# Exercise on using the NMRlipids databank

NMRlipids summer school 2022

June 2<sup>nd</sup> 2022 Espoo, Finland

#### Goal

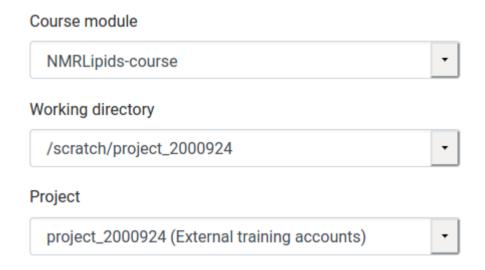
- Demonstrate the basic ideas on how to use the NMRlipids databank to automatically analyse large sets of MD simulation data
- Spark interest on using NMRlipids databank on your research

#### Learning outcome

- Analysing correlations between membrane properties from the NMRlipids databank
- Analysing changes of a property (area per lipid as an example) as function of lipid composition from the NMRlipids databank

## Setting up the exercise

- OPTION 1: CSC interface (recommended for people with no Python exprience)
  - Login to Puhti web interface with the given account: https://www.puhti.csc.fi/public/login.html
  - Go "Jupyter for courses" and select:



- Click "Connect to Jupyter"
- You can run the code in each cell with "shift"+enter

### Setting up the exercise

- **OPTION 2:** Running on your own computer (recommended for people with python environments already set up)
  - Clone these two repositories in the same folder: https://github.com/NMRLipids/DatabankExercises https://github.com/NMRLipids/Databank
  - Install python packages that are imported in the beginning of the exercise notebook

#### Exercise

https://github.com/NMRLipids/DatabankExercises/blob/master/APL/AreaPerLipidAndThicknessExamples.ipynb

- 1) Analyzing correlations between membrane properties. Analyze correlations between area per lipid, membrane thickness, and experimental parameters from the simulations in the NMRlipids databank using cells 2-8. What can you learn from these? You can also try to analyze other correlations.
- 2) Analyzing Changes in area per lipids as a function of membrane composition. Cell 9 plots the area per lipid as function ADDITIVE added to matrix. Investigate are area per lipid changes, for example, by adding POPS, POPE, POPG, or cholesterol to POPC membrane. You can also try other mixtures using molecule names used in the NMRlipids project. In the cases where force fields give different predictions, try to use quality evaluation results to decide which is the most reliable.