



Tutorial: Protein-Ligand Complex MD Setup with Jupyter Notebooks and BioBB

Exploring Biomolecular Modeling & Simulations 9-10/04/2025

EuroCC4 & BioExcel training workshop

Adam Hospital

adam.hospital@irbbarcelona.org

Agenda



THURSDAY, 10 APRIL

10:00 → 11:00	Interoperable and reproducible biomolecular simulation workflows using BioExcel Building Blocks (BioBB)	⌚ 1h
Speaker: Dr Adam Hospital (Institute for Research in Biomedicine in Barcelona (IRB-Barcelona) and Spanish National Institute of Bioinformatics (INB, ELIXIR-ES))		
11:00 → 11:30	Coffee Break	⌚ 30m
11:30 → 13:00	Hands on - Interoperable and reproducible biomolecular simulation workflows using BioExcel Building Blocks (BioBB)	⌚ 1h 30m
Speaker: Dr Adam Hospital (Institute for Research in Biomedicine in Barcelona (IRB-Barcelona) and Spanish National Institute of Bioinformatics (INB, ELIXIR-ES))		
13:00 → 14:00	Lunch Break	⌚ 1h
14:00 → 15:00	Fundamentals and practical use cases of free energy calculations with PMX	⌚ 1h
Speaker: Dr Sudarshan Behera (Max Planck Institute for Multidisciplinary Sciences, Goettingen)		
15:00 → 15:30	Coffee Break	⌚ 30m
15:30 → 17:00	Hands on - Fundamentals and practical use cases of free energy calculations with PMX	⌚ 1h 30m
Speaker: Dr Sudarshan Behera (Max Planck Institute for Multidisciplinary Sciences, Goettingen)		

A red box highlights the first two sessions: "Interoperable and reproducible biomolecular simulation workflows using BioExcel Building Blocks (BioBB)" and "Hands on - Interoperable and reproducible biomolecular simulation workflows using BioExcel Building Blocks (BioBB)". A red arrow points from the bottom of this box down to the "Session 1 (lecture)" box. Another red arrow points from the right side of the box down to the "Session 2 (hands-on)" box.

Session 1 (lecture):

Interoperable and reproducible biomolecular simulation workflows using BioBB

- Workflows & Biomolecular workflows
- BioBB library
- BioBB workflows
 - Demonstration workflows (JN)
 - Pre-exascale workflows (HPC)

Session 2 (hands-on):

Hands-on session on BioBB workflows:

- BioBB demonstration workflow tutorial: GROMACS Protein-ligand complex MD setup
- Quick view on BioBB workflow collection
 - Jupyter Notebooks
 - Pure Python (High Throughput)



BioBB Demonstration Workflows

GROMACS PROTEIN-LIGAND COMPLEX MD SETUP

2024.1



WorkflowHub

Launch

Download

View tutorial

Open Github repository

Open documentation

gmx
ligand
md
protein

(*) MyBinder provides a **free**, online version of **Jupyter Lab**. Take into account that the provided **resources** are **finite** and, in some occasions, it can take a long time to load or to execute your notebooks. **Please be patient** and don't try to execute several notebooks at the same time.

<https://mmb.irbbarcelona.org/biobb/workflows>

- ***MD setup (Protein / DNA) (AMBER / GROMACS)***
- ***Ligand parameterization***
- ***Protein-Ligand Docking***
- ***Free energy calculations***
- ***DNA helical parameters***
- ***Conformational Ensemble generation***

GROMACS Protein-Ligand MD Setup

GROMACS PROTEIN-LIGAND COMPLEX MD SETUP

2024.1



This tutorial aims to illustrate the process of setting up a simulation system containing a protein in complex with a ligand, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the T4 lysozyme L99A/M102Q protein (PDB code 3HTB), in complex with the 2-propylphenol small molecule (3-letter Code JZ4).

WorkflowHub

Launch

Download

View tutorial

Open Github repository

Open documentation

gmx
ligand
md
protein

(*) MyBinder provides a **free**, online version of **Jupyter Lab**. Take into account that the provided **resources** are **finite** and, in some occasions, it can take a long time to load or to execute your notebooks. **Please be patient** and don't try to execute several notebooks at the same time.

- **Complex: T4 Lysozyme (3HTB) with 2-propylphenol (JZ4)**
- **Ligand parameterization (ACPype)**
- **GROMACS MD Setup (Min + NVT eq + NPT eq + unrestrained short MD)**

Questions (1)

YES =

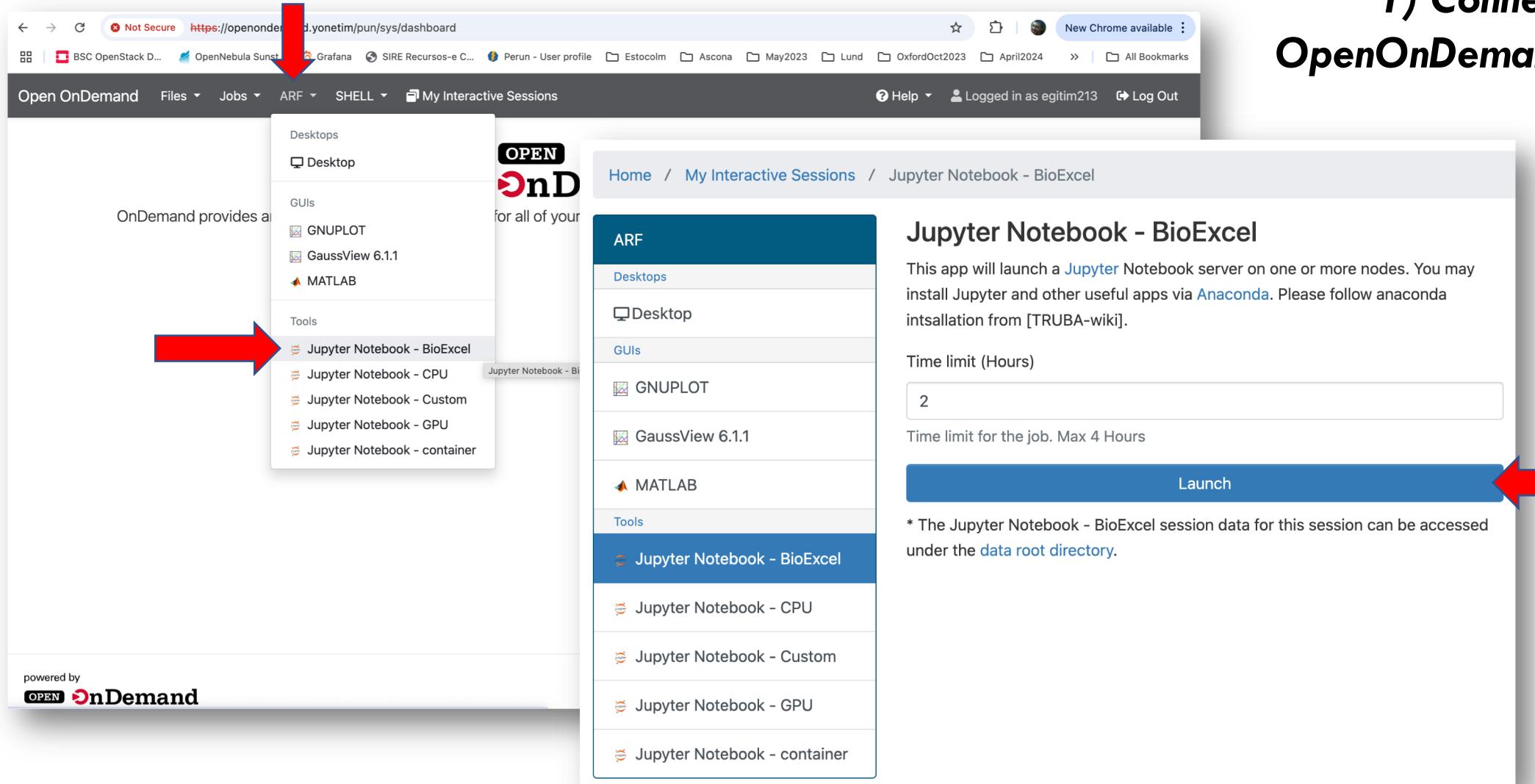
Raise Your Hand



- Have you ever used Jupyter Notebooks before? (Y/N)**
- Have you ever used OpenOnDemand platform before? (Y/N)**

TRUBA OpenOnDemand

<https://openondemand.yonetim/pun/sys/dashboard>



The screenshot shows the OpenOnDemand platform interface. At the top, the URL <https://openondemand.yonetim/pun/sys/dashboard> is visible in the browser bar. The ARF menu is open, with the 'Desktops' option selected. A red arrow points to the 'Desktops' link in the ARF menu. Another red arrow points to the 'Jupyter Notebook - BioExcel' link in the 'Tools' section of the ARF menu. The main content area shows the 'Jupyter Notebook - BioExcel' session configuration page. The session is titled 'ARF' and is set to use a 'Desktop' environment. The 'Tools' section lists 'Jupyter Notebook - BioExcel' (selected), 'Jupyter Notebook - CPU', 'Jupyter Notebook - Custom', 'Jupyter Notebook - GPU', and 'Jupyter Notebook - container'. The configuration form includes a 'Time limit (Hours)' input field set to '2' and a note: 'Time limit for the job. Max 4 Hours'. A large red arrow points to the 'Launch' button at the bottom of the form. The footer of the page includes the text 'OnDemand provides a' and 'powered by OPEN OnDemand'.

1) Connect to OpenOnDemand platform

Jupyter Notebook - BioExcel (1945646)

Queued

Created at: 2025-04-07 17:07:46 +03

Time Requested: 2 hours

Session ID: 39bb61bf-103c-4fe0-ada9-460c610fe22c

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.

✖ Delete

2) Jupyter Notebook – BioExcel

+

Connect to Jupyter

Jupyter Notebook - BioExcel (1945511)

1 node | 4 cores | Running

Host: >_ barbun10.yonetim

Created at: 2025-04-07 16:43:52 +03

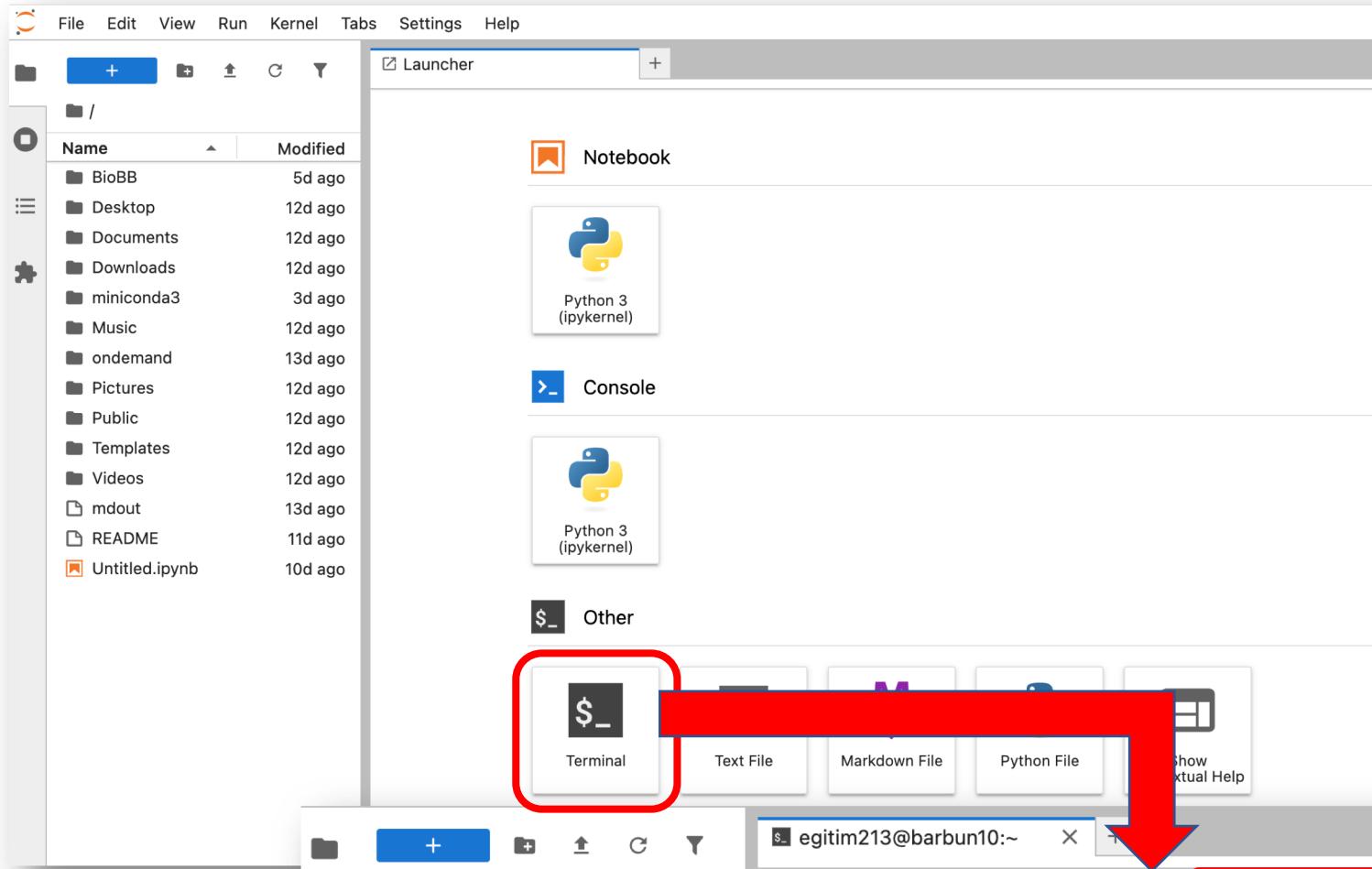
Time Remaining: 1 hour and 59 minutes

Session ID: 56ffddd7-d492-4806-b8a5-7d7b55b1b8f1

👁 Connect to Jupyter

✖ Delete

```
git clone https://github.com/bioexcel/biobb_wf_protein-complex_md_setup.git
```

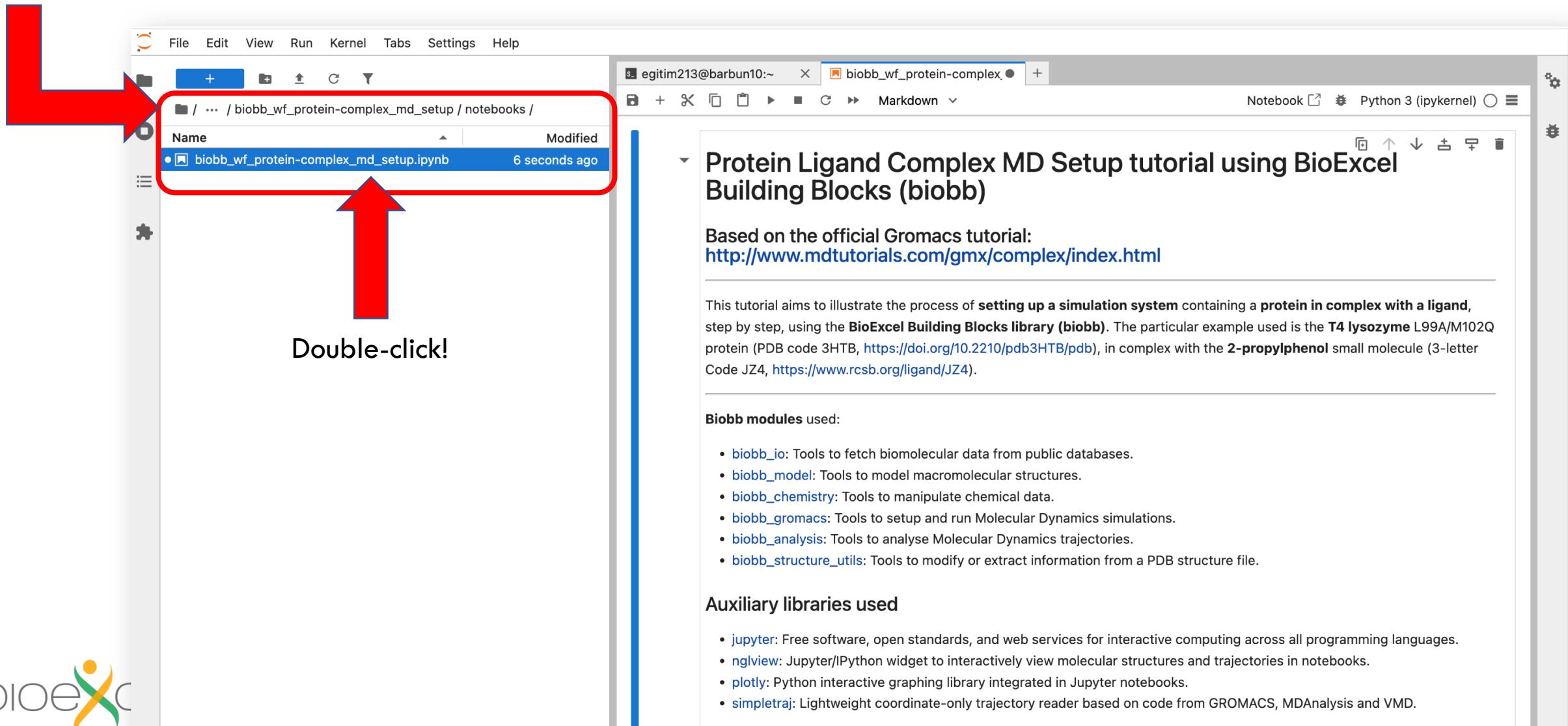


3) Clone (or copy) BioBB Workflow

4) Deploy BioBB Workflow Jupyter Notebook

With the **file explorer**, go the path:

Path: **biobb_wf_protein-complex_md_setup/biobb_wf_protein-complex_md_setup/notebooks/**



Part 1: Install and launch the workflow

Using TRUBA resources:

- 1) Connect to **OpenOnDemand** platform
- 2) Deploy ARF - **Jupyter Notebook – BioExcel** tool
- 3) Clone/copy **BioBB Workflow**
- 4) Open **Jupyter Notebook**



At home:

Conda Installation and Launch

```
git clone https://github.com/bioexcel/biobb_wf_protein-complex_md_setup.git
cd biobb_wf_protein-complex_md_setup
conda env create -f conda_env/environment.yml
conda activate biobb_wf_protein-complex_md_setup
jupyter-notebook biobb_wf_protein-complex_md_setup/notebooks/biobb_wf_protein-complex_md_setup.ipynb
```



Questions (2)

YES =

Raise Your Hand



- Have you ever setup/run a MD simulation using GROMACS before? (Y/N)**
- Have you ever setup/run a MD simulation using other MD packages (AMBER, NAMD, DESMOND...)? (Y/N)**
- Have you ever setup/run a MD simulation of a protein-ligand complex? (Y/N)**

Index

- **Part 1: Install and launch the workflow**
 - Conda Environment
 - Jupyter Notebook
- **Part 2: Download PDB and fixing structure**
- **Part 3: Generating topologies**
 - Protein
 - Ligand
- **Part 4: Generating protein-ligand complex structure**
 - Ligand Position Restraints
 - MD ready structure
- **Part 5: System Topology**
- **Part 6: MD Setup & Run**
- **Part 7: Trajectory (basic) analyses**

GROMACS PROTEIN-LIGAND COMPLEX MD SETUP

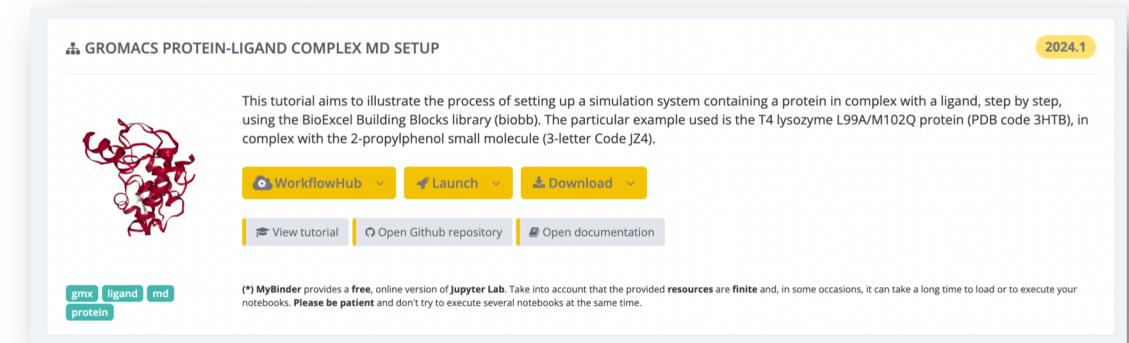
2024.1

This tutorial aims to illustrate the process of setting up a simulation system containing a protein in complex with a ligand, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the T4 lysozyme L99A/M102Q protein (PDB code 3HTB), in complex with the 2-propylphenol small molecule (3-letter Code JZ4).

WorkflowHub ▾ Launch ▾ Download ▾

View tutorial Open Github repository Open documentation

(* MyBinder provides a free, online version of Jupyter Lab. Take into account that the provided resources are finite and, in some occasions, it can take a long time to load or to execute your notebooks. Please be patient and don't try to execute several notebooks at the same time.



Questions (3)

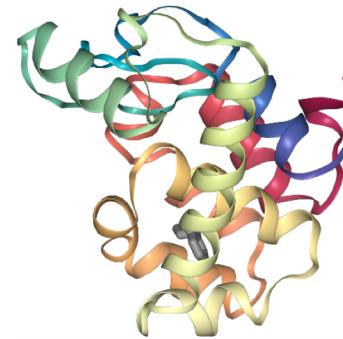
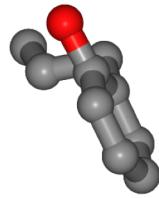
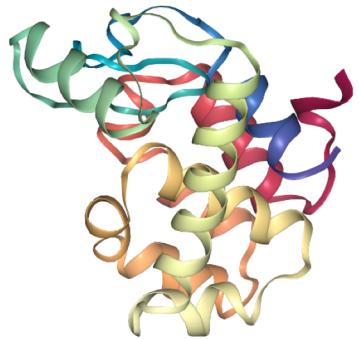
YES =

Raise Your Hand



- Are you familiar with PDB files and their content? (Y/N)**
- Do you use graphical interfaces in your everyday work? (Y/N)**

Part 2: Download PDB, check the structure, and generate protein topology



1. Input Parameters
2. Fetching PDB Structure
3. Fix Protein Structure
4. Create Protein System Topology

CHECK STRUCTURE

Reset

System Configuration

- Detect/Select Models: Single one
- Detect/Select Chains: Single one (A)
- Detect/Select Alt Locations
- Detect/Remove Heavy Metals
- Detect/Remove Ligands
- This structure doesn't contain DNA / RNA

Fix Structure Errors

- Detect/Fix Amide Assignment
ASN A140.OD1 Orig Fixed
- Detect/Fix Improper chirality

<https://mmb.irbbarcelona.org/biobb-wfs/structure/step1>

Questions (4)

YES =

Raise Your Hand



- Have you been following so far? (Y/N)**
- Have you ever parameterized a small molecule? (Y/N)**

```
[2]: # Ligand: Download ligand structure from MMB PDB mirror REST API (https://mmbrbarcelona.org/api/)
# Import module
from biobb_io.api.ligand import ligand

# Create prop dict and inputs/outputs
input_structure = ligandCode + '.pdb'

prop = {
    'ligand_code' : ligandCode
}

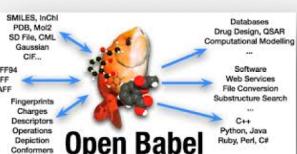
#Create and launch bb
ligand(output_pdb_path=input_structure,
       properties=prop)
```



```
[4]: # Babel_add_hydrogens: add Hydrogen atoms to a small molecule
# Import module
from biobb_chemistry.babelm.babel_add_hydrogens import babel_add_hydrogens

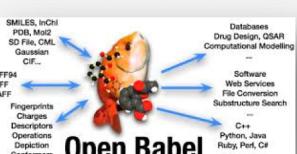
# Create prop dict and inputs/outputs
output_babel_h = ligandCode + '.H.mol2'

prop = {
    'ph' : pH,
    'input_format' : 'pdb',
    'output_format' : 'mol2'
}
```



```
[5]: # Create and launch bb
babel_add_hydrogens(input_path=input_structure,
                     output_path=output_babel_h,
                     properties=prop)
```

```
[6]: # Babel_minimize: Structure energy minimization of a small molecule after being modified adding hydrogen atom:
# Import module
from biobb_chemistry.babelm.babel_minimize import babel_minimize
```



```
[7]: # Create and launch bb
babel_minimize(input_path=output_babel_h,
                output_path=output_babel_min,
                properties=prop)
```

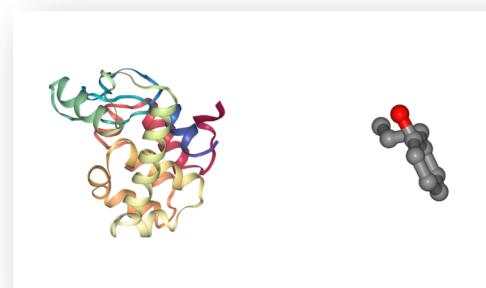
```
[8]: # Acyppe_params_gmx: Generation of topologies for GROMACS with ACYPE
# Import module
from biobb_chemistry.acpype.acpype_params_gmx import acpype_params_gmx
```



```
# Create prop dict and inputs/outputs
output_acpype_gro = ligandCode + 'params.gro'
output_acpype_itp = ligandCode + 'params.itp'
output_acpype_top = ligandCode + 'params.top'
output_acpype = ligandCode + 'params'
prop = {
    'basename' : output_acpype,
    'charge' : mol_charge
}
```

```
#Create and launch bb
acpype_params_gmx(input_path=output_babel_min,
                   output_path_gro=output_acpype_gro,
                   output_path_itp=output_acpype_itp,
                   output_path_top=output_acpype_top,
                   properties=prop)
```

Part 3: Parameterizing small molecule (topology)



5. Create ligand system topology

Questions (5)

YES =

Raise Your Hand



- Have you been following so far? (Y/N)**

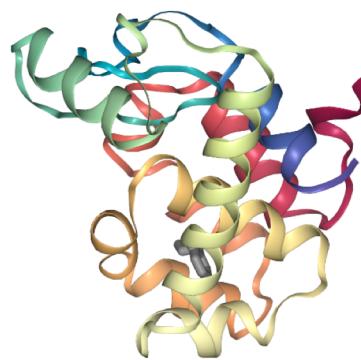
- Are you familiar with the concept of restraints in MD? (Y/N)**

- Are you familiar with the concept of force field parameters? (Y/N)**

Part 4: Generating protein-ligand complex structure

6. Preparing Ligand Restraints

7. Create new protein-ligand complex structure file



Ligand Restraints

Terms	Abbreviation
Molecule and residue information	units
Atom names	names
Atom types	types
Atomic charges	charges
Atomic connectivities	connects
Atomic coordinates	coords
Atomic masses	masses
Bonded parameters (bond, angle, dihedral)	bond params
Nonbonded parameters (electrostatic, VDW)	nonb params

Structure – Topology atom names matching

Questions (6)

YES =

Raise Your Hand



- Have you been following so far? (Y/N)**
- Have you ever setup/run a MD simulation of a protein-ligand complex using GROMACS? (Y/N)**

Part 5: Generating protein-ligand complex topology

```
; Include forcefield parameters
#include "amber99sb-ildn.ff/forcefield.itp"

; Including ligand ITP
#include "JZ4params.itp" ←
```

[moleculetype]
; Name nrexcl
Protein_chain_A 3

8. Create new protein-ligand complex topology file

```
; Include topology for ions
#include "amber99sb-ildn.ff/ions.itp"

[ system ]
; Name
Protein

[ molecules ]
; Compound #mols
Protein_chain_A 1
JZ4params 1
```

Questions (7)

YES =

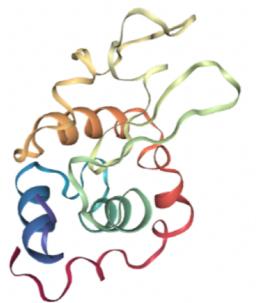
Raise Your Hand



- Have you been following so far? (Y/N)**
- Did you follow the previous GROMACS hands-on session? (Y/N)**

Part 6: MD Setup

 **GROMACS PROTEIN MD SETUP** 2024.1



This tutorial aims to illustrate the process of setting up a simulation system containing a protein, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Lysozyme protein (PDB code 1AKI).

[WorkflowHub](#) [Launch](#) [Download](#)

[View tutorial](#) [Open Github repository](#) [Open documentation](#)

(*) MyBinder provides a free, online version of Jupyter Lab. Take into account that the provided resources are finite and, in some notebooks. Please be patient and don't try to execute several notebooks at the same time.

- 9. Create Solvent Box**
- 10. Fill the Box with Water Molecules**
- 11. Adding Ions**
- 12. Energetically Minimize the System**
- 13. Equilibrate the System (NVT)**
- 14. Equilibrate the System (NPT)**
- 15. Free Molecular Dynamics Simulation**

Questions (8)

YES =

Raise Your Hand

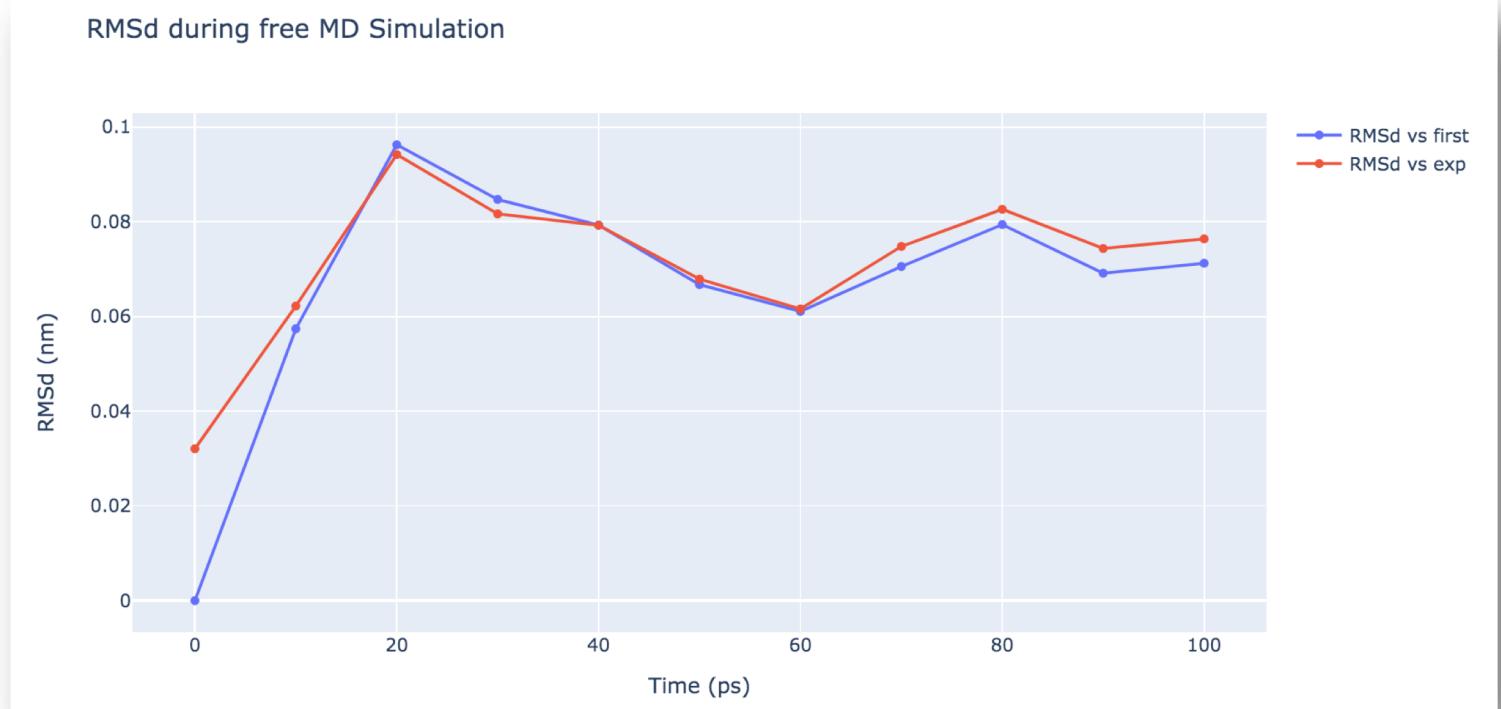


- Have you been following so far? (Y/N)**
- Are you familiar with GROMACS analysis tools? (Y/N)**
- Are you familiar with imaging process (periodicity issues)? (Y/N)**

Part 7: Trajectory post-processing

16. Post-processing and Visualizing Resulting 3D Trajectory

17. Output Files



Final Questions

YES =

Raise Your Hand



- Have you been able to follow the tutorial? (Y/N)**
- Would you like to know more about the BioExcel Building Blocks? (Y/N)**

Follow up:

- <https://ask.bioexcel.eu/c/BioExcel-Building-Blocks-library>
- <https://github.com/bioexcel/biobb/issues>

Acknowledgments



INSTITUTE
FOR RESEARCH
IN BIOMEDICINE



Federica Battistini



Genís Bayarri



Modesto Orozco



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



Pau Andrio



Josep Ll. Gelpí



Centre of Excellence for Computational Biomolecular Research

BioExcel Center of Excellence, funded from the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 675728, 823830, 101093290 (BioExcel-1, BioExcel-2 and BioExcel-3).



**Co-funded by
the European Union**



EuroHPC
Joint Undertaking