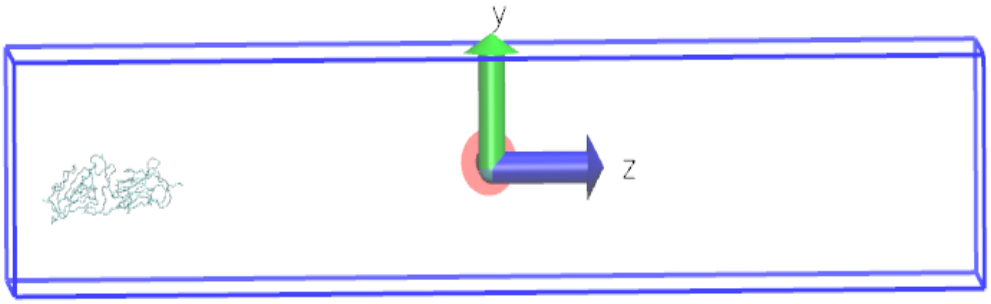
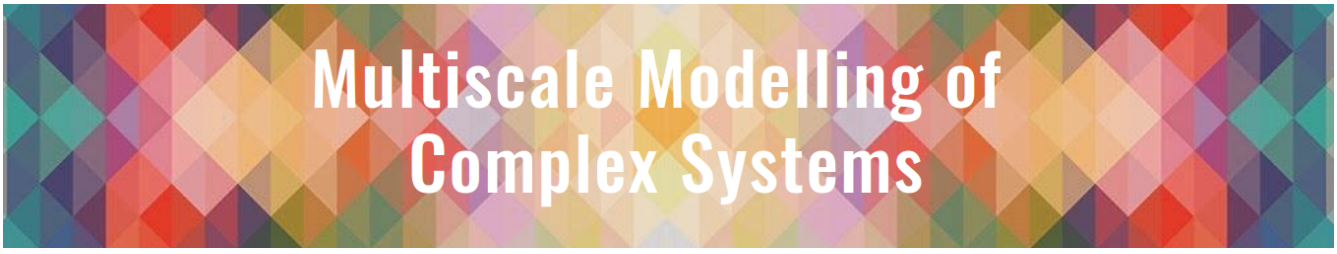
GōMartini
Pulling tutorial 2.0



Version 2. April, 2024
Authored by *Luis F. Cofas-Vargas* and *Adolfo B. Poma*
Institute of Fundamental Technological Research, Polish Academy of Sciences
Department of Biosystems and Soft Matter
PomaLab



Software requirements

Pip
```
>> sudo apt-get -y install python3-pip
```

Anaconda
https://www.anaconda.com/download

PDBFixer
```
>> conda install -c omnia pdbfixer -y
```

Vermouth
```
>> pip install vermouth
```

Martini 3 force field for proteins and other biomolecules
http://cgmartini.nl/index.php/martini-3-0

DSSP
```
>> conda install -c salilab dssp -y
```

GROMACS (version >= 2020)
https://manual.gromacs.org/

Chimera UCSF
https://www.cgl.ucsf.edu/chimera/download.html

Files needed for the tutorial

PDB file: 6ZH9 (file attached)
https://www.rcsb.org/structure/6ZH9

Contact map
http://pomalab.ippt.pan.pl/GoContactMap/

create_gomartini.py script (attached)

Parameter files (attached):
-empty.mdp
-minimization.mdp
-min_vac.mdp
-npt.mdp
-nvt.mdp
pull.mdp

Martini 3 force field:
-martini_v3.0.0.itp
-martini_v3.0.0_ions_v1.itp
-martini_v3.0.0_solvents_v1.itp

Structure preparation

Various software tools can be used to prepare structures, including pymol, vmd, and UCSF Chimera. This step describes a method for doing it with UCSF Chimera.

Open Chimera UCSF

From Favorites, activate command line

From File, Open the pdb file 6ZH9.pdb

This complex contains two chains: nanobody H11-H14 and RBD, named F, and E, respectively.

We must ensure that the numbering of each chain begins with 1. To show the sequence of each chain:

Tools… Sequence… Sequence… Show

To change the number.

Tools… Structure Editing… Renumber residues… (select chain)… starting from 1… OK

Now, we need to align both chains to the axis where we will do the pulling, in this case, the z axis. To accomplish this task, we will select the Calpha of the residues for which we will conduct the pulling (LYS 195 from the E chain and LYS 128 from the F chain)

select: 128.F@CA, select: 195.E@CA

Note: the selection order is very import.

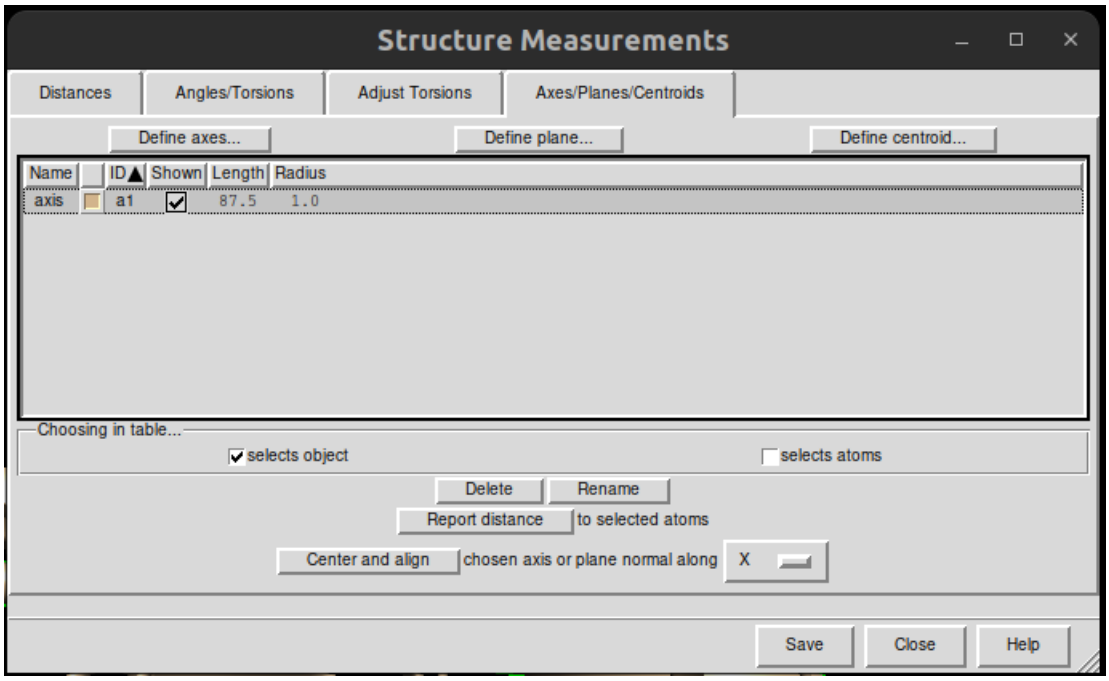
To align to z-axis:

Tools… Structure Analysis… Axes/Planes/Centroids

This will open Structure Measurements menu

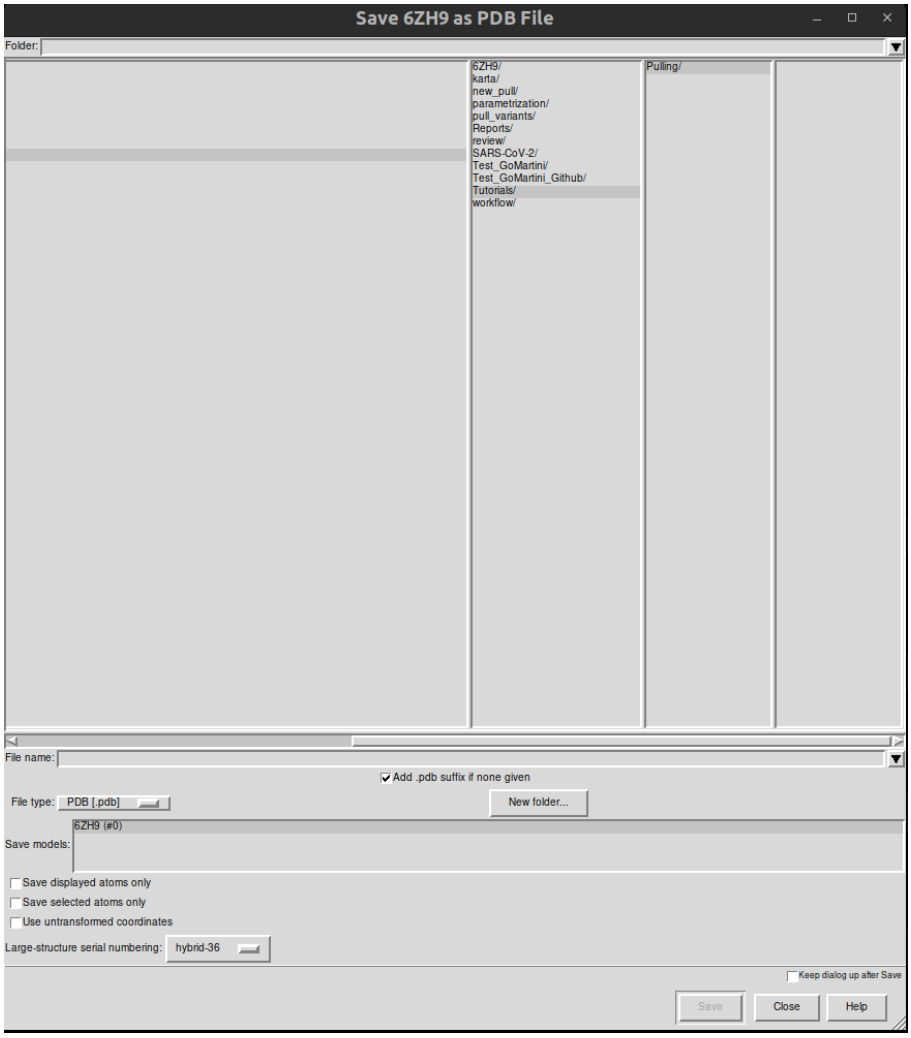
Here, click con Define axes, and then, click con Selected atoms/centroids and then OK. This will draw an axis

In the Structure Measurements menu, click over the axis and then in Center and Align chosen axis or plane normal along, select Z and then, click Center and align. This will shift the structure over the Z-axis.



Next, we need to save the complex with this new coordinates. To do that, click on File.. Save PDB…
Untick the box Use untransformed coordinates, and choose a file name.
In this case, the aligned PDB file was named 6ZH9_aligned.pdb

Note: If you use VMD, Pymol or any other protein structure toolkit you should complete this section with a PDB structure that best represent your system with conventional atom names according to PDB definition and not missing residues.



Before you close, open the structure in a new Chimera session and make sure its aligned correctly. Then, we need to add the missing atoms and hydrogens. Then we can use PDBFixer to do it.

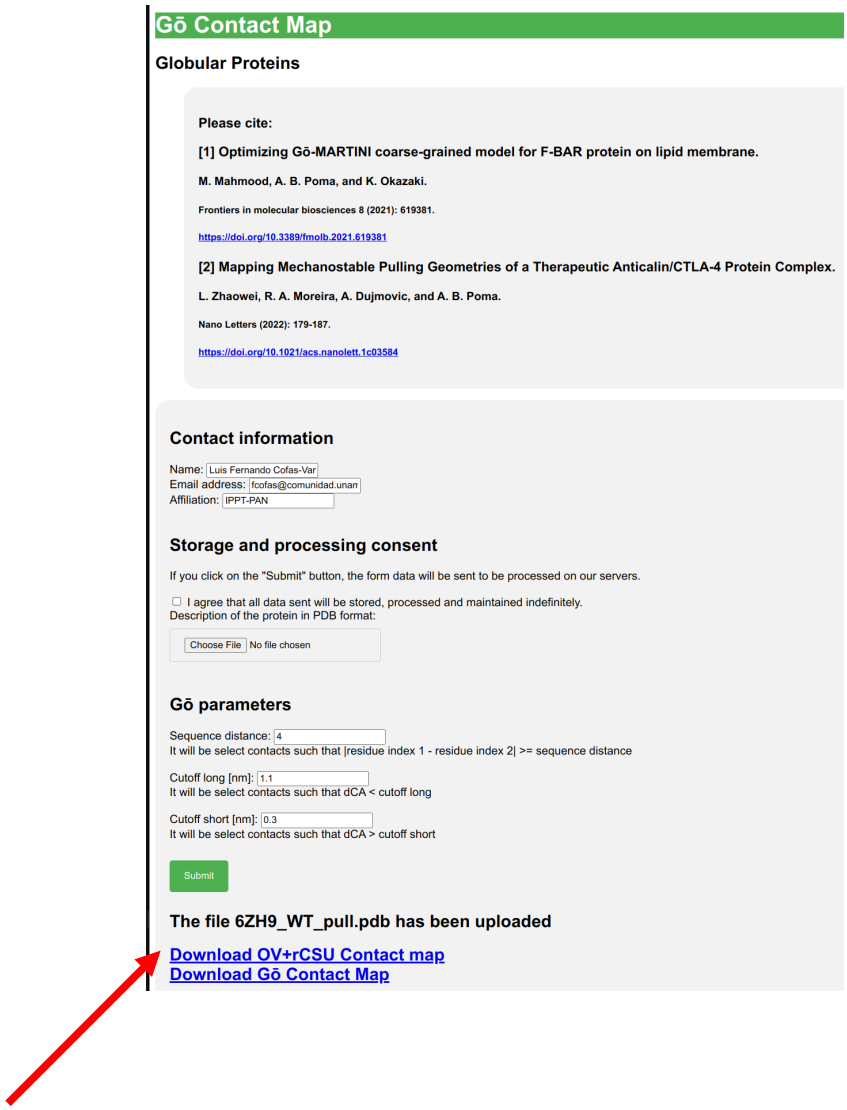
```
pdbfixer 6ZH9_aligned.pdb --output=6ZH9_H.pdb --ph=7.4 --keep-heterogens=none
```

Using the keep-heterogens=none option, PDBFixer will get rid of all the molecules that are not made up of standard amino acids. This is imperative for the creation of the coarsely grained representation.

From this new file we must create individual files for each chain in our system. To accomplish this, simply copy the file n times, and then remove all but the desired chain. Another approach involves employing Chimera, removing the unwanted links via the 'delete' command, and saving the resulting file. Each new file may be designated with a name corresponding to the chain it comprises, such as E.pdb, F.pdb, and so forth.

Calculate contact map based on OV+rCSU for a protein structure

In order to obtain the contact map, we will use the 6ZH9_H.pdb file on the contact map server: http://pomalab.ippt.pan.pl/GoContactMap/. The contact information must be filled out, the storage and processing consent must be accepted, and the PDB file must be uploaded. We must download the OV+rCSU Contact map file. The default name of this file is contact_map.out.



contact_map.out. : Download and to be used in next steps. Alternatively, a locally executable version of the webserver is also available (contactMaps_Analysis.tar.gz). The source files for

the contact_map executable can be found at https://doi.org/10.5281/zenodo.3817447 If you use the locally executable version of the contact map, please cite:

[1] Optimizing Gō-MARTINI coarse-grained model for F-BAR protein on lipid membrane.
M. Mahmood, A. B. Poma, and K. Okazaki.
Frontiers in molecular biosciences 8 (2021): 619381. https://doi.org/10.3389/fmolb.2021.619381
[2] Mapping Mechanostable Pulling Geometries of a Therapeutic Anticalin/CTLA-4 Protein Complex.
L. Zhaowei, R. A. Moreira, A. Dujmovic, and A. B. Poma.
Nano Letters (2022): 179-18. https://doi.org/10.1021/acs.nanolett.1c03584

It is necessary to unzip the contactmap.tar folder (it is inside contactMaps_Analysis.tar.gz). The executable can be used by running the following command.

```
>> '/route/to/folder/contact_map' 6ZH9_H.pdb > contact_map.out
```

Gromacs preparation for GoMartini SMD simulation

We will then convert each chain from the representation of all atoms into the Martini representation.

```
>> martinize2 -f E.pdb -o E_CG.top -x E_CG.pdb -dssp /route/to/dssp/dssp -p backbone -ff martini3001 -cys auto -ignh -from amber -maxwarn 1


>> mv molecule_0.itp E_CG.itp

>> martinize2 -f F.pdb -o F_CG.top -x F_CG.pdb -dssp /route/to/dssp/dssp -p backbone -ff martini3001 -cys auto -ignh -from amber -maxwarn 1

>> mv molecule_0.itp F_CG.itp
```

Note: In certain instances, Martinize2 may encounter difficulty determining the DSSP version, resulting in an error. This issue typically arises due to an incomplete installation of the library required by DSSP. If you encounter this error, please execute the following command:

conda install anaconda::libboost=1.73.0 -y

This command should resolve the issue and enable the correct determination of the required secondary structure files."

Now, we will use the GoMartini method with the create_gomartini.py script. Note that it is essential to maintain the order of the chains. (./create_gomartini.py can found at http://pomalab.ippt.pan.pl/web/gomartini/create_gomartini.py)

```
>> ./create_gomartini.py -s E_CG.pdb,F_CG.pdb -f contact_map.out --go_eps 12.00
```

The default setting for the interaction potential is 9.414 kJ/mol. This value can be changed to match experimental results or simulations for all atoms.

We will obtain the following files: go_system.gro, go_martini.itp, go_molecule1.itp and go_molecule2.itp.

We need to create a text file and rename it topol.top, which should contain the following.

```
#include "martini_v3.0.0.itp"
#include "go_martini.itp"
#include "martini_v3.0.0_ions_v1.itp"
#include "martini_v3.0.0_solvents_v1.itp"
#include "go_molecule1.itp"
#include "go_molecule2.itp"


[ system ]
SYSTEM NAME

[ molecules ]
go_molecule1   1
go_molecule2   1
```

The next step is to fabricate the box, for this system, we will use a box with dimensions of 10 x 10 x 60 nm, with the complex centered at coordinates 5, 5, and 5.5.

```
>> gmx editconf -f go_system.gro -center 5.000 5.000 5.500 -box 10.000 10.000 60.000 -o 6ZH9_WT_pull_CG.gro
```

The next step is to run a minimization in a vacuum.

```
>> gmx grompp -p topol.top -f min_vac.mdp -c 6ZH9_WT_pull_CG.gro  -o minimization-vac.tpr -normvsbds
```

```
>> gmx mdrun -deffnm minimization-vac -v -nt 2 -pin on -pinoffset 0
```

Now, let us proceed with adding water to the system.

>> **gmx solvate -cp minimization-vac.gro -cs water.gro -radius 0.21 -o solvated.gro**

```
Generating solvent configuration
Will generate new solvent configuration of 3x3x17 boxes
Solvent box contains 56629 atoms in 56629 residues
Removed 8767 solvent atoms due to solvent-solvent overlap
Removed 465 solvent atoms due to solute-solvent overlap
Sorting configuration
Found 1 molecule type:
      W (   1 atoms): 47397 residues
Generated solvent containing 47397 atoms in 47397 residues
Writing generated configuration to solvated.gro

Back Off! I just backed up solvated.gro to ./#solvated.gro.1#

Output configuration contains 48511 atoms in 48043 residues
Volume                 :        6000 (nm^3)
Density                :     2417.79 (g/l)
Number of solvent molecules:  47397
```

The topol.top file should be copied and renamed to system.top. We need to add a line to this new file to indicate that the system contains water and the number of beads it contains.

```
[ molecules ]
go_molecule1   1
go_molecule2   1
W           47397
```

The system has been neutralized by adding counterions, and the ionic strength can be adjusted.

>> **gmx grompp -f empty.mdp -c solvated.gro -p system.top -o ions.tpr -maxwarn 1 -normvsbds**

>> **echo W | gmx genion -s ions.tpr -o system_solv_ions.gro -p system.top -pname NA -nname CL -neutral -conc 0.15**

```
Command line:
  gmx genion -s ions.tpr -o system_solv_ions.gro -p system.top -pname NA -nname
CL -neutral -conc 0.15

Reading file ions.tpr, VERSION 2023.1 (single precision)
Reading file ions.tpr, VERSION 2023.1 (single precision)
Will try to add 1041 NA ions and 1048 CL ions.
Select a continuous group of solvent molecules
Group     0 (         System) has 93387 elements
Group     1 (        Protein) has  1114 elements
Group     2 (      Protein-H) has  1114 elements
Group     3 (        C-alpha) has   323 elements
Group     4 (       Backbone) has   323 elements
Group     5 (      MainChain) has   323 elements
Group     6 (   MainChain+Cb) has   323 elements
Group     7 (    MainChain+H) has   323 elements
Group     8 (      SideChain) has   791 elements
Group     9 (    SideChain-H) has   791 elements
Group    10 (    Prot-Masses) has  1114 elements
Group    11 (    non-Protein) has 92273 elements
Group    12 (          Other) has 92273 elements
Group    13 (              W) has 92273 elements
Select a group: 13
```

We choose group 13 to replace water beads with ion beads.

In order to proceed with the subsequent steps, it is imperative that we include the line **define = -DPOSRES** in all mdp files (except for pull.mdp) to maintain the system in its initial coordinates and prevent any potential drift of the complex.

To achieve minimization, we will utilize the following command:

```
>> gmx grompp -p system.top -c system_solv_ions.gro -f minimization.mdp -r system_solv_ions.gro -o minimization.tpr -normvsbds

>> gmx mdrun -deffnm minimization -v -nt 2 -pin on -pinoffset 0
```

An NVT equlibration is performed next, followed by an NPT equlibration.

```
>> gmx grompp -p system.top -c minimization.gro -r minimization.gro -f nvt.mdp -o nvt.tpr -maxwarn 2 -normvsbds

>> gmx mdrun -deffnm nvt -v -nt 2 -pin on -pinoffset 0

>> gmx grompp -p system.top -c nvt.gro  -f npt.mdp -r nvt.gro -o npt.tpr -maxwarn 2 -normvsbds
```

```
>> gmx mdrun -deffnm npt -v -nt 2 -pin on -pinoffset 0
```

We now create an index containing two groups. A freeze procotol will be implemented. One group will be used to pull from, and the other group will be kept stationary. The last three residues of each chain will be found in each group.

```
>> gmx make_ndx -f npt.gro -o index.ndx


a 980 981 982 983 984 985 986
name 15 Pull_A
a 471 472 473 474 475 476
name 16 Pull_B
q
```

Now we run the production.

```
>> gmx grompp -p system.top -c npt.gro -f pull.mdp -o pull1.tpr -
maxwarn 1 -normvsbds -n index.ndx -r npt.gro

>> gmx mdrun -deffnm pull1 -v -nt 2 -pin on -pinoffset 0 -pf
pullf1.xvg -px pullx1.xvg
```

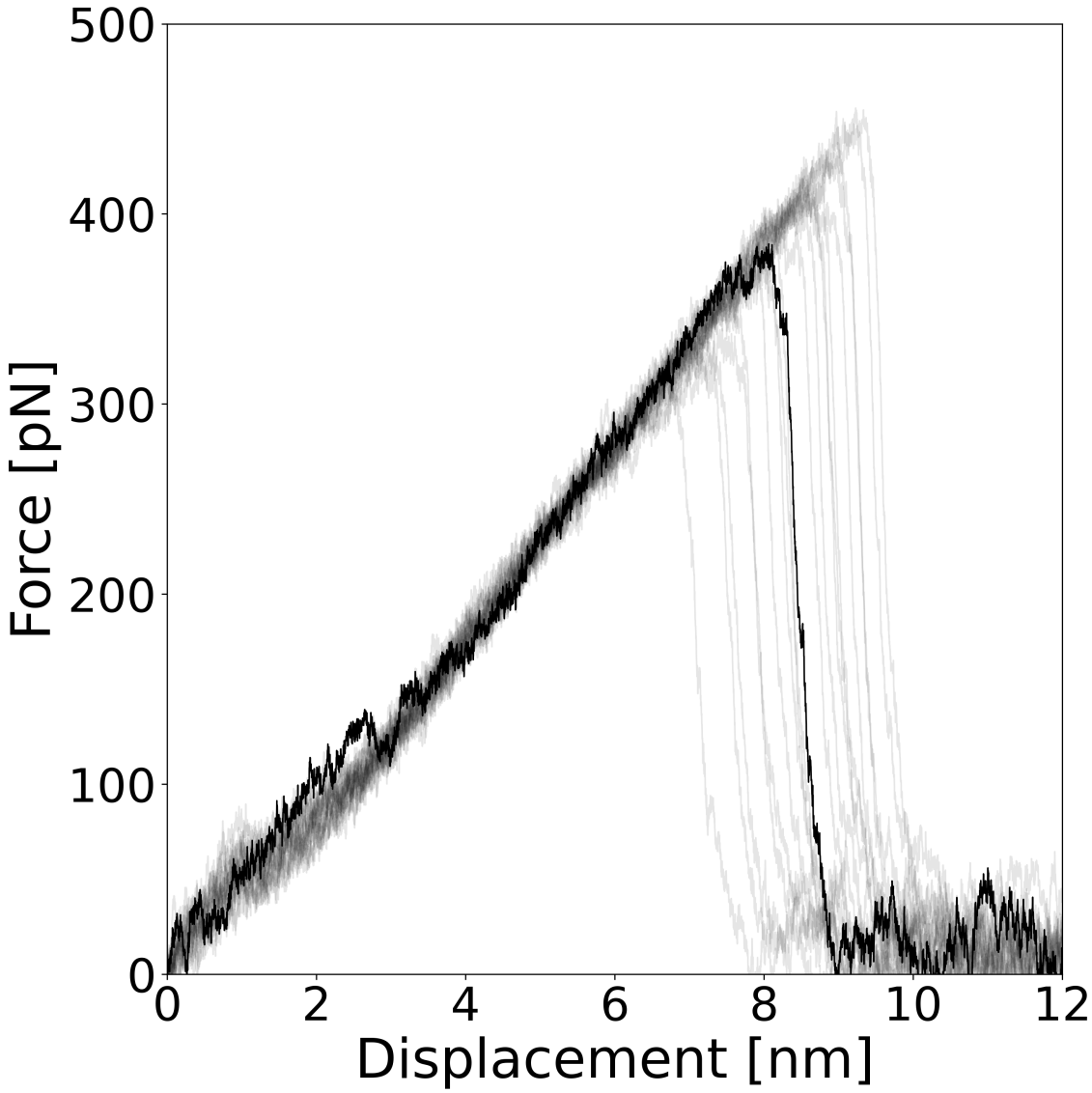
The commands below can be used to center the complex and remove the solvent.

```
>> echo '0' | gmx trjconv -f pull1.xtc -s pull1.tpr -pbc whole -o
pull_whole.xtc
>> echo '0' | gmx trjconv -s pull1.tpr -f pull_whole.xtc -o
pull_nojump.xtc -pbc nojump
>> echo '1 0' | gmx trjconv -s pull1.tpr -f pull_nojump.xtc -o
pull_center.xtc -center
>> echo '1 0' | gmx trjconv -s pull1.tpr -f pull_center.xtc -o
pull_compact.xtc -ur compact -center -pbc mol
>> echo '1 1 1' | gmx trjconv -s pull1.tpr -f pull_compact.xtc -o
pull_fit1.xtc -fit rot+trans -center
>> rm -r pull_whole.xtc
>> rm -r pull_nojump.xtc
>> rm -r pull_center.xtc
>> rm -r pull_compact.xtc
>> rm -f \#*
```

To generate the topology without solvent and ions.

```
>> gmx convert-tpr -s pull1.tpr -o pull1_dry.tpr
```

The force vs distance profile should look similar to the following:



Note: the performance of this system on an RTX A4000 graphics card is approximately 4391 nanoseconds per day.