Molecular Dynamics Exercise VII: Potential of Mean Force of an Amyloid Layer Separation via Umbrella Sampling

Tristan Alexander Mauck

December 23, 2024

This report deals with the calculation of the binding free energy between a amyloid β-protofibril and a an amyloid β-peptide. For this the Umbrella sampling method was used and the results of the simulations were analyzed using the WHAM algorithm. All simulations were performed using GROMACS.

1 Introduction and Procedure

To find the free energy of binding ΔG_{bind}, the potential of mean force needs to be found for the given molecules.

$$F_R(R') = F_R(\infty) + \int_{\infty}^{R'} dR'' \frac{dF_R(R'')}{dR''} \tag{1}$$

This is the free energy depending on the reaction coordinate $R(r^N)$ evaluated at a specific position $R'(r^N)$ [1]. While in principle the samples necessary for this calculation can be extracted from a normal simulation of the two molecules, this is rather inefficient as the necessary values of R' might only appear very rarely. Hence, to facilitate sampling at a given distance R_0, a biasing potential is introduced to keep the molecules close to the desired spacing.

$$U^{res}(R(\mathbf{r}^N), R_0) = \frac{1}{2}(R(\mathbf{r}^N) - R_0)^2 \tag{2}$$

In this simulation $F_R(R')$ can be found by recording the average restraint force as this corresponds to minus the average physical force.

$$F_R(R') = -\int_{R'_0}^{R'} \left\langle \frac{\partial U^{res}(R'', R_0)}{\partial R''} \right\rangle dR'' + \text{const} \tag{3}$$

This in then done for multiple values of the reaction coordinate (umbrellas). To construct the final curve out of the obtained pieces $F_{R_i}(R'_i)$ using the (weighted histogram analysis method) WHAM it is important that the distances sampled in adjacent simulations overlap each other sufficiently.

For the simulation, the amyloid β-protofibril was placed in a box with dimensions $6.560\,\text{nm} \times 4.362\,\text{nm} \times 12\,\text{nm}$, which was then filled with SPC water molecules. Additionally, electrostatic neutrality was achieved by adding NA- and CL-ions into the solvent. Following a steepest descent energy minimization, a NPT equilibration was conducted. This simulation conducted

50 000 steps of $\Delta t = 0.002$ ps and used the Berendsen weak coupling method for both temperature and pressure correction.

To find the initial configurations of the systems for the umbrella sampling, a pull simulation was conducted in GROMACS. In this simulation, running for 500 ps, the amyloid β-peptide was pulled away from the rest of the protofibril. (This was acheived by fixing a chain of the protofibril.) Using the bash script shown below, the trajectory from this simulation was used to take systems configurations with the center of mass distance between both fragments ranging from 0.5 nm to 5.0 nm in steps of 0.2 nm.

```
#!/bin/bash

gmx grompp -f md_pull.mdp -c npt.gro -p topol.top -r npt.gro -n index.ndx -t
    npt.cpt -o pull.tpr -maxwarn 4
gmx mdrun -v -ntomp 4 -deffnm pull -pf pullf.xvg -px pullx.xvg

gmx pairdist -f pull.xtc -s pull.tpr -n index.ndx -sel "com of group Chain_A
   " -ref "com of group Chain_B" -o dist.xvg

sed '/^\@/d' dist.xvg |\
sed '/^\#/d' |\
awk '$2 > i+0.2{a[k++]=$1;i=i+0.2}END{for(j=0;j<501;j++)if(a[j] != 0.000 &&
   j > 0 || j < 1 )printf("%7.3f%1s",a[j],"\n")}' > important_frames.txt

for frame in $(cat important_frames.txt)
do
        printf "0\n" | gmx trjconv -s pull.tpr -f pull.xtc -o conf${frame
   %%.*}.gro -b $frame -e $frame
done
```

Requested Comments:

1. Preparation of the pulling simulation taking the results of the NPT equilibration performed before. As for the correct execution of the pulling the chains need to be labeled correctly, the index.ndx file is included defining 'Chain A' and 'Chain B'.

2. The pulling simulation is run. 'ntomp' defines the number of OpenMP threads to start with. Additionally to the results from the run, the xvg-files 'pullx' and 'pullf' are returned. They show the development of the force as well as the distance over the simulated 500 ns.

3. The GROMACS command 'pairdist' returns the distance between the reference ('ref') and selection ('sel'). Using the results from the previous runs, this is calculated and stored for the centres of mass of the chains A and chain B

4.a.5. The sed command is used to delete the patterns inside the /'s from 'dist.xvg'.

6. 'awk' allows for the using of logical expressions to edit text. In this case the output file 'importantframes.txt' is created. The following program is implemented: When the distance between the centers of mass of chains A and B is larger than i+0.2, then a[k+1] is defined to be the simulation time at that moment and i is increased be 0.2. Then a for loop in j runs from 0 to 500, going through the indices of a. If a and j are not 0, a[j] is printed in the output file with 7 digits and 3 digits after the comma.

7. This for loop goes over all the important frames (0.2 nm apart) defined in the previous steps.

8. For each of the frames running through the for loop, the respective .gro configuration is extracted using 'gmx trjconv'.

Having obtained the conformations of the individual umbrella windows, a NPT equilibration and a production run were conducted for each of them. The equilibration run conducted 25 000 steps of time step $\Delta t = 0.002\,\mathrm{ps}$ and used again the Berendsen method for both temperature and pressure control. For the production run 250 000 steps of the same time step were calculated. Temperature and pressure control were done using the Nose-Hoover thermostat and the Parrinello-Rahman barostat.

The final analysis was done using GROMACS implementation of the WHAM method.

2 Simulation Results and Discussion

The results from the pull simulation are shown below. As expected the molecule initially resists to the force and the displacement remains almost negligible for the first 100 ps. After applying $\approx 1300\,\mathrm{kJ\,nm^{-1}\,mol^{-1}}$, the resistance of the molecule ceases and after a fast separation of the molecules, the velocity of disintegration becomes almost constant. Additionally, a graphic of the separation process is provided (figure 2). Using the WHAM method the mean force was

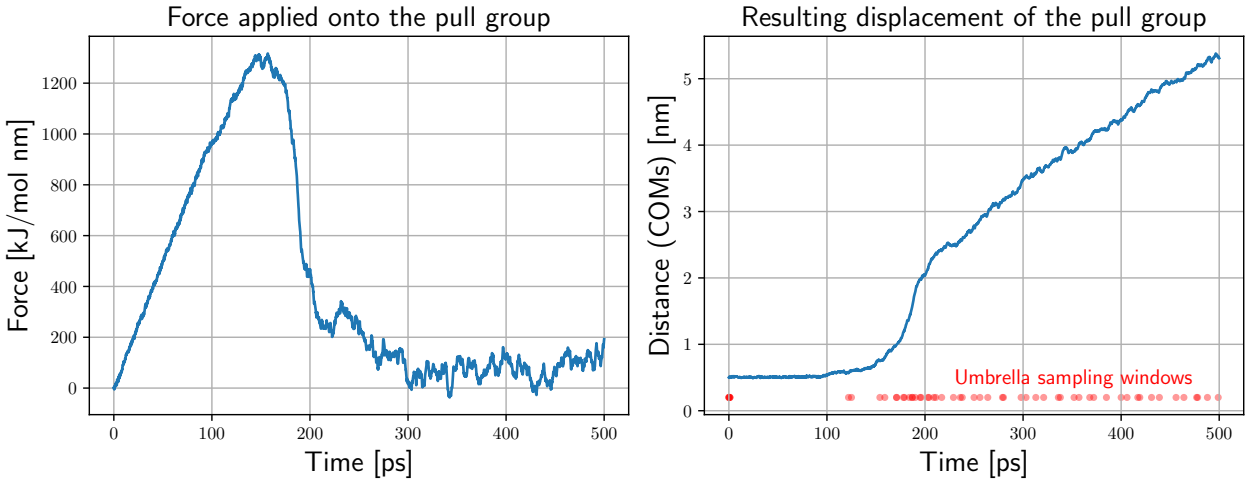

Figure 1: The force on the pull group and resulting displacement over the 500 ps of the initial simulation. The small red dots denote the conformations used as umbrella windows.

constructed from the individual umbrella simulations. The results are shown below. Clearly, the result for ΔG_{bind} differs from the literature value. The obtained value is $-38.6\,\frac{\mathrm{kcal}}{\mathrm{mol}}$, which differs from the literature value by roughly 23% [2]. There are likely several reasons for this. Firstly, there are three visible instances of the umbrella simulations being not close enough (see histograms in figure 3 (1.2 nm, 1.6 nm and 3.4 nm)). For the Wham algorithm to work well, the sampled spaces from adjacent umbrella simulations need to have sufficient overlap. The consequence of the missing overlap can be seen in the PMF curve. With WHAM being unable to join the curves together, the free energy profile shows unphysical drops. Furthermore, the short simulation time results in fairly low count rates. Even at the maxima of most histograms only about 1000 samples are taken.

While the free energy profile is likely not yet a correct representation of reality, some of its features are likely to be correct. Comparing it to the provided example curve, one finds that

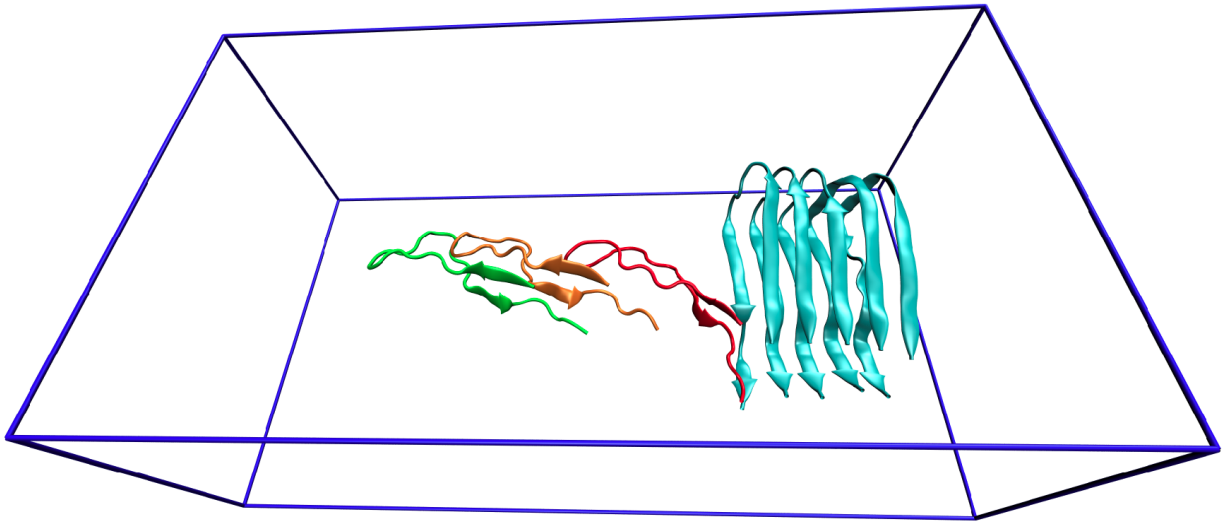

Figure 2: Visualization of the pulling simulation using VMD. From the initial 'blue' configuration, the *beta*-peptide is seen to slowly separate from the fibril over time. The configurations shown are snapshots at 0 ps, 190 ps, 336 ps and 464 ps.

the PMF behaves well in the range from 0.5 nm to 3.2 nm. The slope observed in this part should be continued up to roughly 4 nm and reach then a plateau (bond fully broken). In this simulation this plateau was however already reached already at about 3.4 nm. Potentially, this is due to some unfortunate positioning of the ligand with respect to the complex. This could be fixed by controlling the parameters of the pulling simulation more tighly.

3 Bibliography

[1] Zacharias, Martin et Al. Molecular Dynamics: From Basics to Applications Lecture Notes; 2023
[2] Vollmers, Luis; Zacharias, Martin; Reif, Maria; Molecular Dynamics: Exercise 7 - Potential of a Mean Force on an Amyloid Layer Separation via Umbrella sampling; 2023

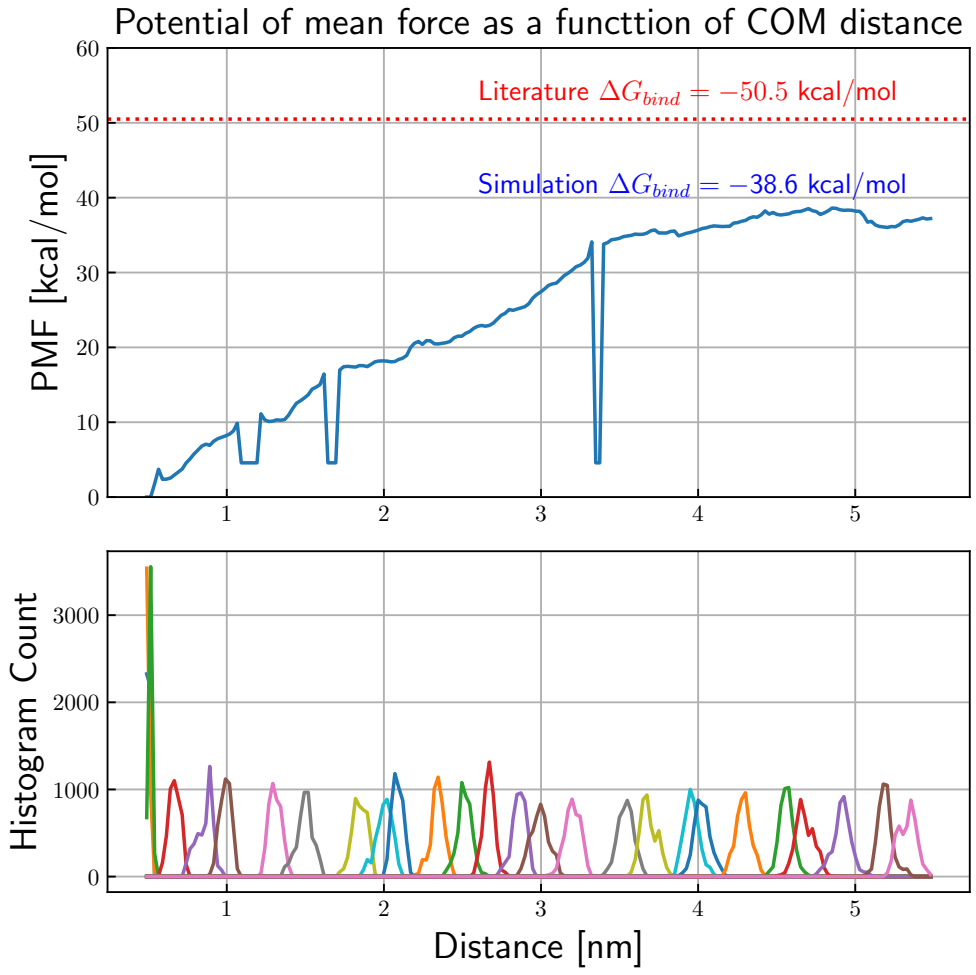

Figure 3: Free Energy Profile (PMF) along the separation of the COM of the two molecules.