FDA Tutorial

Contents

FDA measures forces of interactions as calculated by Molecular Dynamics (MD) Simulations. It reports changes in interatomic forces (F_{ij}) between pair of atoms (i and j) from MD trajectories. Most commonly, changes in interatomic forces upon some external pertubation (e.g. ligand binding) or pulling forces are of interest. Forces sampled by MD simulations under two conditions (with and without pertubation) are compared.
In this tutorial we are going to compare a two-protein system either bound or unbound with a ligand.

You need:

- a protein system
  This example uses a two-protein system.
- two conditions to compare
  e.g. a ligand bound and unbound, a protein under a constant force and the absence of force.
- trajectories of equilibrium MD simulations in both conditions.
  A recommendation would be, per condition, 5-20 replicas of several hundred ns for a typical globular protein system. However, the complete simulation time highly depends on the system and on the question you are asking.

Be sure to always keep the two conditions separate and repeat everything with the second condition until we are explicitly comparing the output (see Calculating the force differences).

1 FDA Versions

First, you need MD trajectories from GROMACS. Make sure that the GMX version used to generate the trajectories matches the FDA Version. To install your own FDA Version, find the instructions in the fda-manual.
Please keep in mind, that newer versions of FDA will follow in the future and might change parts of this tutorial.

2 Trajectory processing and concatenation

For the preparation of the trajectories we need: one trajectory (.xtc), one atomic coordinate file (.gro) and the compressed gromacs input file(.tpr).
Keep in mind, if your trajectory only contains the protein and not the water because of the .mdp options you set, your coordinate and run-input files most probably will have the water present. So first we need to convert both files:

```
gmx trjconv -f md.gro -s md.tpr -o md_woWater.gro
gmx convert-tpr -s md.tpr -o md_woWater.tpr
```

Both times select 1 (for Protein) in the selection prompt.

During the MD we approximate a large continuum by copying our small simulation box in every direction. So it can happen that the protein crosses from one box into its periodic neighbor box. This is however difficult for FDA to handle so we convert this back and also center the protein in the box:

```
gmx trjconv -f md.xtc md_woWater.tpr -o md_proc.xtc -center -pbc mol -ur compact
```

Here you also select 1 both for centering and for the output.

FDA requires a specification for which groups of atoms the forces should be calculated. So the last step here is to generate an index-file:

```
gmx make_index -f md_woWater.gro -o index.ndx
```

Since we do not want to insert any new index-groups you can just type "q" to close the index-maker without changes.

Note that the conversions to the .gro file and the .tpr file and the generation of the .ndx file only have to be done for one replicate run since these do not change over the replicas. However, we need to concatenate each replica trajectory of one condition as this is easier for FDA to handle. If

you have large trajectories it might be sometimes wise to run the FDA on each replicate trajectory individually and average the results later.

```
gmx trjcat -f <file_name_1>.xtc <file_name_2>.xtc ... -o md_concat.xtc -cat
```

We have now completed only the preparations for the system without a ligand, in a real project you would also need to prepare the system with the ligand. Keep in mind that we can ignore the ligand for these conversions (again you would need to select only the protein). This is because the effect of the ligand is already represented in the trajectory itself and FDA can only compare residues that are present in both conditions.

3 FDA input file

The FDA input file is given by the input file with the file ending .pfi. It contains all the parameters that FDA needs to run. Below you find an example input file.

`input.pfi`

```
onepair = summed
group1 = Protein
group2 = Protein
atombased = no
residuebased = pairwise_forces_scalar
type = all
time_averages_period = 0
residuesrenumber = no
```

Next is a brief description of the options used here, but more information and additional keywords can be found in the FDA manual.

`onepair = summed` : indicates that all the pairwise forces on one atom will be summed up.

The **`groups`** definition is taken from the index file (.ndx). Both groups are set to Protein, because we want to compare all the intramolecular interactions. If you had a two-protein system and you were only interested in the forces between the two proteins, you would make one index group for each protein and put protein1 for group1, and protein2 for group2.

You can print out either the **`atom- or residuebased`** forces, or both at the same time. Here, only residuebased is printed. The format of the atom- or residuebased forces is in this case `pairwise_forces_scalar`: this means that the pairwise forces are saved as a single, singed scalar. In other words, you get a list of each interacting residue-pair and its corresponding force value. Other options can be found in the manual.

`Type` indicates the interaction types to include. Options are: `bond, angle, dihedral,`

`polar, coulomb, lj, nb14, bonded, nonbonded, all`. In this example we are including all interactions and do not filter for specific types. Be aware that with this option the forces of all interaction types are summed up. If you have more than one interaction type per residue-pair, you cannot distinguish which interaction type has the strongest influence.

`time_averages_period`: if this is set to 0 the average over all frames of the (concatenated) trajectory is calculated. You will only get one value for the complete simulation time. If you set the average to 1 it will print out the forces for every frame and if you set it to *n* it will average over *n* frames. This way you could get a time-resolved FDA i.e. dynamic force information.

`residuerenumbering`: This option is only needed when you have overlapping indices, like for example in a two-protein system it would be likely that the two proteins have some residue indices in common. This is a problem for FDA so with the **`residuerenumber`** option it will renumber the residues. The residue renumbering works as follows: it will start with the first residue of the first protein in the .gro file at sets it as residue number 0. Then it just counts on to the last residue of the second protein. The default behavior of **`residuerenumber`** is to perform the renumbering only if a conflict is detected.

4 Running FDA

Now it is time to run the FDA. You have to provide the concatenated trajectory file (md_concat.xtc), the .tpr file of one of the replica-runs and an index-file .ndx including the groups selected in the input .pfi file.

```
gmx_fda mdrun -nt 1 -rerun md_concat.xtc -s md_woWater.tpr -pfi input.pfi -pfn index.ndx
-pfr FDA1.pfr
```

FDA formally uses the mdrun engine, but it only performs a rerun (`-rerun` flag) so it is considerably faster. FDA can only run on one core. The number of threads has to be set to 1 (`-nt 1`).

It is important to use the same input file (.pfi) for both of the conditions otherwise the output would be useless.

5 Understanding the output file

Attention!: The FDA output file comes in the format of a backup file: #FDA1.pfr.1#. So we have to convert them back into normal files:

```
mv ./#FDA1.pfr.1# FDA1.pfr
```

The FDA output file (`FDA.pfr`) contains four columns:

1. residue index of the first interacting partner
2. residue index of the second interacting partner
   both residue indices are renumbered when the residuerenumber option is turned on
3. the force value (for this example given as `pairwise_forces_scalar`)
4. the type of interaction, one of the following numbers:

None	0
Bond	1
Angle	2
Dihedral	4
Bonded (bond + angle + dihedral)	7
Polar	8
Coulomb	16
LJ	32
Coulomb + LJ	48
NB14	64
Nonbonded (coulomb + LJ)	112
Bondend+Nonbonded	119
All (bonded, nonbonded, polar)	127

6 Calculating the force differences

Once both FDA runs have finished (one for with the ligand, and the other for without ligand), we compare and subtract the forces. There is a built-in function to subtract the forces in the visualization tool (see Visualization) BUT it only blindly subtracts the lines. This option can only safely be used if every residue is involved in some kind of interaction in both conditions!

The more adjustable option is to write an own script to subtract exactly the force values you are interested in. Such a script for example reads in the .pfr files from both conditions and compares the residue indices of both interacting residues between the two files. It only subtracts the force values if the interacting residues from both files are the same. If in one of the files one residue combination is missing it will just set it to 0 and then subtract. The output file is in the style of a .pfr FDA output file.

7 Visualization

Now we are almost done. In the last step we visualize the force differences with the `fda_graph` tool.

```
gmx_fda fda_graph -ipf difference_pfr.pfr -s md_woWater.gro -n index.ndx
-o difference_graph_4_100.pdb -convert -frame 0 -min 4 -t 100 -big
```

for residuebased analysis: the best index group to select is the backbone C-alpha atoms (normally this should be 3).

The other options used in the command are explained below.

- `-frame 0` option is crucial for the mapping to be correct (single frame number or "average n" to take the mean over every n-th frame)
- `-convert` convert the forces from gromacs units (kJ/mol/nm) to pN
- `-t` force threshold, only forces larger than the cutoff are shown.
- `-min` minimal number of nodes in a force network to be visualized. This reduces noise by selecting larger connected networks only.
- `-big` only the largest connected force network (number of nodes) is shown. (Please note that this option seems to be non-functional in the FDA 2016 and 2018 versions, but it should be working in the 2020 version.)

The options `-t` and `-min` are highly sensitive to the system you have so try a few out.

Load first the atom coordinate file into VMD.

```
vmd md_woWater.gro
```

Go to **Graphics** and select **Representations**. In the pop-up window go under **Drawing Method** and select **NewCartoon**. Now go to **File** and select **New Molecule**. In the pop-up window browse for the **difference_graph** file and click on **load**.

Go back to the **Graphical Representations** pop-up window. Now you see under **Selected Molecule** that the difference graph is selected. Go again to **Drawing Method** and select a **Licorice** representation for the force network. You can adjust the color to your liking under the **Coloring Method** option.

You can save the picture by going to **File** and selecting **Render**. In the pop-up select **Tachyon (internal, in-memory rendering)** option and change the Filename to something more conclusive but leave the file extension .tga. Now just go to **Start Rendering** and you are done.

Display the other force networks you generated with varying thresholds in the same way. Observe the changes in the force network depending on the respective cutoff.