**1 Модифицируем DRY.ndx**

После того как извлекли траектории и получили DRY.ndx его необходимо несколько модифицировать.

**1.1** Во-первых создадим группу, которая будет содержать только тяжёлые атомы лиганды, для

этого пишем:

**gmx make\_ndx -f <dry\_structure>.gro -n DRY.ndx -o DRY.ndx**

Внутри программы пишем: **<номер\_группы\_лиганда>&!aH\***

обычно номер группы лиганда это **13**.

**1.2** Создаем группу, содержащую боковые цепи сайта связывания (LBP) без водородов:

**gmx make\_ndx -f <dry\_structure>.gro -n DRY.ndx -o DRY.ndx**

Внутри пишем:

**r343|r346|r349|r350|r353|r383|r384|r387|r388|r391|r394|r404|r421|r424|r425|r428|r524|r525&9**

Созданные группы можно переименовать внутри программы при при помощи команды вида:

**name <group\_number> <desired\_group\_name>**.

**2**. **Вычисляем RMSD для основной цепи белка (backbone), лиганда без водородов, сайты связывания без водородов**:

**gmx rms -f <trajectory>.xtc -s <structure>.gro -n DRY.ndx -tu ns -o <output>.xvg**

Лиганд выравниванием на **backbone**!!! В остальных случаях выравниваем группы сами на себя.

**3.** Делаем кластеризацию для backbone, лиганда (без Н) и сайта связывания (без Н). О там как это делать читаем здесь: <https://vk.com/@501384604-klasternyi-analiz-v-gromacs-gmx-cluster>

Команда для кластеризации:

**gmx cluster -f <name>.xtc -s <name>.gro -n DRY.ndx -cutoff <num> -b <start\_num> -e <end\_num> -tu ns -dt 0.1 -method gromos -clid <name>\_md\_clid.xvg -dist clusters\_dist -o -g -cl**

**-cutoff** — отсечка в нм(!!!), его значение нужно изменять для получения оптимальной кластеризации. **-b** и **-e** — то с какой и по какую наносекунды (или пс, если возле флага -tu не стоит значение ns) хотим проводить кластеризацию. **-dt** — шаг с которым мы будем идти по траектории, в данном случае анализируется каждый 10-й кадр.

После кластеризации извлекаем репрезентативные структуры для каждого из случаев (осн., цепь лиганд, LBP).

**4**. **Вычисляем энергии связывания (энтальпийные) методом MMGBSA.**

В общем виде команда выглядит так:

**mpirun -np 16 gmx\_MMPBSA -O -i mmpbsa.in -cs <name>.pdb -ci DRY.ndx -cg 1 13 -ct <name>\_MD\_fit.xtc -lm <lig\_name>.mol2 -cp topol.top -o <name>\_MMPBSA.dat -eo <name>\_MMPBSA.csv -do <name>\_DECOMP\_MMPBSA.dat -deo <name>\_DECOMP\_MMPBSA.csv**

Однако надо будет внести некоторые правки в **mmpbsa.in**, а именно изменить значения **startframe=** и **endframe=**  . Участок выбираем согласно графику RMSD лиганда и распределению кластеров в **<name>\_md\_clid.xvg.** Значения указываются в виде нс/10, т.е. если мы хотим рассчитать энергию на участке с 400 по 450 нс, то пишем startframe=40000 и endframe=45000.

Будем брать участок в 50 нс, также можно попробовать рассчитать энергию на нескольких стабильных участках, соответствующих разным кластерам, или даже одному и тому же.

!!!Для запуска расчетов в таком в виде, помимо файлов, указанных в команде понадабятся: atoms.itp, <ligand.itp>, 7uj7.itp.

Кидаем все необходимые файлы в одну папку и запускаем расчёт.

**5.** **Анализ взаимодействий**

**5.1** Для анализа нам понадобится траектория (.xtc) и структура (.tpr). Tpr-файл нам нужен без воды! Для этого открываем файл nvt.mdp и убираем в разделе tc-grps группу Water\_and\_ions, а также соответствующие ей значения tau\_t и ref\_t. Сохраняем под именем test.mdp (или любым) другим. Открываем topol.top и убираем строки, соответствующие воде и иона в **самом низу — раздел [ molecules ], группы SOL, NA, CL и схраняем под именем top.top.**

**5.2** Создаем tpr-файл командой:

**gmx grompp -f test.mdp -c <dry\_structure>.gro -r <dry\_structure>.gro -p top.top -n DRY.ndx -o test.tpr -maxwarn 1**

**5.3** Кидаем траекторию, tpr-файл и interactor.py в отдельную папку.

**5.4** Выбираем для анализа тот же участок, который выбрали для расчетов энергий. Для этого в скрипте в строке 20 прописываем необходимые значения в виде [start:end:step]. По умолчанию скрипт выводит лишь те взаимодействия, которые наблюдались более 50% времени. Но иногда бывает полезно увидеть и более редкие взаимодействия, особенно в случае Н-связей, которых с такой отсечкой можно и не увидеть. Для этого в скрипте находим строки 36 и 43 и изменяем в них значения threshold=0.5 и occ2 > 0.5 на желаемые. Запускаем скрипт!

**PS** Я поправил скрипт и теперь в файл NAME+'\_ints.csv' он записывает правильные номера остатков.