根据Hugo的excel从pdb下载两个复合体的pdb文件，检查后发现文件无需CHARMM-GUI修改；在服务器上下载gromacs2023，将pdb文件与Hugo提供的改良版force field放到working directory，使用pdb2gmx将pdb文件转换为gromacs文件，命令报错，在Hugo添加的SEP, TPO and PTR residue基础上向residuetypes.dat文件中加入MSE（硒代甲硫氨酸）后仍然报错，此FF中没有关于MSE的拓扑结构，故在pdb文件中把MSE改为MET（现在只改了D），修改后SE原子仍然存在，修改太麻烦，看看能不能用pdb上面的其他D；先用C进行尝试，但仍然不行，尝试使用其他的FF

Residue 2 named ARG of a molecule in the input file was mapped to an entry in the topology database, but the atom CG used in that entry is not found in the input file. Perhaps your atom and/or residue naming needs to be fixed.

难以找到含MSE的FF，想试试看能不能通过CHARMM-GUI PDB Reader把MSE改成MET，但这一步还是需要MSE的结构文件，不如直接尝试在FF的结构库里加上MSE的结构；OK很难操作MSE的结构，但发现CHARMM-GUI可以在PDB Reader里默认把所有engineered residues都改回正常残基并在此基础上加磷酸化，选项选了pH7.0，Renaming Engineered Residues，Terminal group patching，Phosphorylation

为什么会有六个不同的链？

Pdb2gmx报错：Atom OT1 in residue THR 219 was not found in rtp entry THR with 14 atoms while sorting atoms；解决方法：将所有OT1、OT2都改为O，把原文件中PTR的位置都从TYR改回PTR等，把氨基酸原子标号改成与FF的aminoacids.rtp中一致

在网上找了一个ions.mdp，后续可以改参数

gmx\_mpi genion -s D\_ions.tpr -o D\_solv\_ions.gro -p topol.top -pname NA -nname CL -neutral所选组为12 SOL，Replacing 11 solute molecules in topology file (topol.top) by 11 NA and 0 CL ions

在网上找了一个em.mdp，后续可以改参数

gmx\_mpi grompp -f em.mdp -c D\_solv\_ions.gro -p topol.top -o em.tpr:

The largest distance between excluded atoms is 7.626 nm between atom 6986 and 6998, which is larger than the cut-off distance. This will lead to missing long-range corrections in the forces and energies. If you expect that minimization will bring such distances within the cut-off, you can ignore this warning.

A black screen with white text

Description automatically generated

Test2用了Hugo给的mdp文件，Fmax > 1000，后续可能会出问题

A screenshot of a computer

Description automatically generated

gmx\_mpi grompp -f ions.mdp -c SHP2\_SH2\_CagA\_EPIYA\_D\_CHARMM\_mod\_solv.gro -p topol.top -o ions.tpr：

NOTE 1 [file ions.mdp]:

With Verlet lists the optimal nstlist is >= 10, with GPUs >= 20. Note

that with the Verlet scheme, nstlist has no effect on the accuracy of

your simulation.

将ions.mdp和minim.mdp中的nstlist改为10

gmx\_mpi grompp -f nvt.mdp -c em.gro -r em.gro -p topol.top -o nvt.tpr：

ERROR 1 [file nvt.mdp]:

The largest distance between excluded atoms is 1.141 nm between atom 6517

and 6543, which is larger than the cut-off distance. This will lead to

missing long-range corrections in the forces and energies.

尝试把water box建大来解决这个问题

和Hugo谈过后发现问题：下载的pdb文件里有多个蛋白，应该现根据网页上的信息用CHARMM-GUI挑出合适的蛋白再跑分析

从头开始

整个PTPN11（两个SH2）：

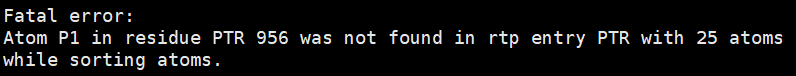
5X94的pdb文件中有两个PTPN11（共四个SH2）和四个与之结合的CagA EPIYpA。选择unmodeled最小的组合：A, C, E。

首先用CHARMM-GUI的PDB Reader选出A C E链，此过程会把engineered residues改回来，也就是MSE和PTR改为MET和TYR；选择model missing residues: PROA 155 – 164, fail to model missing residues.

手动修改输出文件，把C和E的956位TYR改回PTRA black background with white text

Description automatically generated

将所有OT1, OT2都手动改为O



PTR改为和结构文件相同

将ions.mdp和minim.mdp中的nstlist改为10

A black screen with white text

Description automatically generated

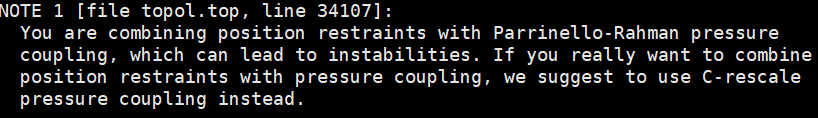
仍然有warning

A black background with white text

Description automatically generated

A screen shot of a computer

Description automatically generated



A black screen with white text

Description automatically generated

0.2 ns结果不理想，尝试2 ns

2 ns结果比较理想，但RMSD较大

PTPN11-N-SH2 + EPIYpA-D:

借助3D模型，预计可以手动在105 ALA与106 ASP中间切断PROA，N端SH2与C链结合，C端SH2与E链结合；A链结束加TER

minim.mdp中的emtol从1000减小到了100

通过RMSD

和gyration的波动程度判断结合的紧密程度

PTPN11-C-SH2 + EPIYpA-D:

2 ns后RMSD仍在上升，尝试模拟4 ns

5x7b没有MSE，CHARMM-GUI可以模拟A链缺失的154-165；手动加上TER，修改B、C链的TYR为PTR，修改PTR的结构与FF中的aminoacids.rtp一致；手动分裂蛋白，N端与B链结合，C端与C链结合，从105、106中间断开

还测试了去磷酸化后的结合强度

D\_EPIYA与SH2的结合可以在2 ns达到稳定，C尝试模拟10 ns

report思路：文章中的结果是Nter与D结合比与C紧密，去磷酸化后Nter与D亲和力降低。以RMSD为指标，先证明这两个结论，把Nter与D以及去磷酸化的D作为positive and negative control。后续通过模拟Cter与C和D的结合，证明Cter与EPIYA的结合没有Nter稳定。