及以内的原子

maxringsize 6 from protein: 蛋白当中所有六元及六元以下环上的原子

same resname as resid 33: 所有与33号残基相同名称的残基

same residue as {protein within 5 of nucleic}: 与核酸的原子相距5埃以内的蛋白的原子，并且把被截断的残基保留完整

x > 15 and not same fragment as {exwithin 8 of protein}: 蛋白质以及蛋白质8埃范围以外的原子，保留完整片段，同时x坐标得大于15埃

# Download VMD


<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

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

NIH RESOURCE FOR MACROMOLECULAR MODELING & VISUALIZATION | UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Type Keywords

THEORETICAL and COMPUTATIONAL  
BIOPHYSICS GROUP



Home

Research

Publications

Software

Instruction

News

Galleries

Facilities

About Us

Home

Overview

Publications

Research

Software

▶ VMD Molecular Graphics Viewer

▶ NAMD Molecular Dynamics Simulator


▶ BioCoRE Collaboratory Environment

▶ MD Service Suite

▶ Structural Biology Software Database

▶ Computational Facility

Outreach



VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting. VMD supports computers running MacOS X, Unix, or Windows, is distributed free of charge, and includes source code.  
([more details...](#))

Spotlight

In 2017, the Royal Swedish Academy of Sciences awarded the Nobel Prize in Chemistry to [Jacques Dubochet](#), [Joachim Frank](#), and [Richard Henderson](#) "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution". We are pleased to celebrate this great triumph for structural biology along with the well-deserved recognition of the Center's long-time collaborator and friend, Joachim Frank. Our center has a long tradition in developing computational methods that enable scientists to build atomistic models of biomolecules. [Molecular Dynamics Flexible Fitting](#) (MDFF), a method developed in close collaboration with Joachim Frank and his group, reconciles high resolution data from X-ray crystallography and functional information from cryo-electron microscopy (cryo-EM). MDFF utilizes molecular dynamics to "naturally" fit each atom into a cryo-EM map. In less than a decade since its development, MDFF has proved instrumental in studying biomolecular systems. A selected list of publications employing MDFF both by our group and others can be found [here](#).  
**Other Spotlights**

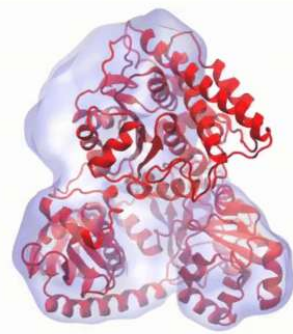


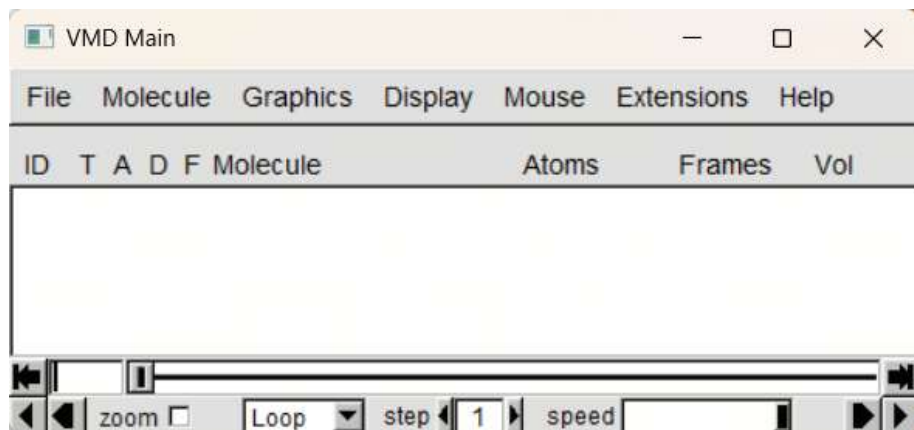
image size:  
made with **VMD**

在绝大多数情况下，我们都使用VMD软件查看模拟轨迹

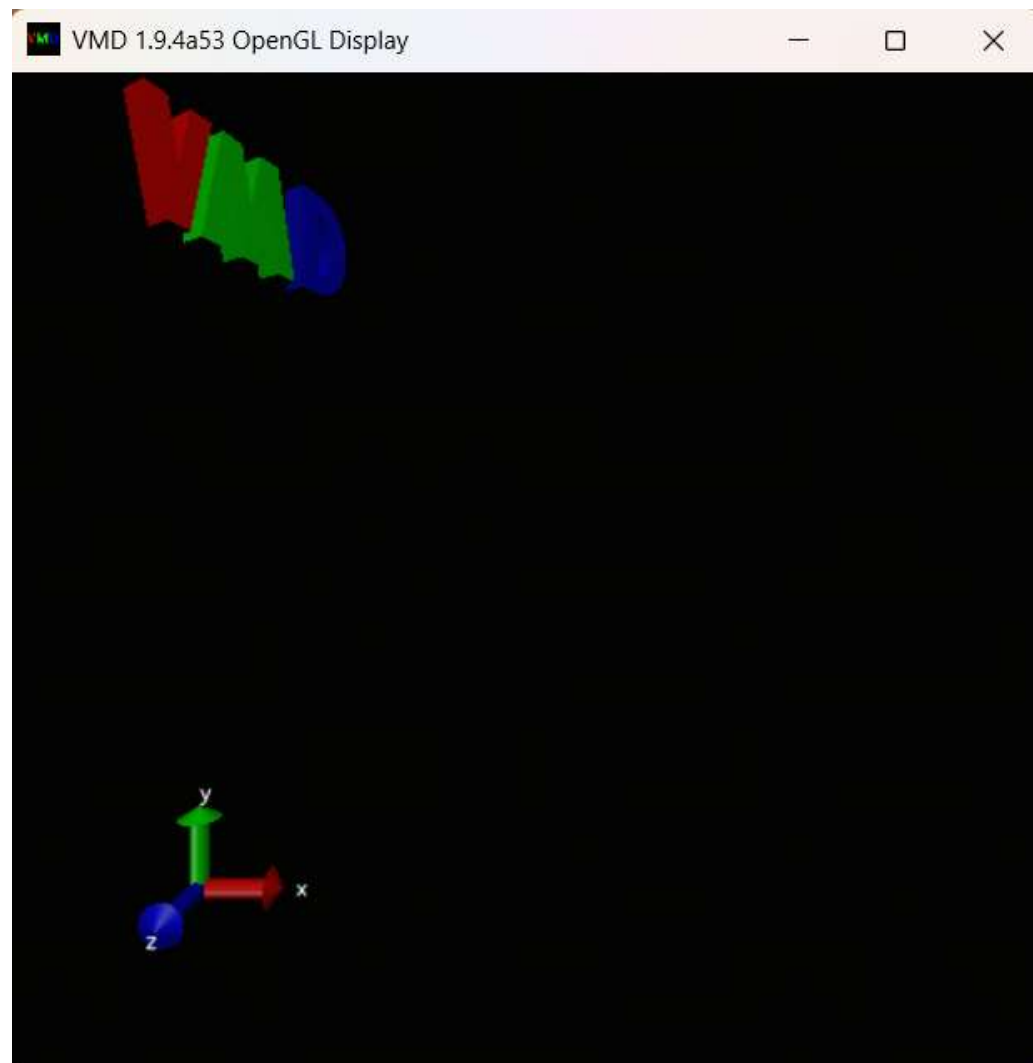


## Main Windows of VMD

<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

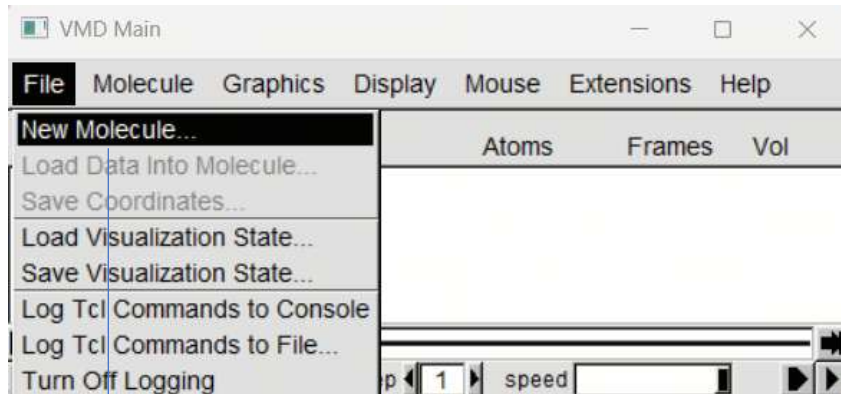


打开VMD，会出现2个窗口，一个进行各种操作的主窗口，一个显示图像的展示窗口

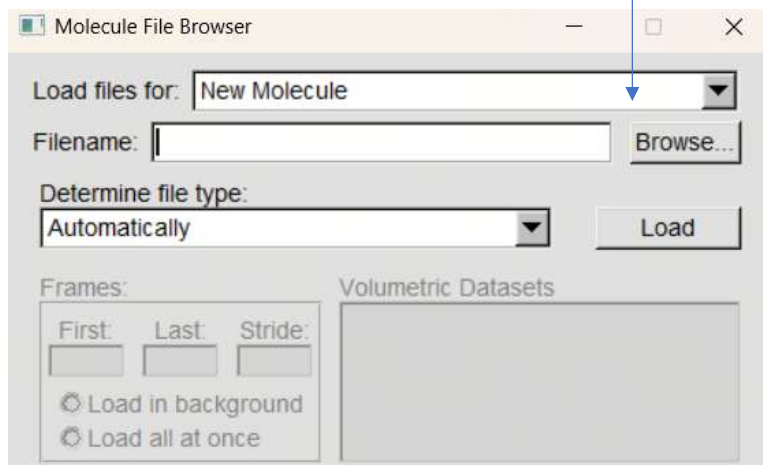


## Load System into VMD

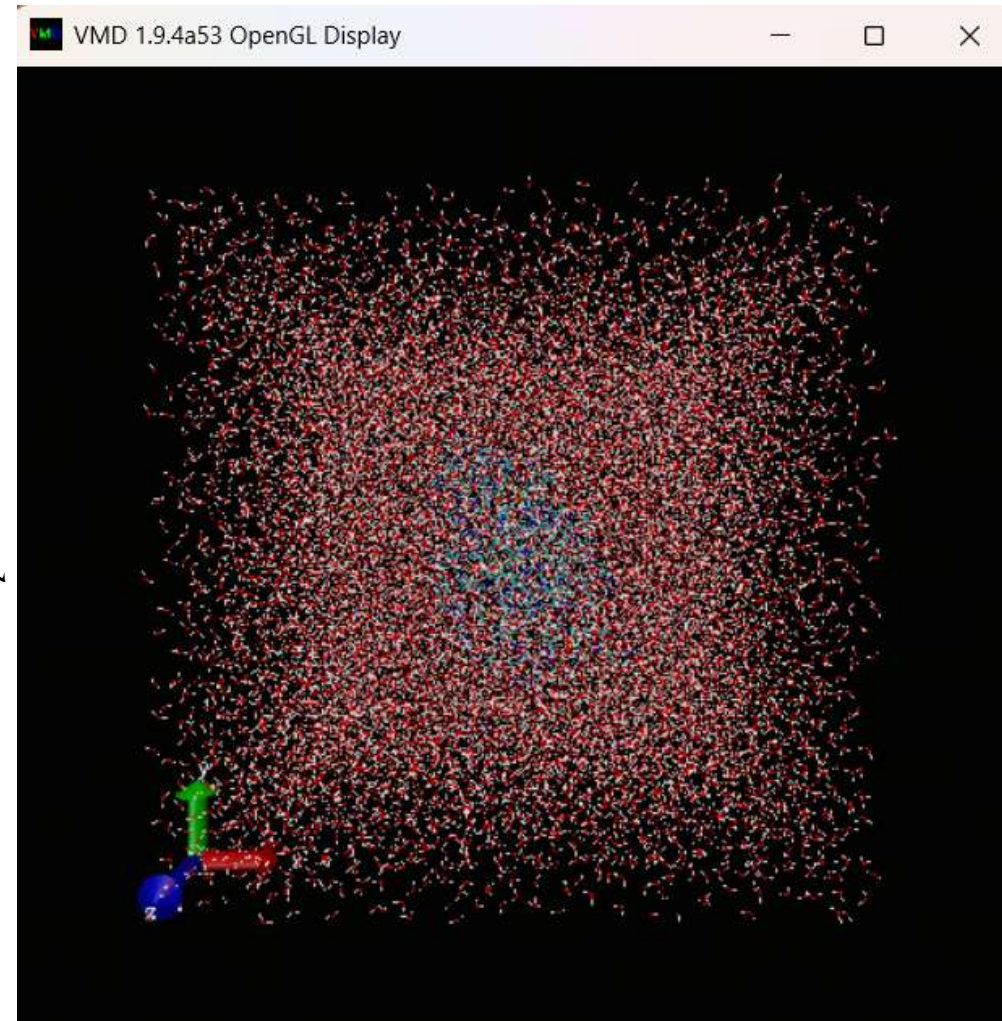
<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>



File – New Molecule – Browse



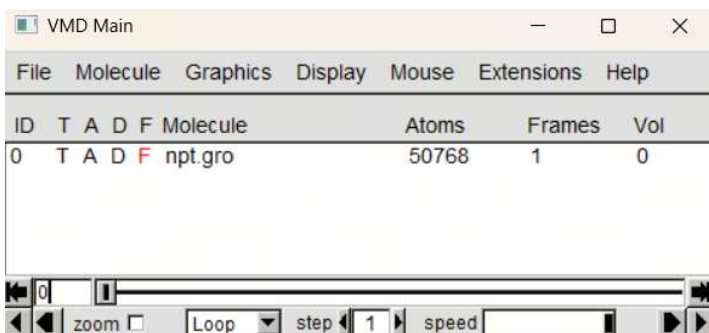
导入pdb/gro等格式的文件，比如我们之前模拟得到的npt.gro/md.gro



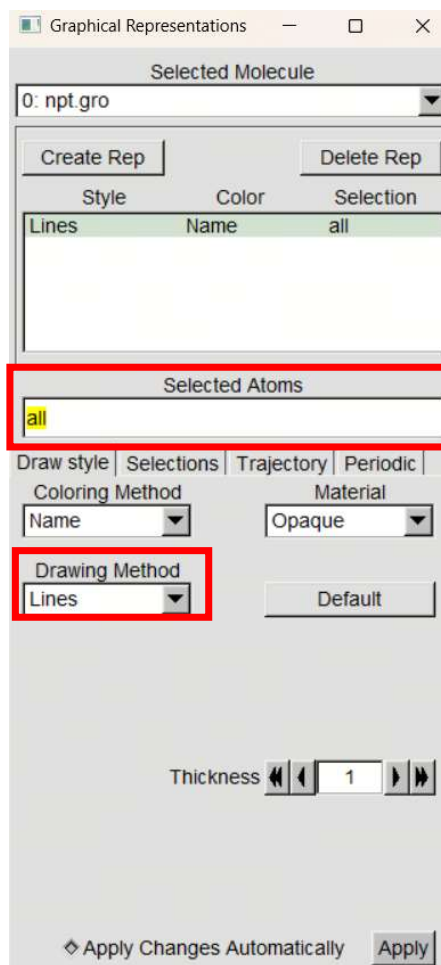
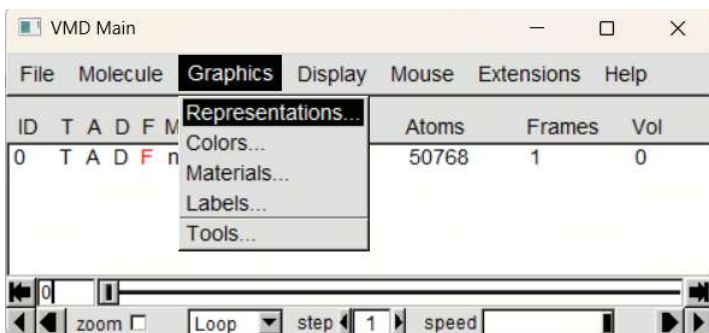
# VMD Basic

<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

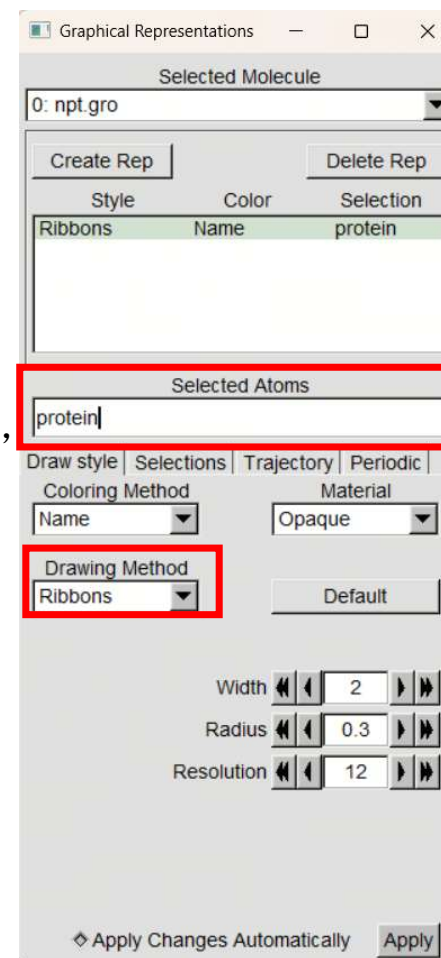
做完上一页的操作以后，主窗口出现了我们导入的体系



Graphic – Representations 打开显示控制



修改红框部分，  
敲击回车，观察VMD显示窗口变化



默认显示

# VMD select language

<http://sobereva.com/504>

有一些关键词可以直接选择特定原子，以下举例一部分：

- all: 所有原子
- none: 不选择任何原子
- noh: 氢以外的原子（即重原子）
- ion: 离子
- water: 水
- backbone: 生物大分子骨架
- sidechain: 生物大分子侧链
- protein: 蛋白
- nucleic: 核酸
- helix: 螺旋
- alpha\_helix: alpha螺旋（是helix中的子集，较长一段螺旋才算）
- sheet: 折叠
- turn: 转角
- coil: 盘绕
- alpha: 蛋白质的alpha碳
- acidic: PH=7时带负电氨基酸
- basic: PH=7时带正电氨基酸
- charged: acidic和basic的并集
- neutral: 电中性氨基酸
- polar: 极性残基
- hydrophobic: 疏水性残基
- bonded: 成键的原子
- hetero: 非蛋白质和核酸的部分
- carbon、hydrogen、oxygen、nitrogen、sulfur: 相应元素。对于其它元素没法这么输入元素名来选择

用下面这些关键词可以通过属性选取原子，都是后面要接参数的

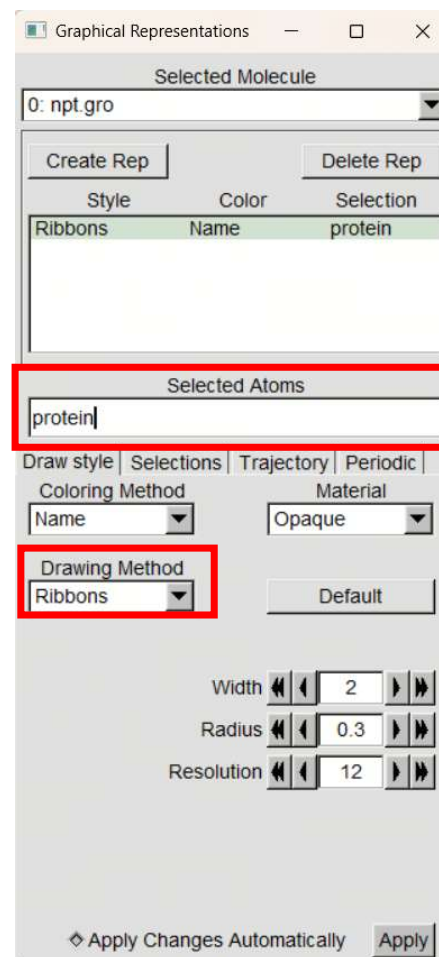
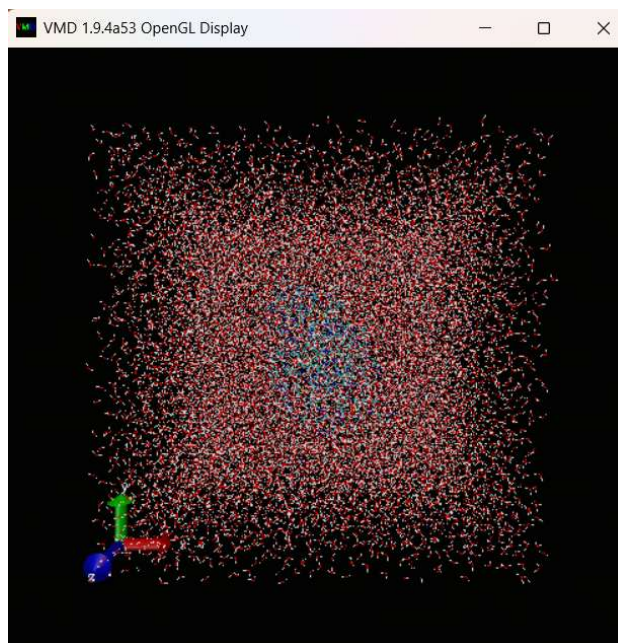
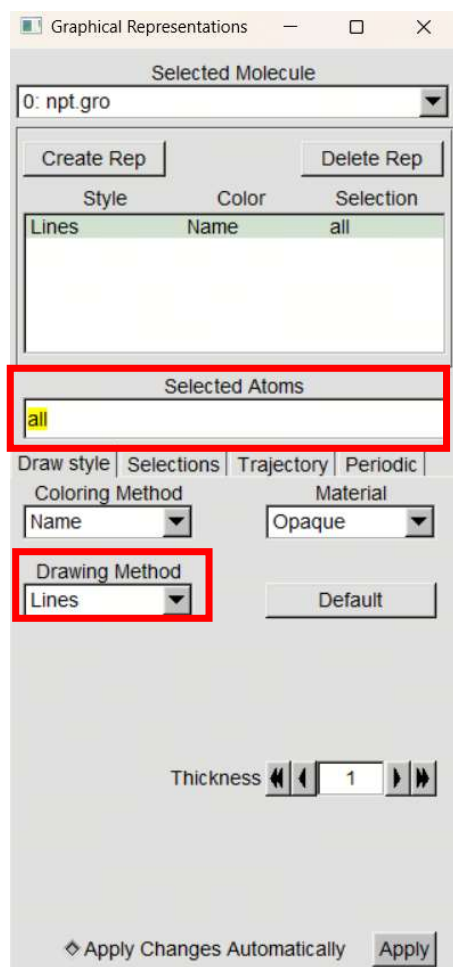
- name: 原子名。例：name OW选择原子名叫OW的原子
- index: 原子序号（从0开始！）。例：index 4
- serial: 原子序号（从1开始）
- type: 原子类型。例：type CA选择CA原则类型
- element: 元素名。例：element P选择磷原子
- resname: 残基名。例：resname ALA代表选择丙氨酸
- residue: 残基编号，从0开始。例resid 372代表选择372号残基
- resid: 残基编号，从1开始。若结构文件里有残基号则与之一致
- chain: 链名。例：chain B代表选择B链
- fragment: 片段编号。VMD对每个键连的片段自动设定一个编号。例：fragment 4代表选择片段4
- numbonds: 成键数目。例：numbonds=2或numbonds 2代表选形成了两个键的原子
- structure: 二级结构。例：structure H代表选择螺旋(helix)区域
- x,y,z: X/Y/Z笛卡尔坐标
- vx,vy,vz: X/Y/Z方向速度
- beta: pdb文件中的beta值
- occupancy: pdb文件中的原子占有率
- mass: 原子质量
- charge: 原子电荷
- phi、psi: 蛋白质骨架角度
- radius: 原子半径
- ...等等

- within 6 of protein: 距离蛋白质6埃以内的原子
- not within 5 of resname ADP: 距离名为ADP的分子5埃以外的原子
- water within 5 of residue 8 to 44: 距离8~44号残基5埃以内的水
- withinbonds 2 of index 31: 距离编号为31原子的两个键及以内的原子
- maxringsize 6 from protein: 蛋白当中所有六元及六元以下环上的原子
- same resname as resid 33: 所有与33号残基相同名称的残基
- same residue as {protein within 5 of nucleic}: 与核酸的原子相距5埃以内的蛋白的原子，并且把被截断的残基保留完整
- x > 15 and not same fragment as {exwithin 8 of protein}: 蛋白质以及蛋白质8埃范围以外的原子，保留完整片段，同时x坐标得大于15埃



# VMD Basic

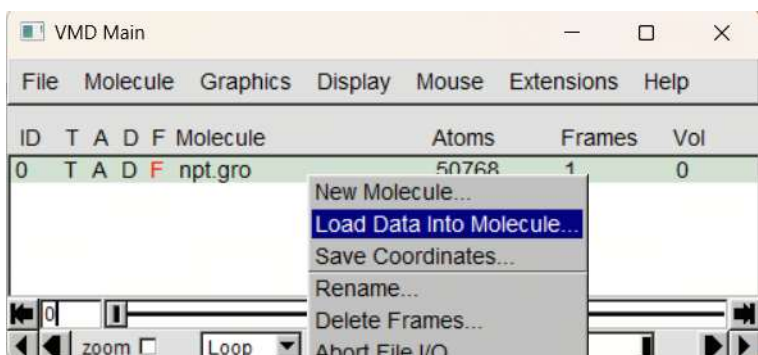
<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>



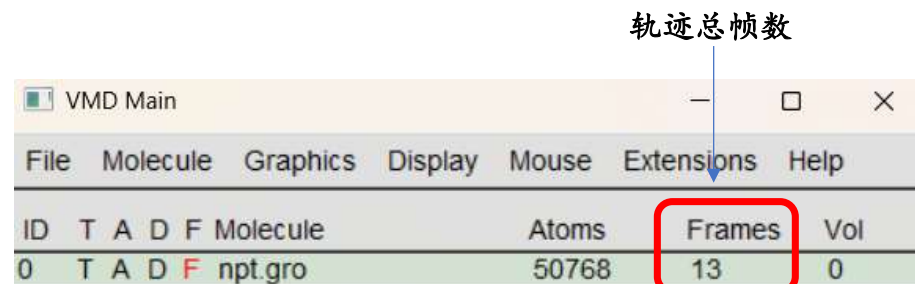
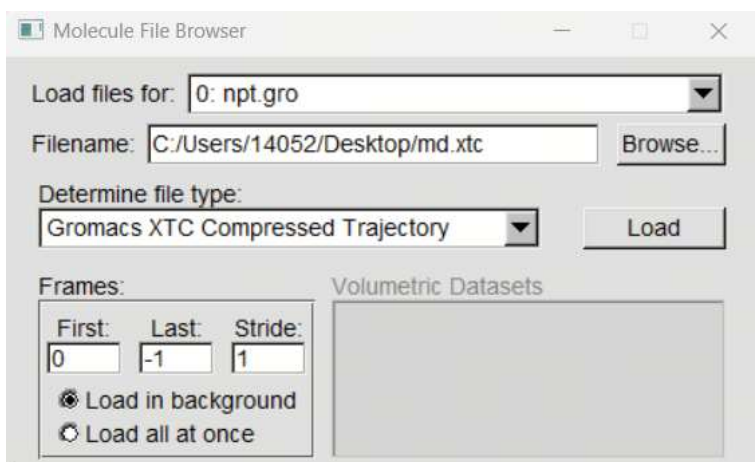
# Load Trajectories into VMD

<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

在npt.gro这一行上点击鼠标右键，选择Load Data into Molecule



点击Browse，选择轨迹文件，比如  
npt.trr / npt.xtc / md.xtc 等，点击Load



11是当前展示窗口为第11帧

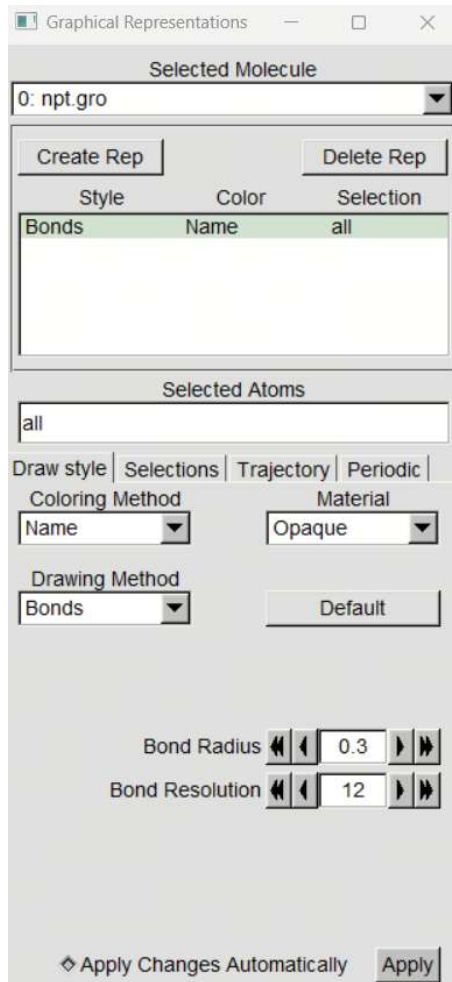


倒放轨迹

正放轨迹

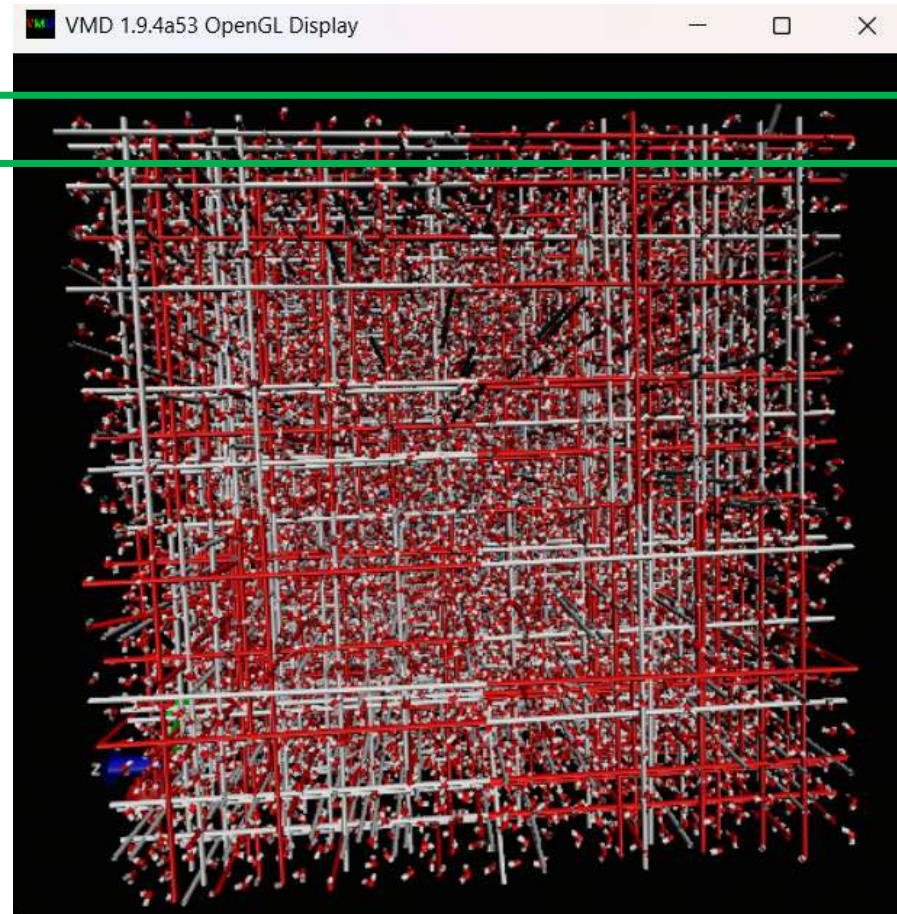


# Periodic Boundary Conditions



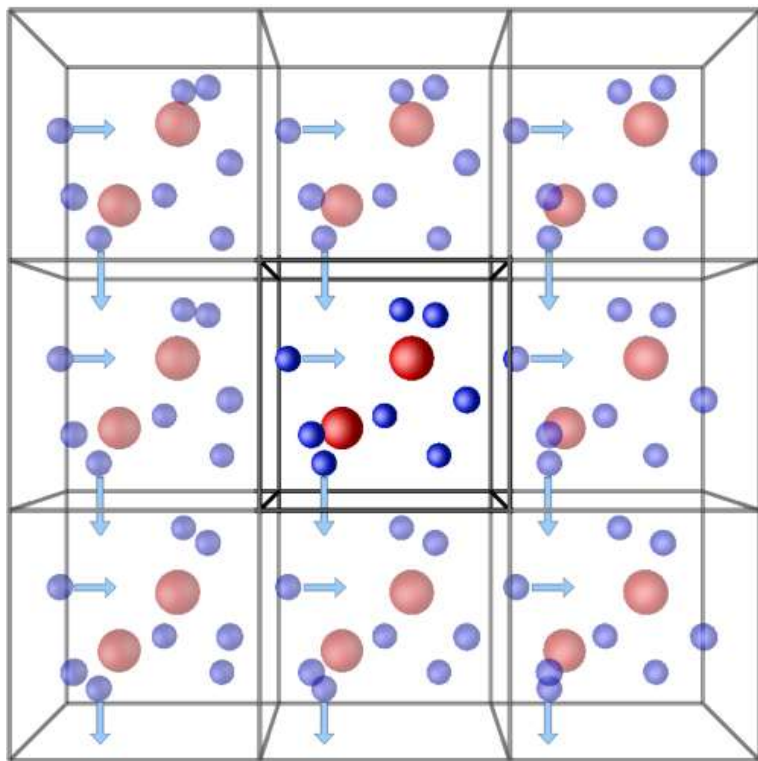
此时，如果把水展示出来，展示方式改成Bonds，即按化学键展示

你可能发现，为什么有的化学键这么长，比如绿框出来的键，横跨了水盒子，看起来一个氢原子在最左边，对应的氧原子在最右边，这是因为周期性边界条件

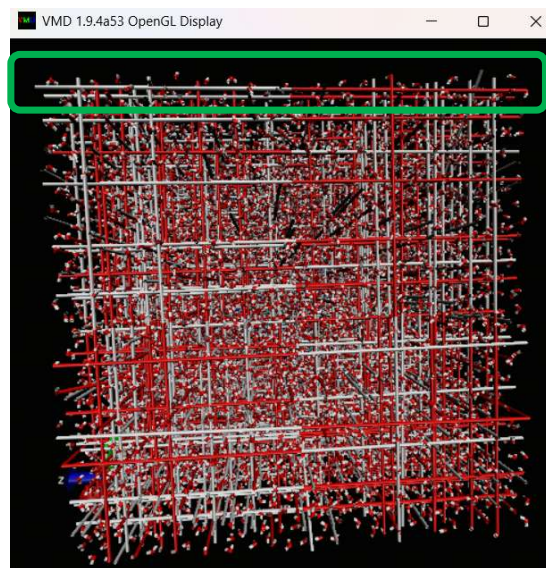




# Periodic Boundary Conditions



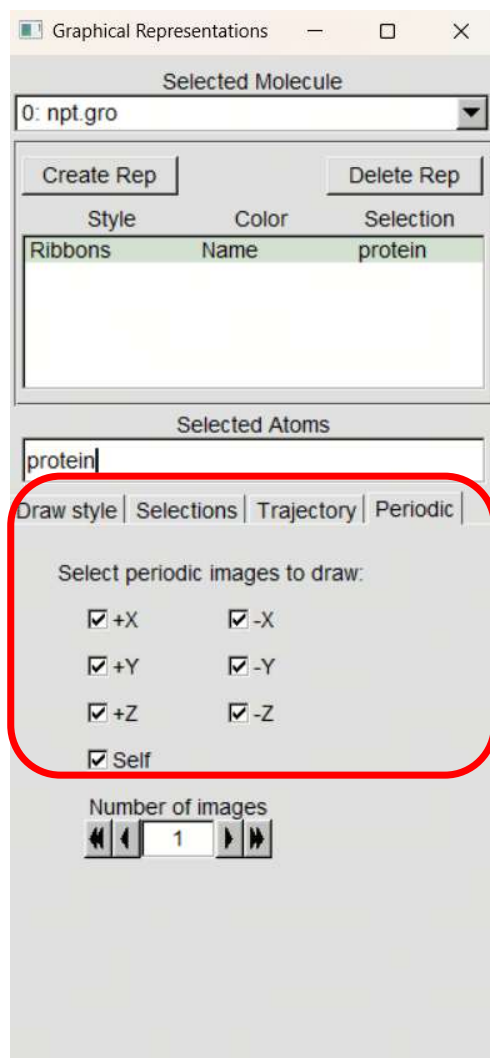
- 问题背景：**实际模拟的体系通常仅包含数千到数百万个原子，远小于宏观物质的尺寸。若直接模拟孤立体系，边界附近的粒子会因缺少相邻粒子而出现非物理行为（如表面张力、粒子逃逸等）。
- PBC解决方案：**将模拟盒子（如立方体）在三维空间无限重复，形成一个周期性排列的体系。当粒子从一个方向移出盒子时，会从相反方向重新进入。



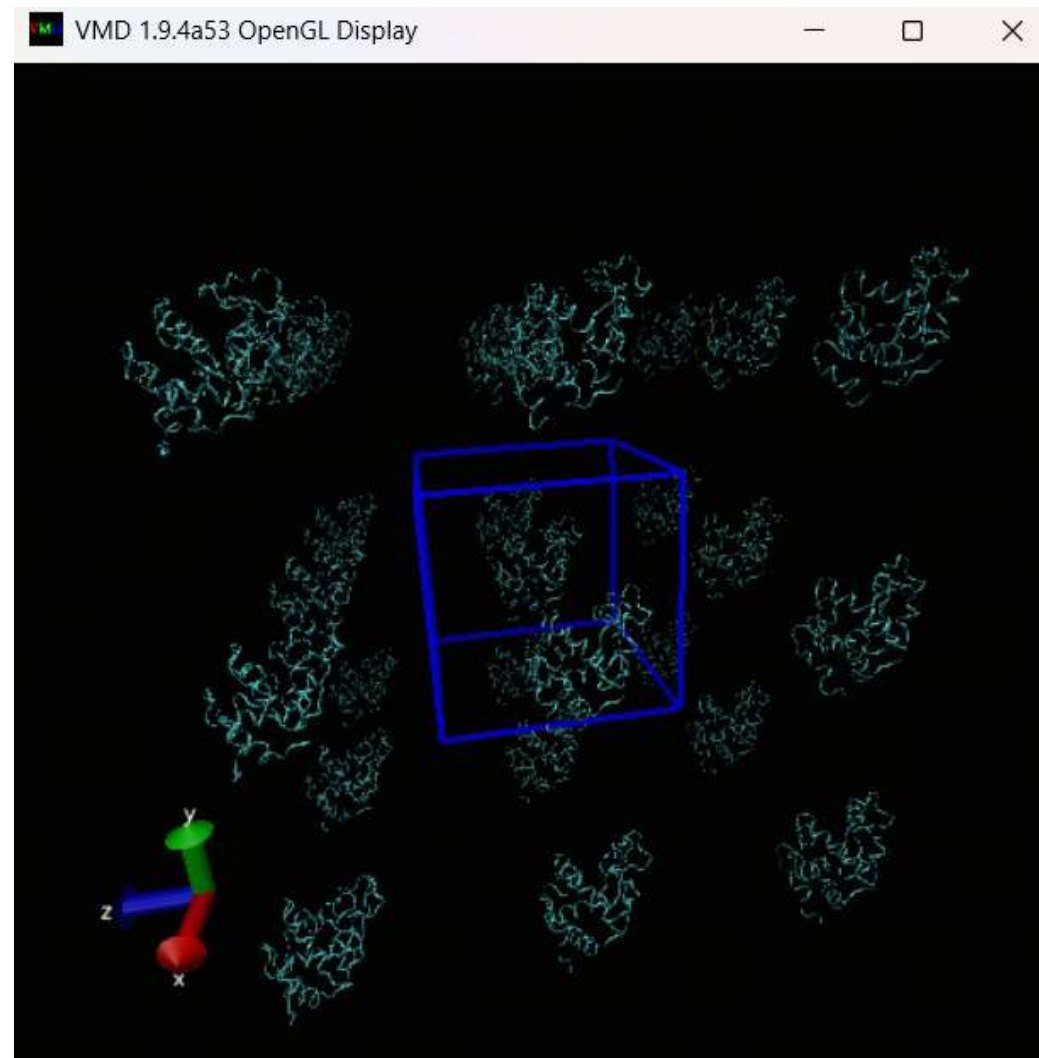
所以绿框处的水正好在盒子的边界，它的一个氢原子已经从右边离开了盒子，通过周期性边界条件，把这个氢原子从盒子的左边又放了进来，造成视觉上这个水横跨了整个盒子



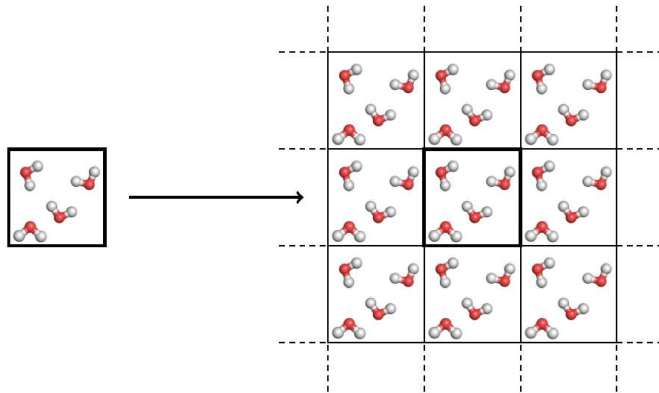
# Periodic Boundary Conditions



在Periodic选项卡把所有附近的周期性镜像打开，能看到有意思的东西



# Long Range Coulomb Interactions



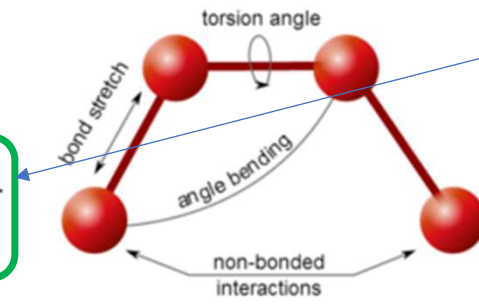
$$V(r) = \sum_{\text{bonds}} k_b(b - b_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2$$

$$+ \sum_{\text{dihedrals}} k_\phi(1 + \cos(n\phi - \phi_0)) + \sum_{\text{impropers}} k_\psi(\psi - \psi_0)^2$$

$$- \sum_{\text{non-bonded pairs}(i,j)} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{\text{non-bonded pairs}(i,j)} \frac{q_i q_j}{\epsilon_D r_{ij}}$$

范德华相互作用

静电相互作用



- 非键作用一般分为静电作用和范德华作用两部分，计算时不仅仅考虑单个周期内，理论上应该计算所有镜像之间的非键作用，但实际操作中当然不可能。
- 最简单的做法是截断，就是当距离超过一个cutoff后，就不计算这两个粒子之间的相互作用了。
- 对范德华作用，其随着距离衰减非常快，所以一般用截断法足矣，vdwtype=Cut-off, rvdw=1.0, 距离超过1nm, 范德华力不算。
- 对静电相互作用，只截断会有很多误差，一般分成短程和长程两部分，短程（距离小于rcoulomb）直接计算，长程部分这里使用PME来进行计算，这是一种比简单的截断精确很多的方法

```

;-----
; 非键相互作用参数
;-----
cutoff-scheme      = Verlet ; Verlet截断方案
nstlist             = 20    ; 邻居列表更新频率
rlist               = 1.0    ; 邻居列表截断半径
vdwtype             = Cut-off ; Amber标准范德华截断
rvdw                = 1.0    ; 范德华截断半径
coulombtype         = pme    ; PME处理长程静电
rcoulomb            = 1.0    ; 直接空间静电截断（与rvdw对齐）
;-----
    
```

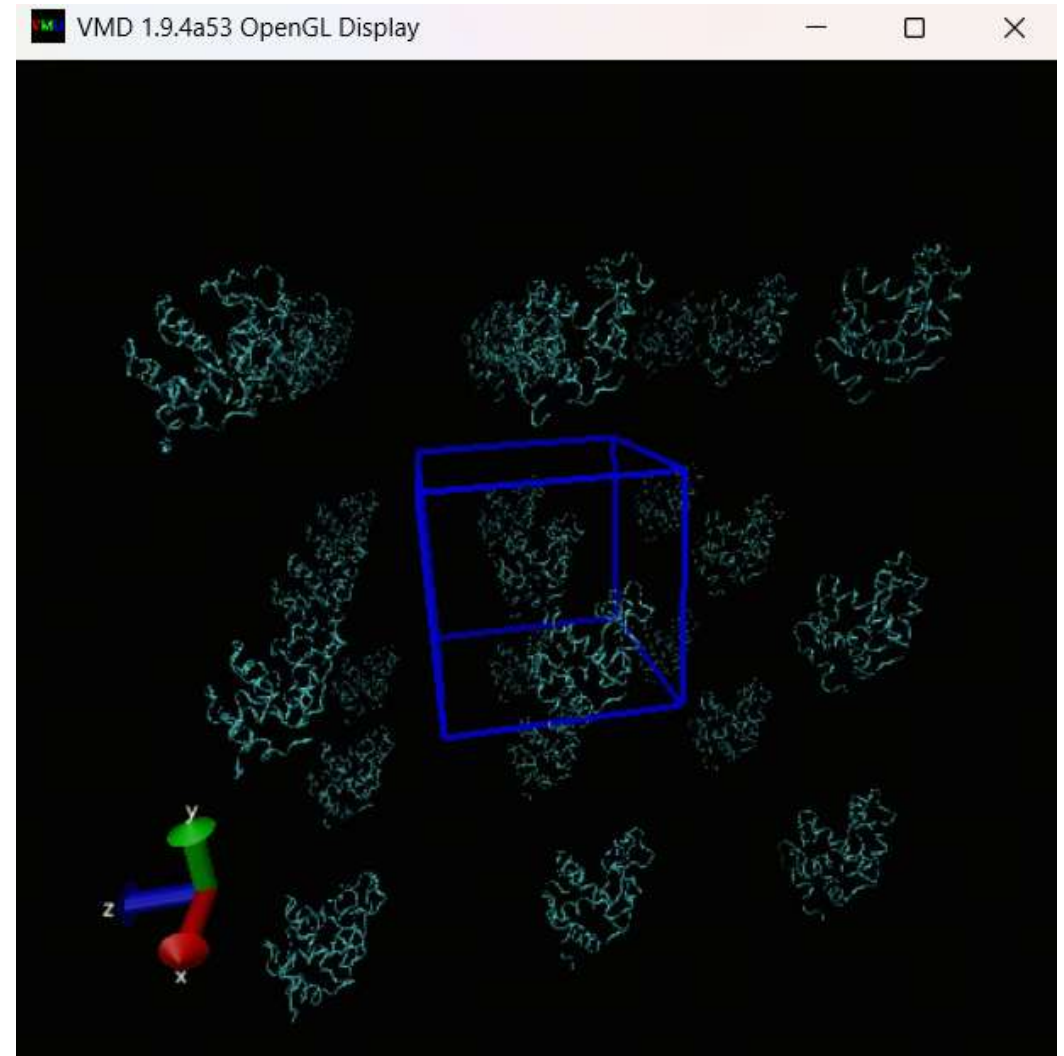
[http://micro.stanford.edu/mediawiki/images/4/46/Ewald\\_notes.pdf](http://micro.stanford.edu/mediawiki/images/4/46/Ewald_notes.pdf)

<https://courses.physics.illinois.edu/phys466/sp2013/lnotes/Ewald/EwaldLecture.pdf>

## Post Process of MD Trajectories

在观察轨迹的时候，有时候蛋白质也会跨越镜像，非常影响观察。因此通常我们在得到轨迹以后，要做一些后处理。

- 解除周期性边界条件，使得蛋白质保持完整
- 对齐蛋白，消除平动和转动
- 删除不那么关心的原子，比如膜蛋白模拟中的膜，比如有的时候我们对水不关心



## Post Process of MD Trajectories

```
gmx trjconv -f md.xtc -s md.tpr -o md_centered_nopbc.xtc -pbc mol -ur compact -center
```

- -pbc mol, 整条链会被平移回主盒子, 保持分子结构的连续性
- -ur compact: 平移后分子质心尽可能靠近盒子中心 (默认推荐)

执行命令以后需要做2次选择

需要选择把谁居中, 且保证分子完整性 (不跨越周期性边界)

```
Select group for centering
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

需要选择输出哪部分原子, 这里我只输出了蛋白, 意味着水和离子都会从轨迹里被删除

```
Select group for output
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

```
gmx trjconv -f md_centered_nopbc.xtc -s md.tpr -o fit.xtc -fit rot+trans
```

- -fit rot+trans: 对齐到参考结构, 消除整体平动/转动

执行命令以后需要做2次选择

选择谁作为参考结构来消除平动和转动

```
Select group for least squares fit
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```

选择输出哪部分原子

```
Select group for output
Group 0 (      System) has 50768 elements
Group 1 (      Protein) has 1960 elements
Group 2 (    Protein-H) has 1001 elements
Group 3 (      C-alpha) has 129 elements
Group 4 (      Backbone) has 387 elements
Group 5 (    MainChain) has 517 elements
```



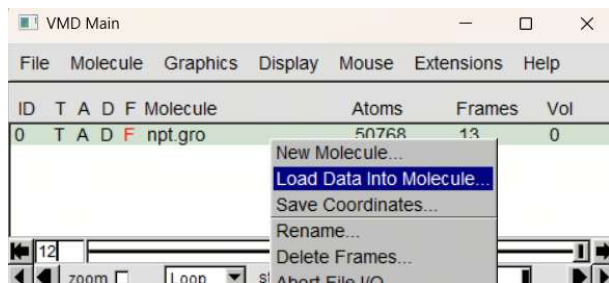
## Load into VMD

<https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

执行完上一页的命令以后，我们得到了fit.xtc，此轨迹：

- 去除了周期性边界条件，将蛋白放到了盒子中心
- 消除了蛋白的平动和转动，删除了体系中非蛋白的所有原子（此例子中是水和离子）

此时，如果我们想将此轨迹fit.xtc导入之前的npt.gro，VMD会报错，因为fit.xtc里只有蛋白，没有npt.gro里那些水和离子的轨迹。此时，我们需要准备一个只有蛋白的结构文件，且与fit.xtc对应。



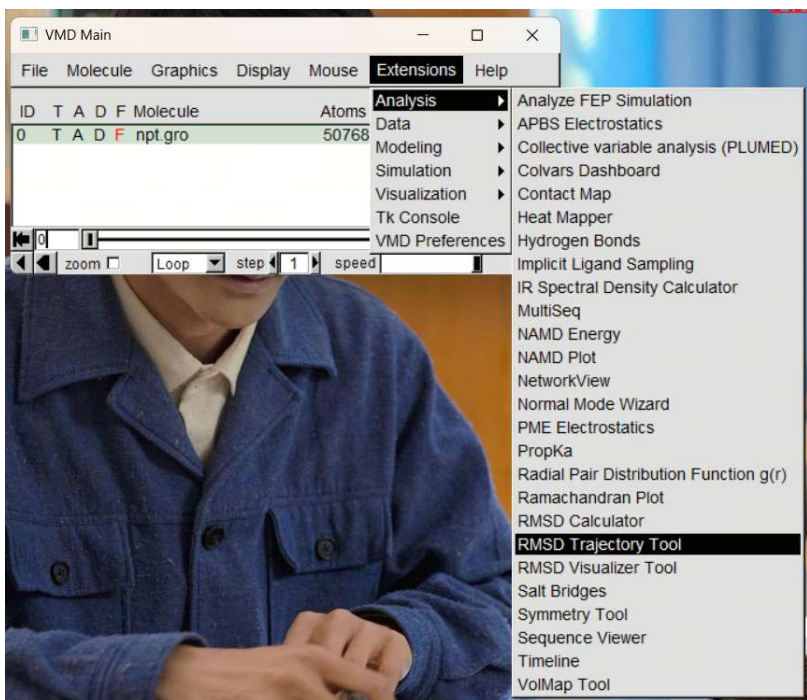
```
INFO) Finished with coordinate file C:/Users/14052/Desktop/fit.xtc.  
ERROR) BaseMolecule: attempt to init atoms while structure building in progress!  
ERROR) Invalid number of atoms in file: 1960  
Info) Using plugin xtc for coordinates from file C:/Users/14052/Desktop/fit.xtc  
ERROR) Incorrect number of atoms (1960) in  
ERROR) coordinate file C:/Users/14052/Desktop/fit.xtc  
ERROR) Mismatch between existing molecule or structure file atom count and coordinate or trajectory file atom count.  
Info) Finished with coordinate file C:/Users/14052/Desktop/fit.xtc.
```

我们用记事本打开npt.gro，手动把蛋白之外的部分全删除，把文件第二行的原子数改成剩余的数目，存储即可。此文件就可以用来载入fit.xtc了。

文件	编辑	查看
16355SQL	OW50636	3.760 6.763 3.541 0.0159 0.3609 0.0089
16355SQL	HW150637	3.759 6.707 3.619 1.3342 -2.0205 -1.5400
16355SQL	HW250638	3.669 6.771 3.516 -0.2541 -0.7389 0.5564
16410NA	NA50713	1.276 2.543 7.534 0.3977 0.1125 0.2732
16411NA	NA50714	2.381 2.030 0.686 -0.2468 0.0102 -0.3763
16412CL	CL50715	3.355 5.000 1.077 -0.4429 -0.3598 0.1964
16413CL	CL50716	1.124 0.294 0.592 0.2360 -0.1963 0.3131

文件	编辑	查看
LYSOZYME in water		
1960		
1LYS	N	1 4.717 3.754 2.791 -0.1441 -0.4398 0.2784
1LYS	H1	2 4.688 3.806 2.873 3.4024 0.9889 0.7778

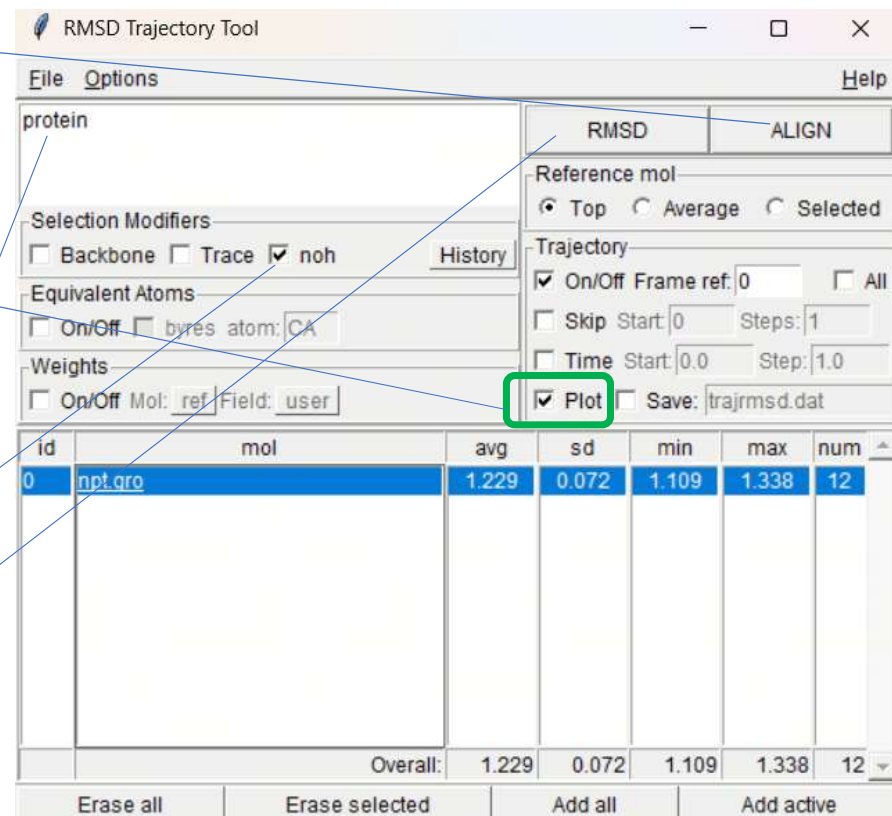
# Calculate RMSD



首先对齐

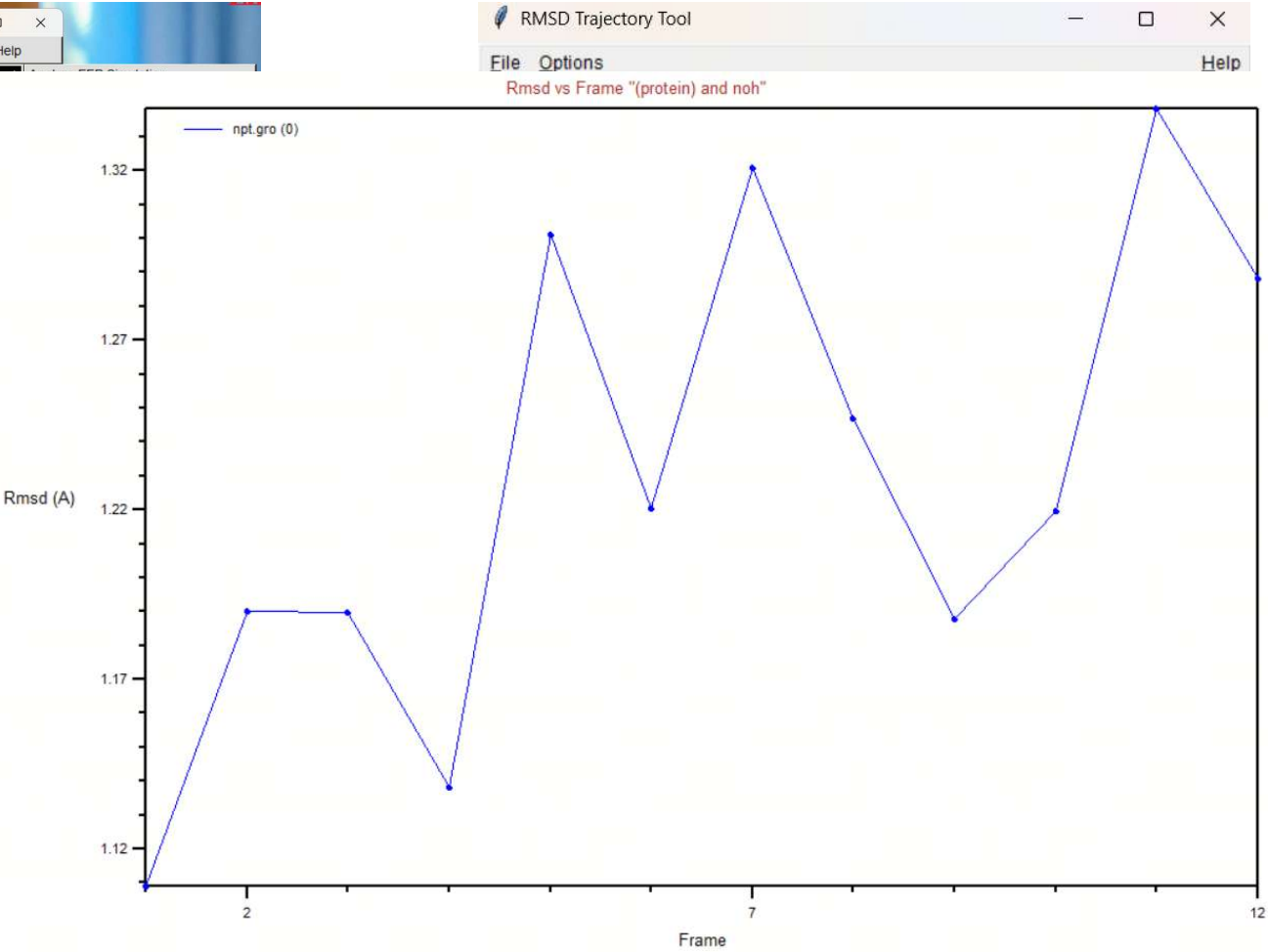
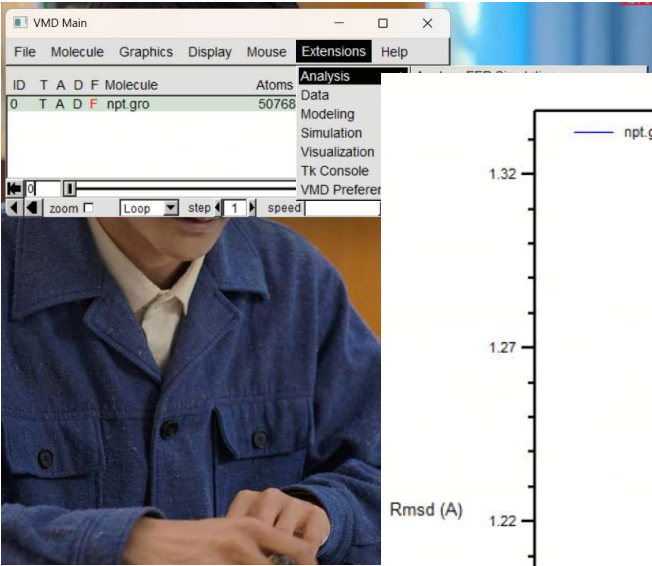
记得勾上

计算蛋白质，  
且非氢部分  
的RMSD



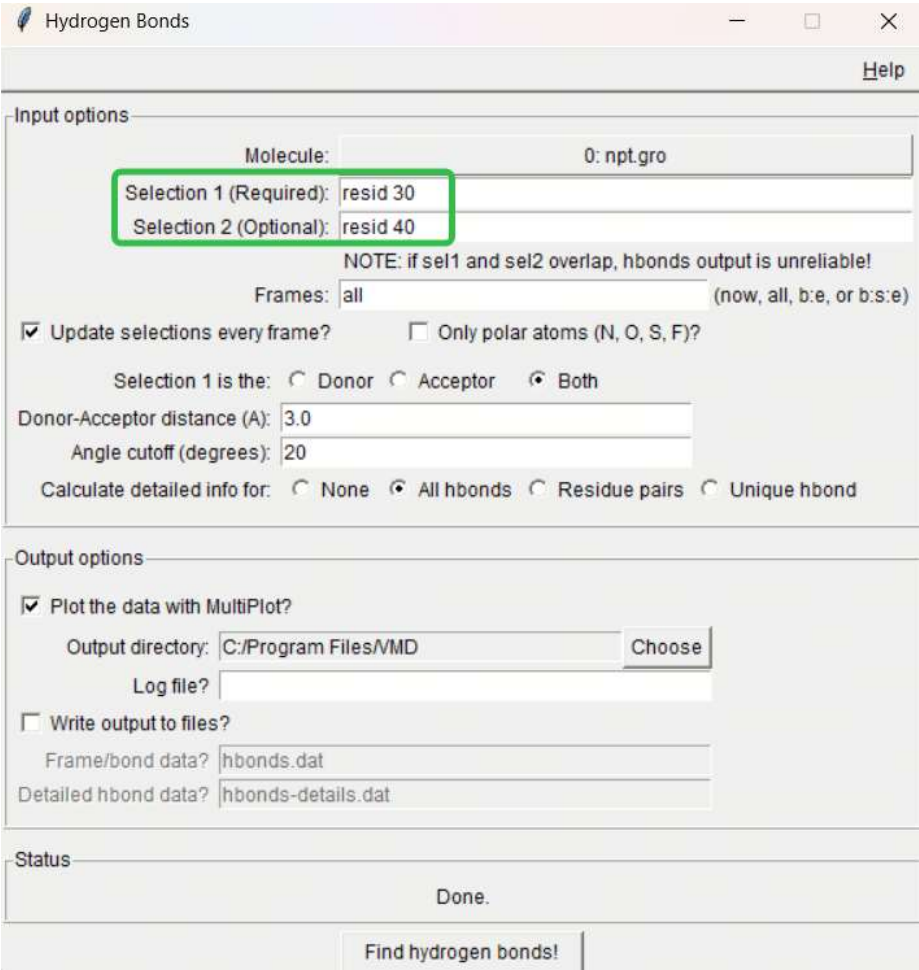
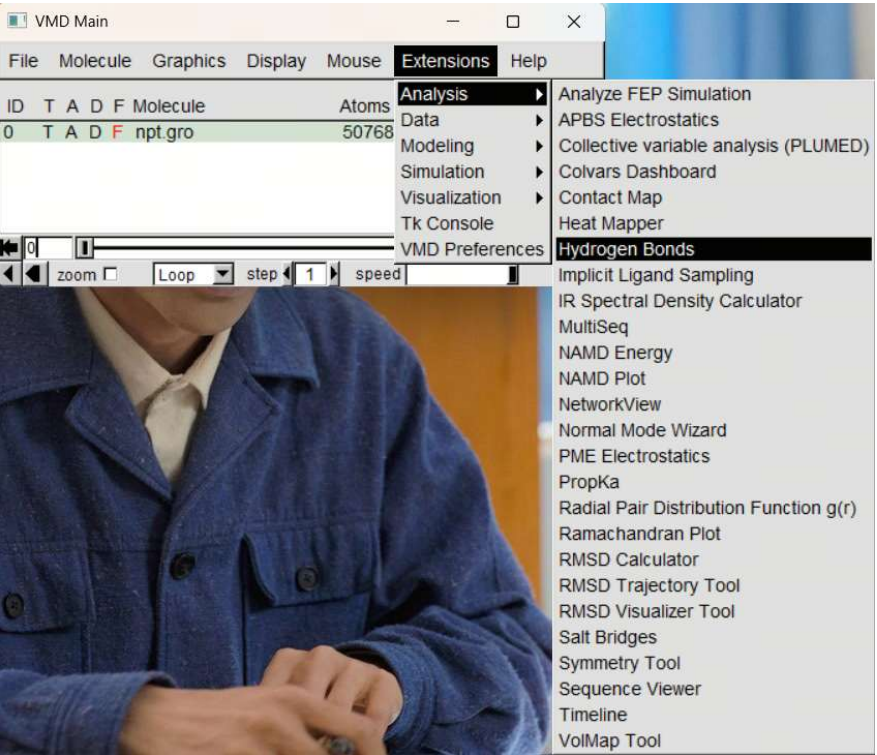
实际情况按需求来，一般计算rmsd的场合都是蛋白backbone的，所有C alpha的；  
或者小分子配体复合物时，小分子的。

# Calculate RMSD



# Calculate HBond

假如我想知道在模拟每一帧中第30号残基和第40号残基之间氢键作用情况



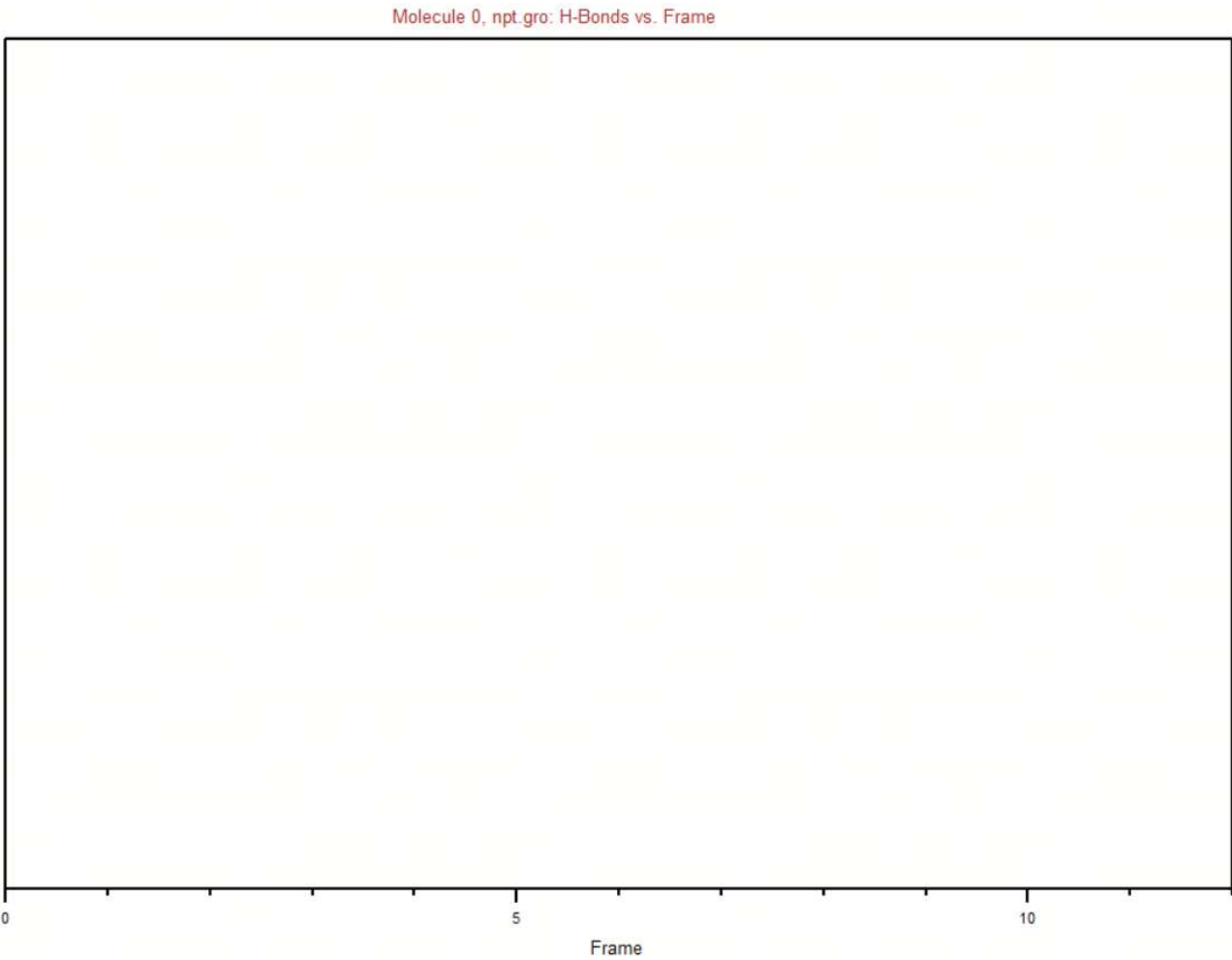


# Calculate HBond

Oh no, 它俩之间一根氢键都没有



No. Bonds



Help

pt.gro

hbonds output is unreliable!  
(now, all, b:e, or b:s:e)  
, O, S, F)?

ue pairs ☐ Unique hbond

Choose