BioBB-Wfs and **BioBB-API**, integrated web-based platform and programmatic interface for biomolecular simulations workflows using the **BioExcel Building Blocks** library

BioExcel Webinar, 2023-05-23

Adam Hospital

adam.hospital@irbbarcelona.org

Index

Index

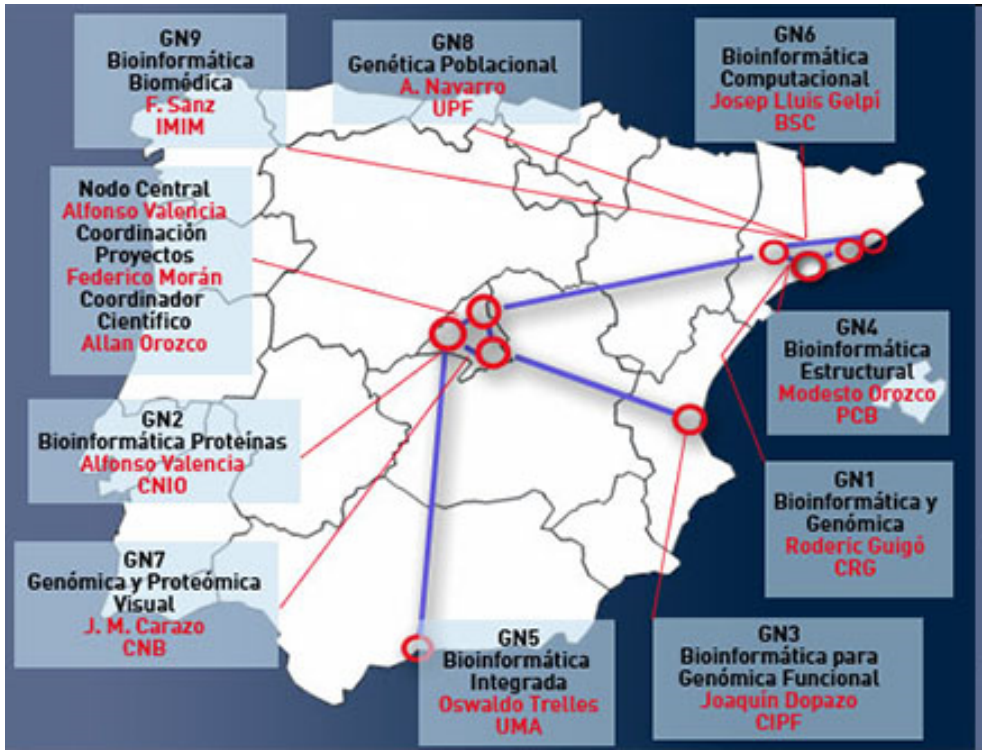
2008 - 2012

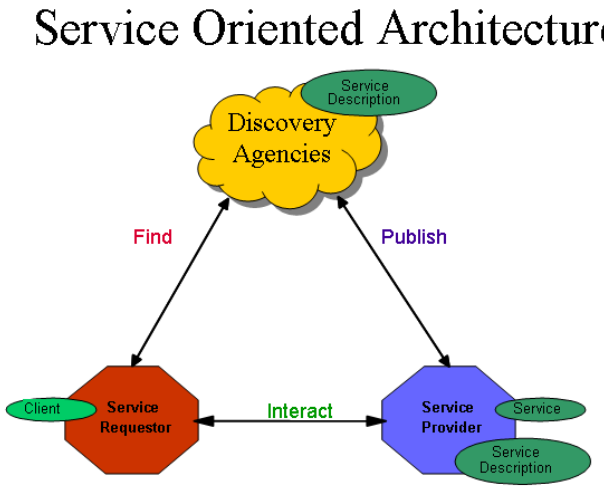
Index

Introduction | **Story** | BioBB | Aim

2008 - 2012

INB: A National Network for coordination, integration and development of Spanish Bioinformatics Resources in genomics and proteomics projects.





Web – Services:
A **software system** designed to support interoperable **machine-to-machine interaction** over a network.

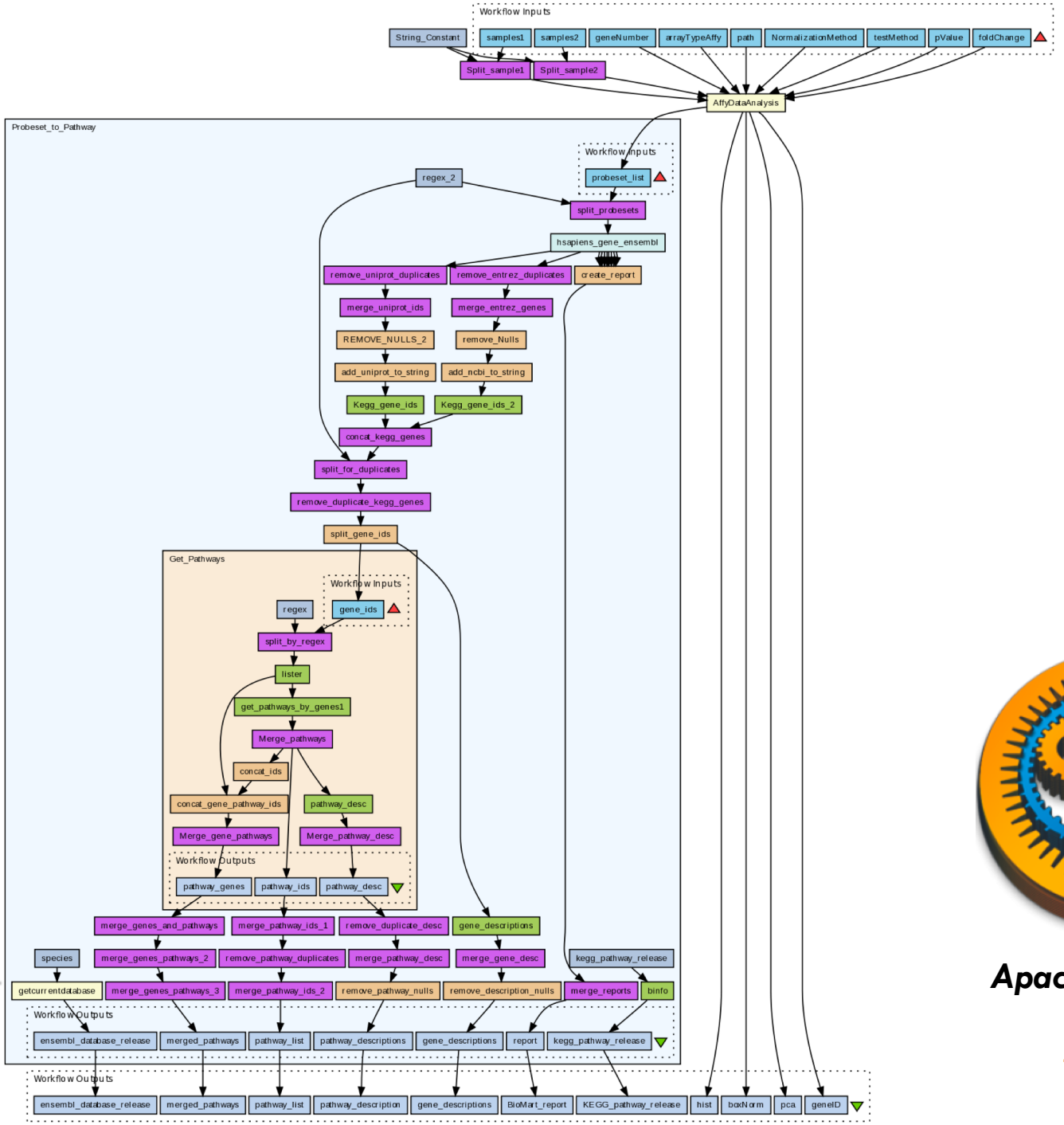
2008 - 2012

'Human Microarray CEL file to candidate pathways.'

- MicroArray Data.
- R statistics.
- Databases Connection.

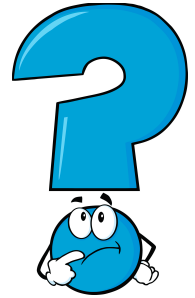
www.myexperiment.org

https://www.myexperiment.org/workflows/10.html

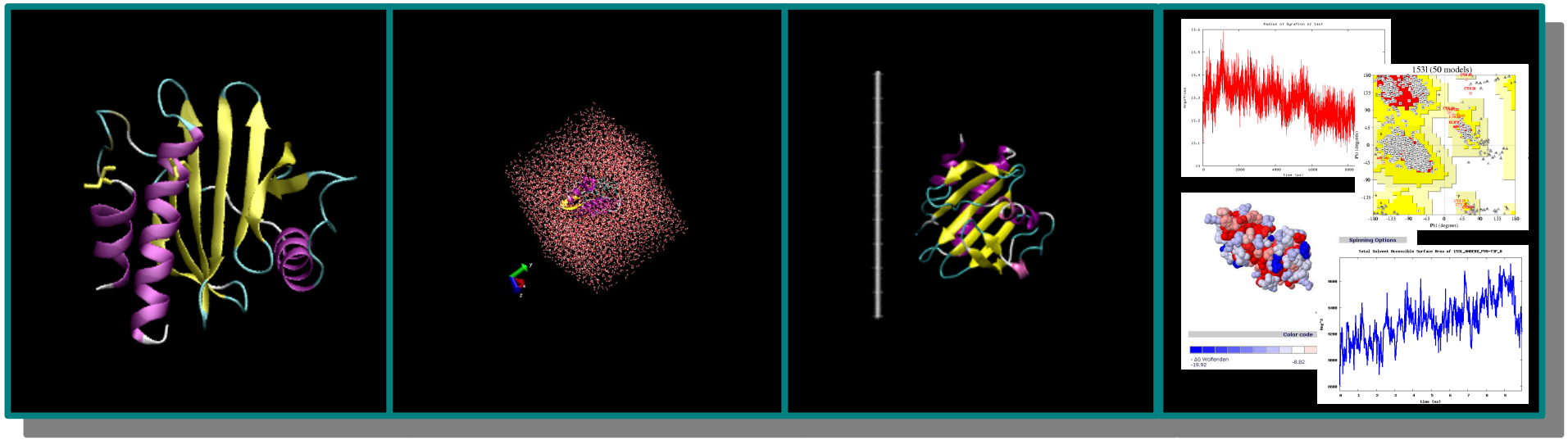


Apache Taverna

2008 - 2012

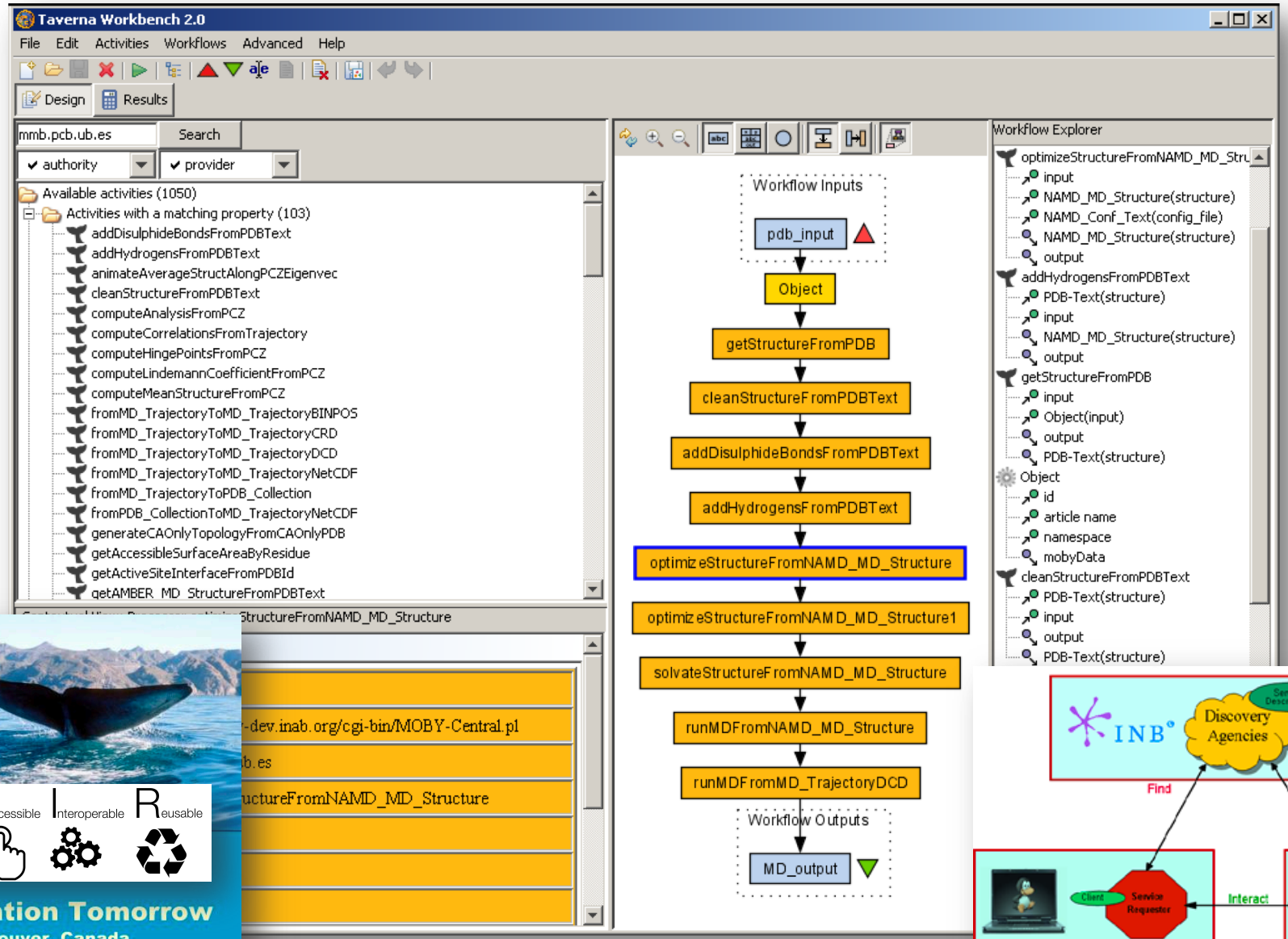


Preparation → MD run → Trajectory Analysis

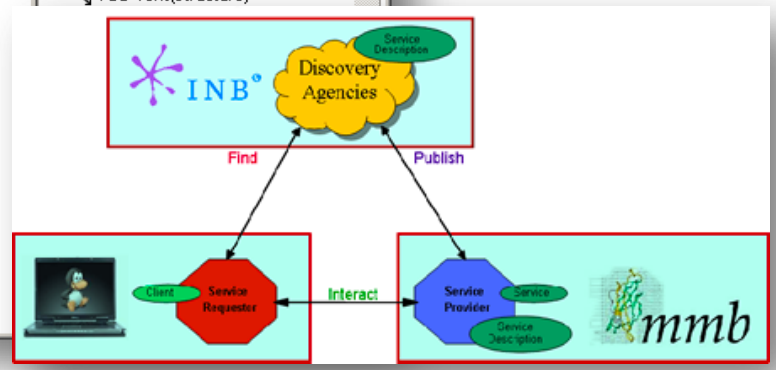


Protein Structure | Protein-Solvent System | Protein Dynamics (Trajectory) | Trajectory Analysis

2008 - 2012


Taverna Workbench 2.0
File Edit Activities Workflows Advanced Help
Design Results
mmb.pcb.ub.es Search
authority provider
Available activities (1050)
Activities with a matching property (103)
addDisulphideBondsFromPDBText
addHydrogensFromPDBText
animateAverageStructAlongPCZEigenvec
cleanStructureFromPDBText
computeAnalysisFromPCZ
computeCorrelationsFromTrajectory
computeHingePointsFromPCZ
computeLindemannCoefficientFromPCZ
computeMeanStructureFromPCZ
fromMD_TrajectoryToMD_TrajectoryBINPOS
fromMD_TrajectoryToMD_TrajectoryCRD
fromMD_TrajectoryToMD_TrajectoryDCD
fromMD_TrajectoryToMD_TrajectoryNetCDF
fromMD_TrajectoryToPDB_Collection
fromPDB_CollectionToMD_TrajectoryNetCDF
generateCAOnlyTopologyFromCAOnlyPDB
getAccessibleSurfaceAreaByResidue
getActiveSiteInterfaceFromPDBId
getAMBER_MD_StructureFromPDBText
Workflow Inputs
pdb_input
Object
getStructureFromPDB
cleanStructureFromPDBText
addDisulphideBondsFromPDBText
addHydrogensFromPDBText
optimizeStructureFromNAMD_MD_Structure
optimizeStructureFromNAMD_MD_Structure1
solvateStructureFromNAMD_MD_Structure
runMDFromNAMD_MD_Structure
runMDFromMD_TrajectoryDCD
Workflow Outputs
MD_output
Workflow Explorer
optimizeStructureFromNAMD_MD_Stru
input
NAMD_MD_Structure(structure)
NAMD_Conf_Text(config_file)
NAMD_MD_Structure(structure)
output
addHydrogensFromPDBText
PDB-Text(structure)
input
NAMD_MD_Structure(structure)
output
getStructureFromPDB
input
Object(input)
output
PDB-Text(structure)
Object
id
article name
namespace
mobyData
cleanStructureFromPDBText
PDB-Text(structure)
input
output
PDB-Text(structure)


moby
Findable Accessible Interoperable Reusable
BioMOBY
Interoperability today, Integration Tomorrow
Mark Wilkinson, iCAPTURE Centre, UBC, Vancouver, Canada


INB
Discovery Agencies
Service Description
Find
Publish
Client
Service Requester
Interact
Service Provider
Service
Service Description
mmb

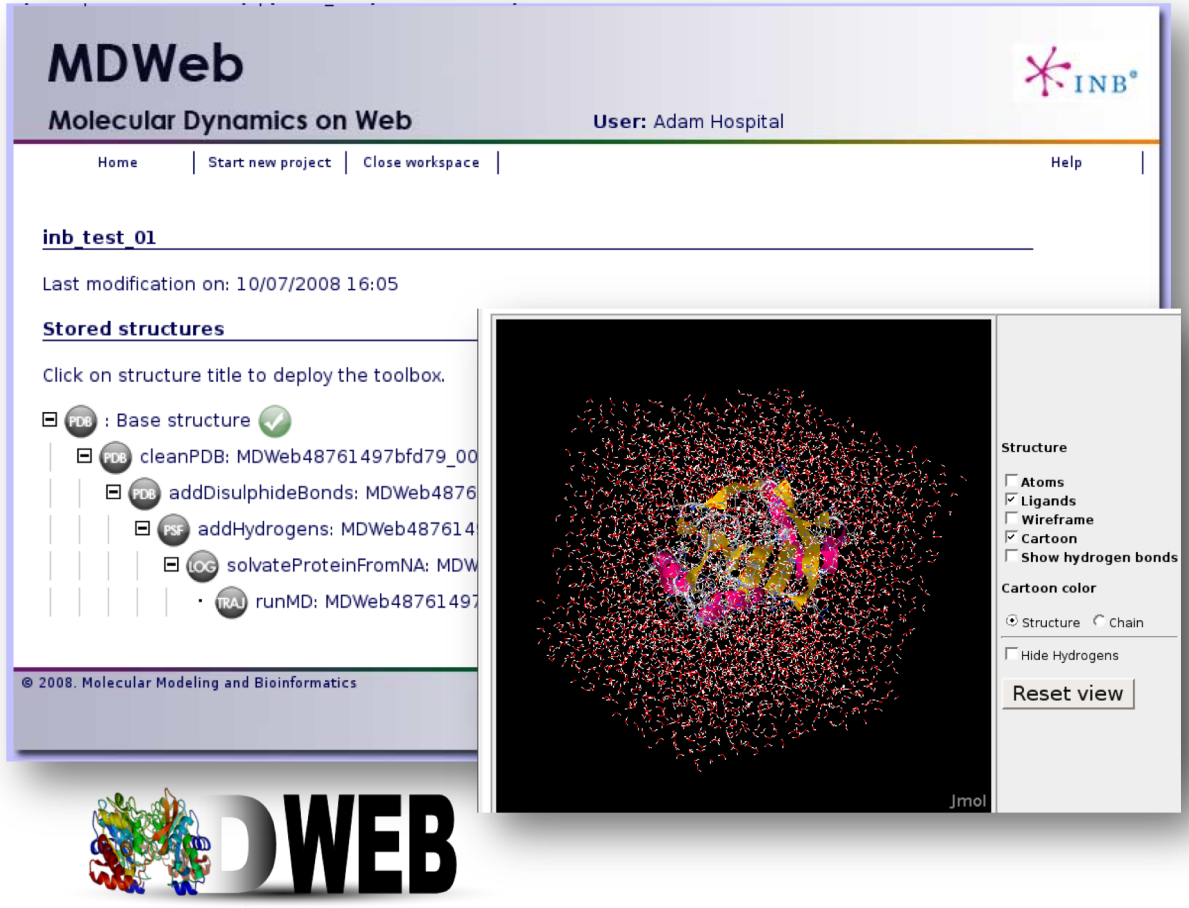
2008 - 2012

Scripting (Perl)

```
########## getStructurePDB Service ##########

print "\n1.- Running getStructurePDB Service...\n";

my $object = Object->new($pdb,'PDB');

my $pdbStruct = getStructureFromPDB (
        'structure' => $object
        ) -> {'structure'};
if ($@) {
        print "Service execution failed:\n$@";
        exit;
}

print OUT "\#getStructureFromPDB:\n";
print OUT $pdbStruct->content;
```

MDMoby

MDWeb (GUI)



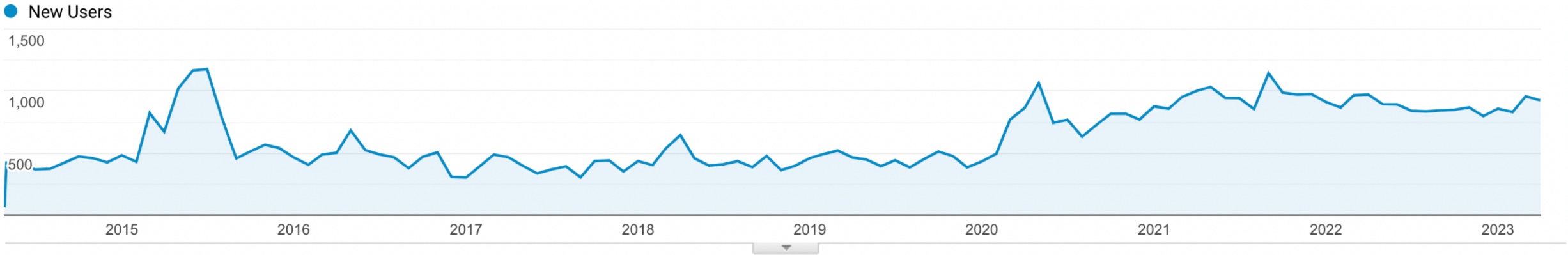
2008 - 2012

JOURNAL ARTICLE

MDWeb and MDMoby: an integrated web-based platform for molecular dynamics simulations

Adam Hospital, Pau Andrio, Carles Fenollosa, Damjan Cicin-Sain, Modesto Orozco ✉,
Josep Lluís Gelpí ✉ Author Notes

Bioinformatics, Volume 28, Issue 9, May 2012, Pages 1278–1279,
https://doi.org/10.1093/bioinformatics/bts139
Published: 21 March 2012 **Article history** ▾



BioExcel Webinar Series #3: ***Atomistic Molecular Dynamics Setup with MDWeb***

2016 - 2020

Index

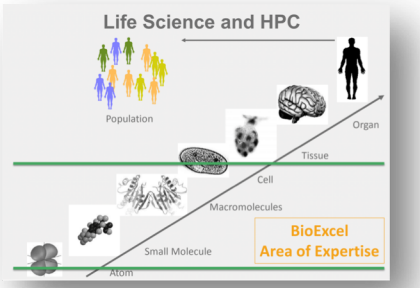
Introduction | Story | **BioBB** | Aim

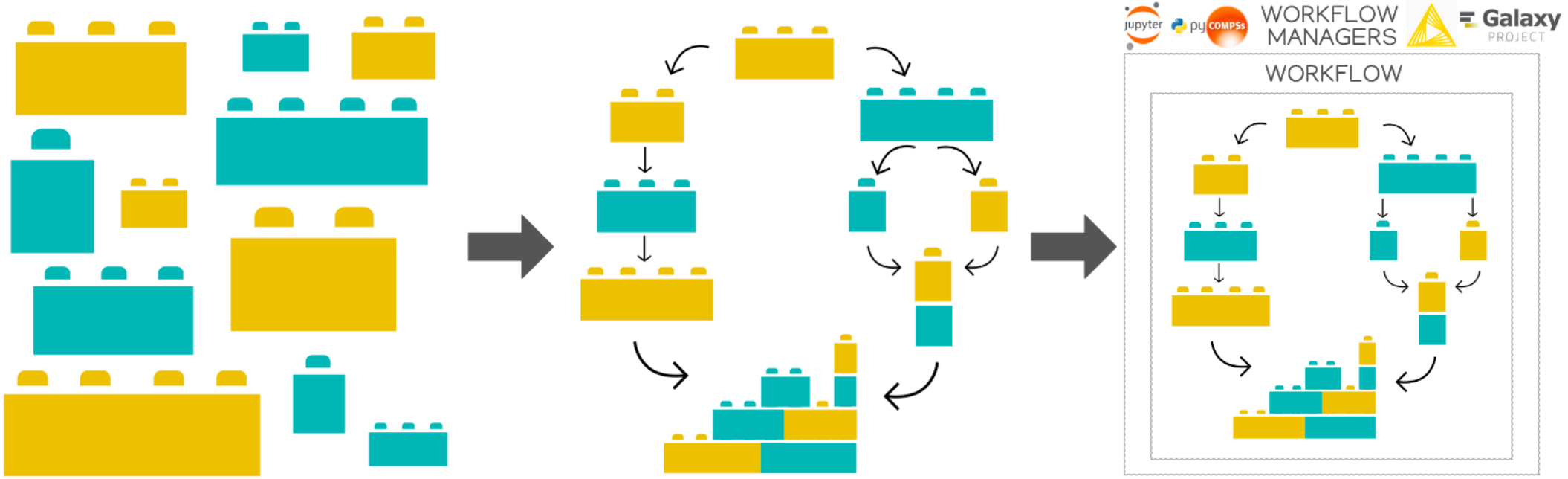
2016 - 2020

bioexcel

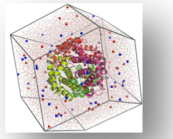
Centre of Excellence for Computational Biomolecular Research

A central hub for biomolecular modelling and simulations

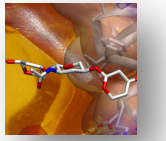




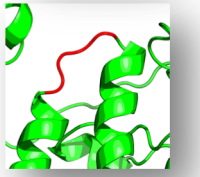
2016 - 2020



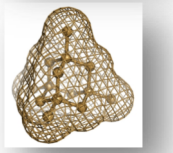
biobb_gromacs
Molecular Dynamics GROMACS



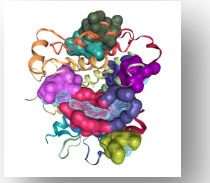
biobb_amber
Molecular Dynamics AMBER



biobb_model
Molecular Modelling



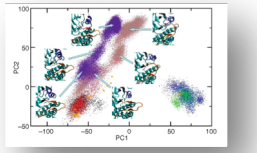
biobb_chemistry
Chemoinformatics functionalities



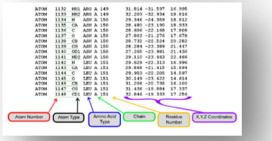
biobb_vs
Virtual Screening

biobb_cp2k
Quantum Mechanics

biobb_io
Biological databases

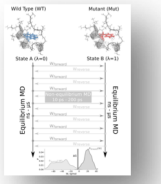


biobb_analysis
MD trajectories analysis

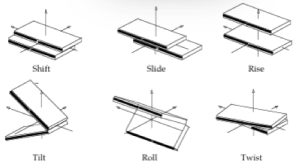


biobb_structure_utils
Modify or extract information from PDB

biobb_ml
Machine learning



biobb_pmx
Free energy calculations



biobb_dna
Nucleic Acids MD Trajectory analyses

biobb_vs

Virtual Screening

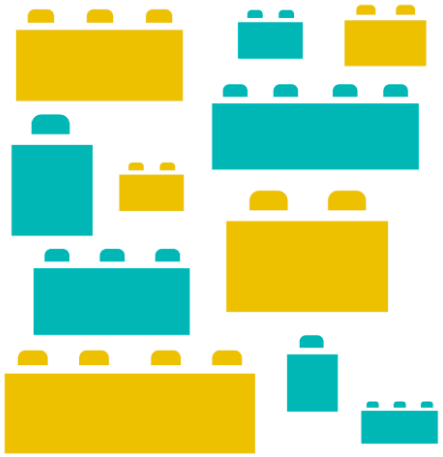
https://mmb.irbbarcelona.org/biobb/documentation/source

−	biobb_vs	Biobb_vs is the Biobb module collection to perform virtual screening studies.					

Building block	Wrapped tool	Description
AutoDockVinaRun	AutoDock Vina	Wrapper of the AutoDock Vina software.
BindingSite	in house using biopython	This class finds the binding site of the input_pdb.
Box	in house	This class sets the center and the size of a rectangular parallelepiped box around a set of residues or a pocket.
BoxResidues	in house using biopython	This class sets the center and the size of a rectangular parallelepiped box around a set of residues.
ExtractModelPDBQT	in house using biopython	Extracts a model from a PDBQT file with several models.
FPocketRun	fpocket	Wrapper of the fpocket software for finding the binding sites of a structure.
FPocketFilter	in house	Performs a search over the outputs of the fpocket building block.
FPocketSelect	in house	Selects a single pocket in the outputs of the fpocket building block..

biobb_gromacs

Molecular Dynamics GROMACS



Building block	Wrapped tool	Description
Pdb2gmx	gmx pdb2gmx	Creates a compressed (ZIP) GROMACS topology (TOP and ITP files) from a given PDB file.
Editconf	gmx editconf	Creates a GROMACS structure file (GRO) adding the information of the solvent box to the input structure file.
Genion	gmx genion	Creates a new compressed GROMACS topology adding ions until reaching the desired concentration to the input compressed GROMACS topology.
Genrestr	gmx genrestr	Creates a new GROMACS compressed topology applying the indicated force restrains to the given input compressed topology.
Grompp	gmx grompp	Creates a GROMACS portable binary run input file (TPR) applying the desired properties from the input compressed GROMACS topology.
Mdrun	gmx mdrun	Performs molecular dynamics simulations from an input GROMACS TPR file.
GromppMdrun	gmx grompp & gmx mdrun	Combination of Grompp & Mdrun blocks. Grompp, creates a GROMACS portable binary run input file (TPR) applying the desired properties from the input compressed GROMACS topology. Mdrun, performs molecular dynamics simulations from an input GROMACS TPR file.
MakeNdx	gmx make_ndx	Creates a GROMACS index file (NDX) from an input selection and an input GROMACS structure file.
Gmxselect	gmx select	Creates a GROMACS index file (NDX) from an input selection and an input GROMACS structure file.
Solvate	gmx solvate	Creates a new compressed GROMACS topology file adding solvent molecules to a given input compressed GROMACS topology file.
Trjcat	gmx trjcat	Concatenates two or more GROMACS trajectory files.
Ndx2resttop	in house	Creates a new GROMACS compressed topology applying the force restrains to the input groups in the input index file to the given input compressed topology.
AppendLigand	in house	Takes a ligand ITP file and inserts it in a topology.

2016 - 2020

CONDA®

Python

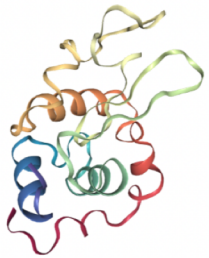
jupyter

COMMON WORKFLOW LANGUAGE

Galaxy

GROMACS PROTEIN MD SETUP

2023.1



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).

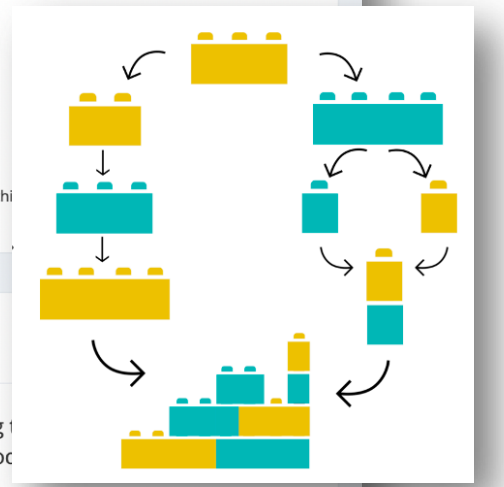
WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

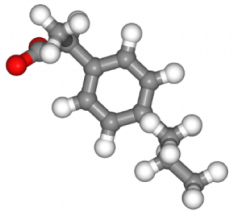
View tutorial | Open Github repository | Open documentation

gmx md protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launchi
the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.



AUTOMATIC LIGAND PARAMETERIZATION



This tutorial aims to illustrate the process of ligand parameterization for a small molecule, step by step, using t
library (biobb). The particular example used is the Ibuprofen small compound (3-letter code IBP, Drugbank coc
anti-inflammatory drug (NSAID) derived from propionic acid and it is considered the first of the propionics.

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

View tutorial | Open Github repository | Open documentation

gmx ligand

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

2016 - 2020

scientific **data**

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific data › articles › article

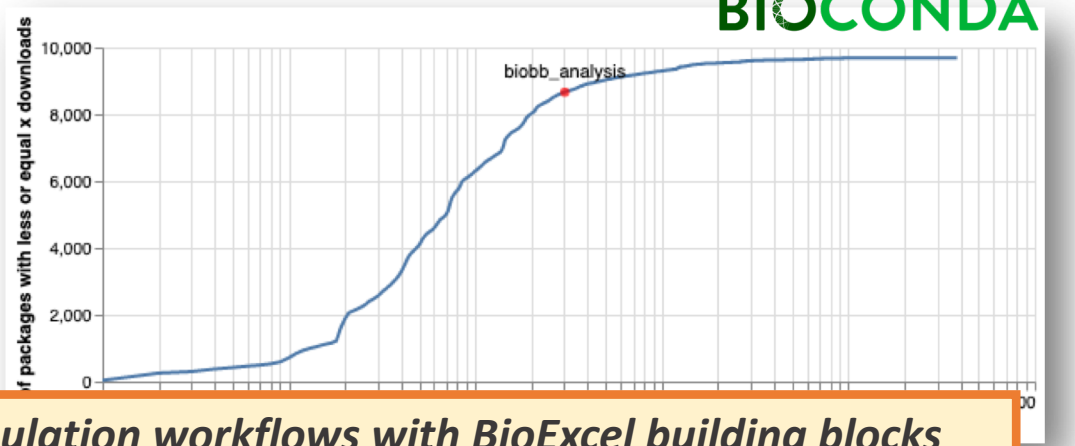
Article | Open Access | Published: 10 September 2019

BioExcel Building Blocks, a software library for interoperable biomolecular simulation workflows

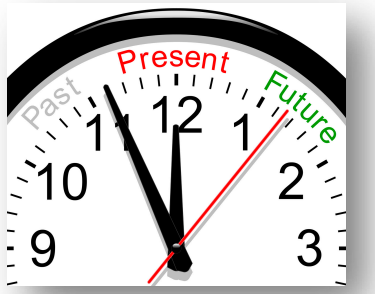
Pau Andrio, Adam Hospital, Javier Conejero, Luis Jordá, Marc Del Pino, Laia Codo, Stian Soiland-Reyes, Carole Goble, Daniele Lezzi, Rosa M. Badia, Modesto Orozco & Josep Ll. Gelpi

Scientific Data **6**, Article number: 169 (2019) | Cite this article

BIOCONDA



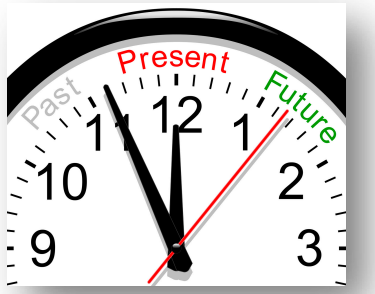
BioExcel Webinar Series #48: ***Computational biomolecular simulation workflows with BioExcel building blocks***



2020 - 2023

Index

Introduction | Story | BioBB | **Aim**



2020 - 2023

Main objectives:

- Use **BioBB** to update/replace old **MDMoby/MDWeb**
  - New **program versions**
  - Extended **functionalities**
  - **REST-API** vs old SOAP Web Services
  - New **web-based technology** (GUI)

- Offer **easy access** to **biomolecular simulation** processes

BioBB REST-API

BioBB Workflows GUI

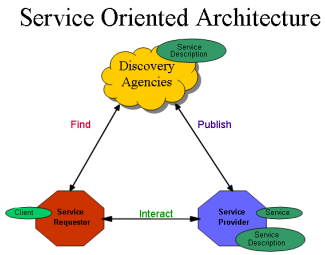
Index

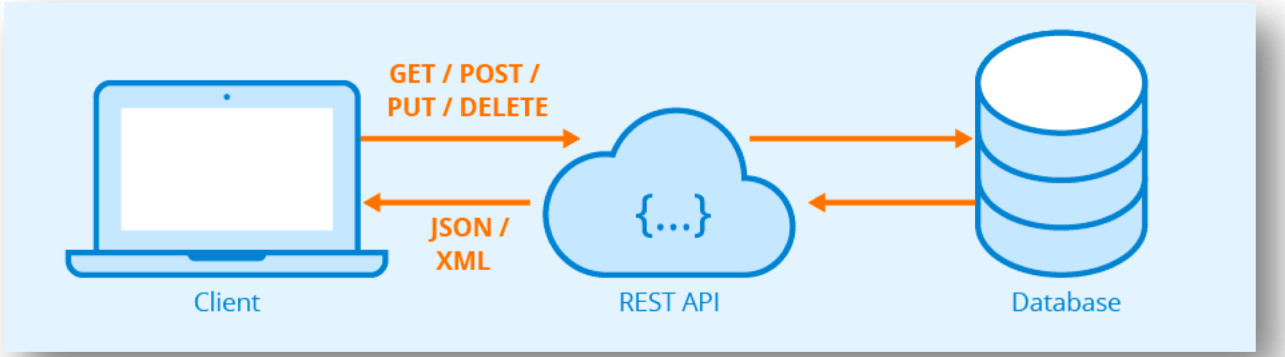
Index

BioBB API | Concept | Implementation | Usage

BioBB REST-API

- **Remote** and **programmatic access** to biomolecular simulation processes
- Processes are run in the **provider's infrastructure** (with limitations)
- **No needs** of any **installation** / **deployment**





Programmatic Access

What is an API endpoint?

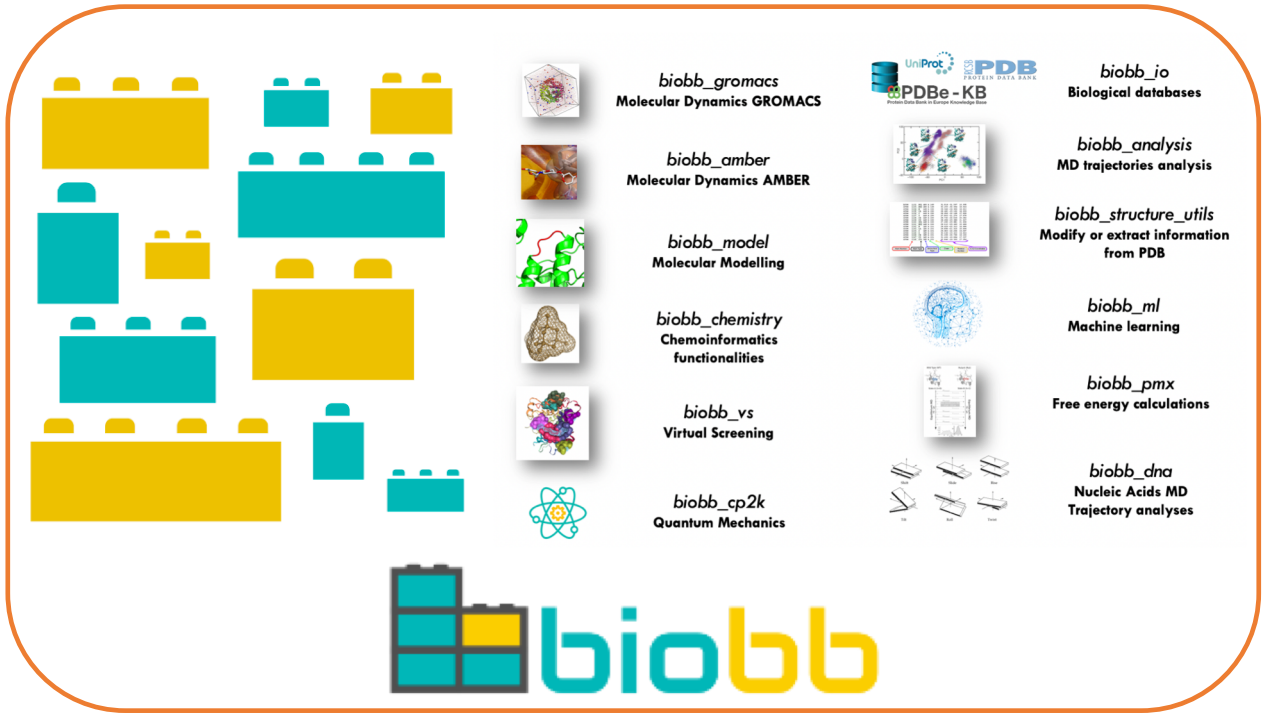
An API endpoint is a digital location where an API receives requests about a specific resource on its server. In APIs, an endpoint is typically a uniform resource locator (URL) that provides the location of a resource on the server.

https://blog.hubspot.com/

PDBe REST API

Programmatic access to PDBe data

https://www.ebi.ac.uk/pdbe/api/pdb/entry/summary/2vgb



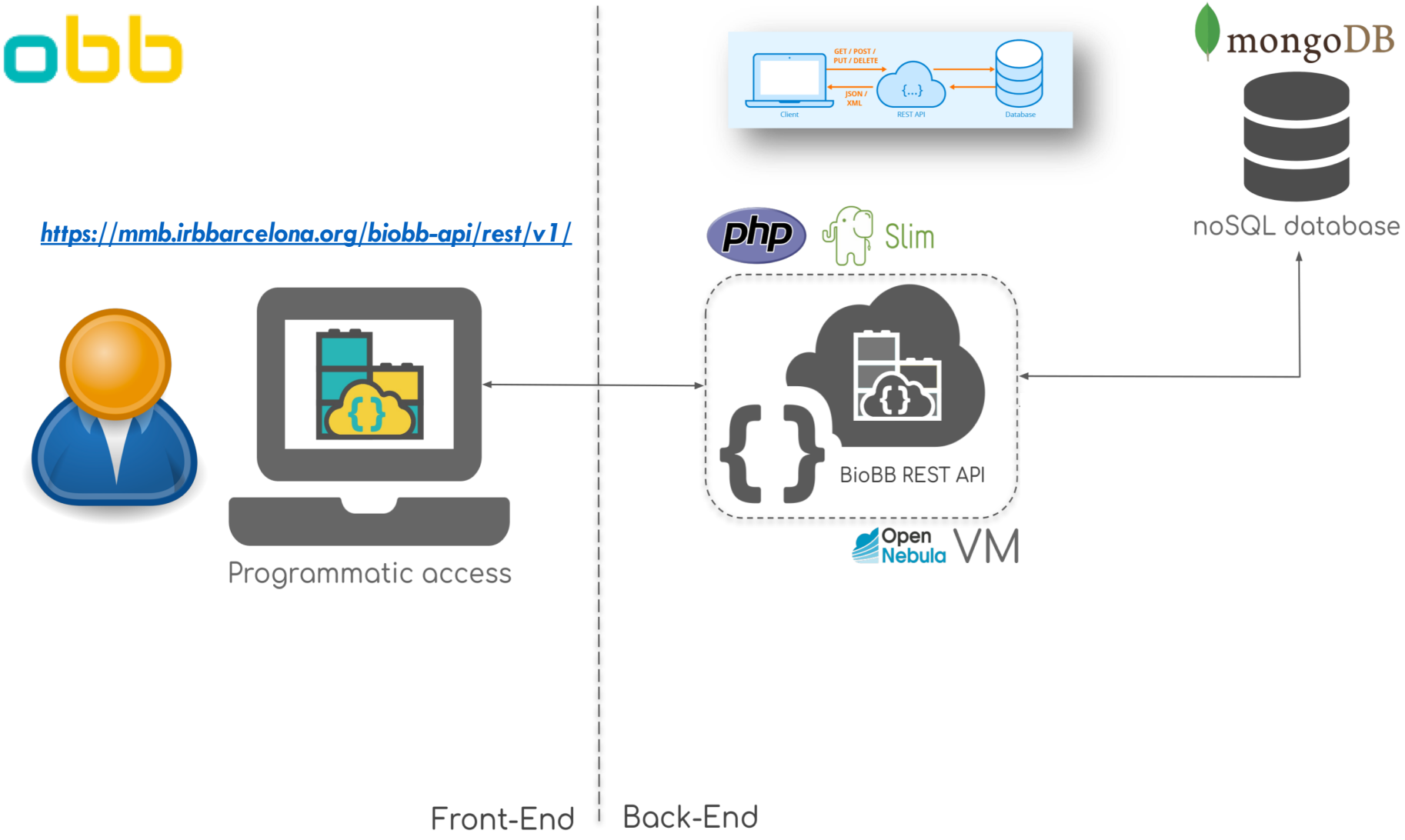
https://mmb.irbbarcelona.org/biobb-api/rest/v1/

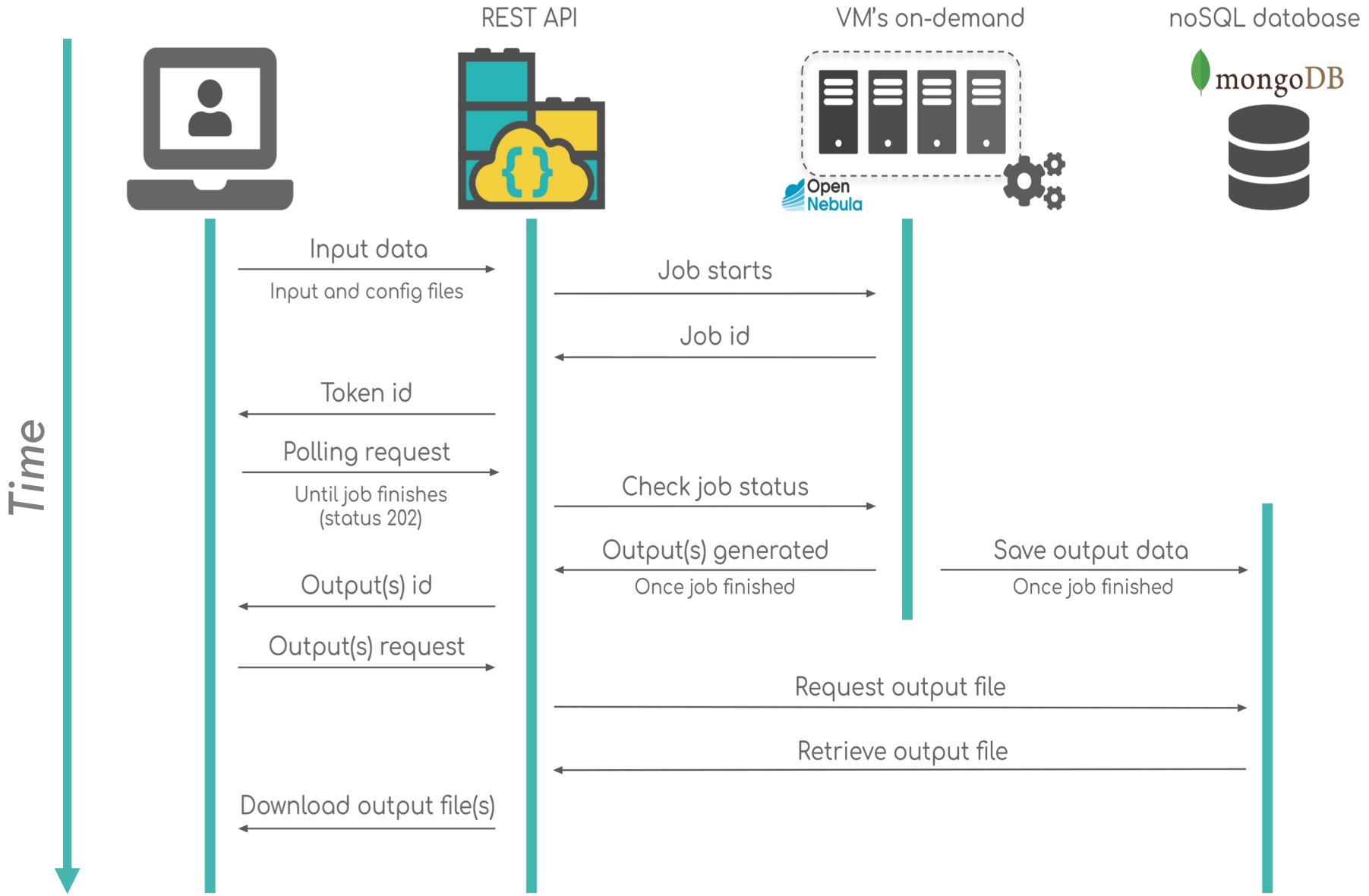
Package biobb_vs

GET	`/launch/biobb_vs/autodock_vina_run`	Get information of autodock_vina_run tool
POST	`/launch/biobb_vs/autodock_vina_run`	Launch a new job with autodock_vina_run tool
GET	`/launch/biobb_vs/bindingsite`	Get information of bindingsite tool
POST	`/launch/biobb_vs/bindingsite`	Launch a new job with bindingsite tool
GET	`/launch/biobb_vs/box`	Get information of box tool
POST	`/launch/biobb_vs/box`	Launch a new job with box tool
GET	`/launch/biobb_vs/box_residues`	Get information of box_residues tool
POST	`/launch/biobb_vs/box_residues`	Launch a new job with box_residues tool

BioBB API | Concept | **Implementation** | Usage

BioBB REST-API


https://mmb.irbbarcelona.org/biobb-api/rest/v1/
GET / POST / PUT / DELETE
JSON / XML
Client
REST API
Database
mongoDB
noSQL database
php
Slim
BioBB REST API
Open Nebula VM
Programmatic access
Front-End
Back-End


REST API
VM's on-demand
noSQL database
mongoDB
Open Nebula
Time
Input data
Input and config files
Job starts
Job id
Token id
Polling request
Until job finishes (status 202)
Check job status
Output(s) generated
Once job finished
Save output data
Once job finished
Output(s) id
Output(s) request
Request output file
Retrieve output file
Download output file(s)

BioBB API | Concept | Implementation | Usage

BioBB REST-API

https://mmb.irbbarcelona.org/biobb-api/

Home | Availability | Tutorials | Release Notes | Contact

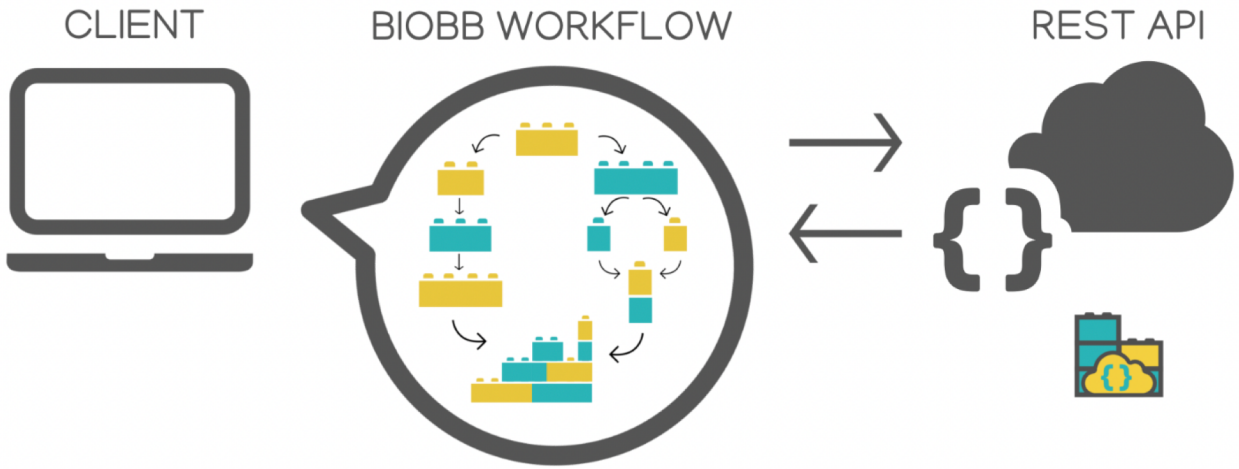
Welcome to BioBB REST API

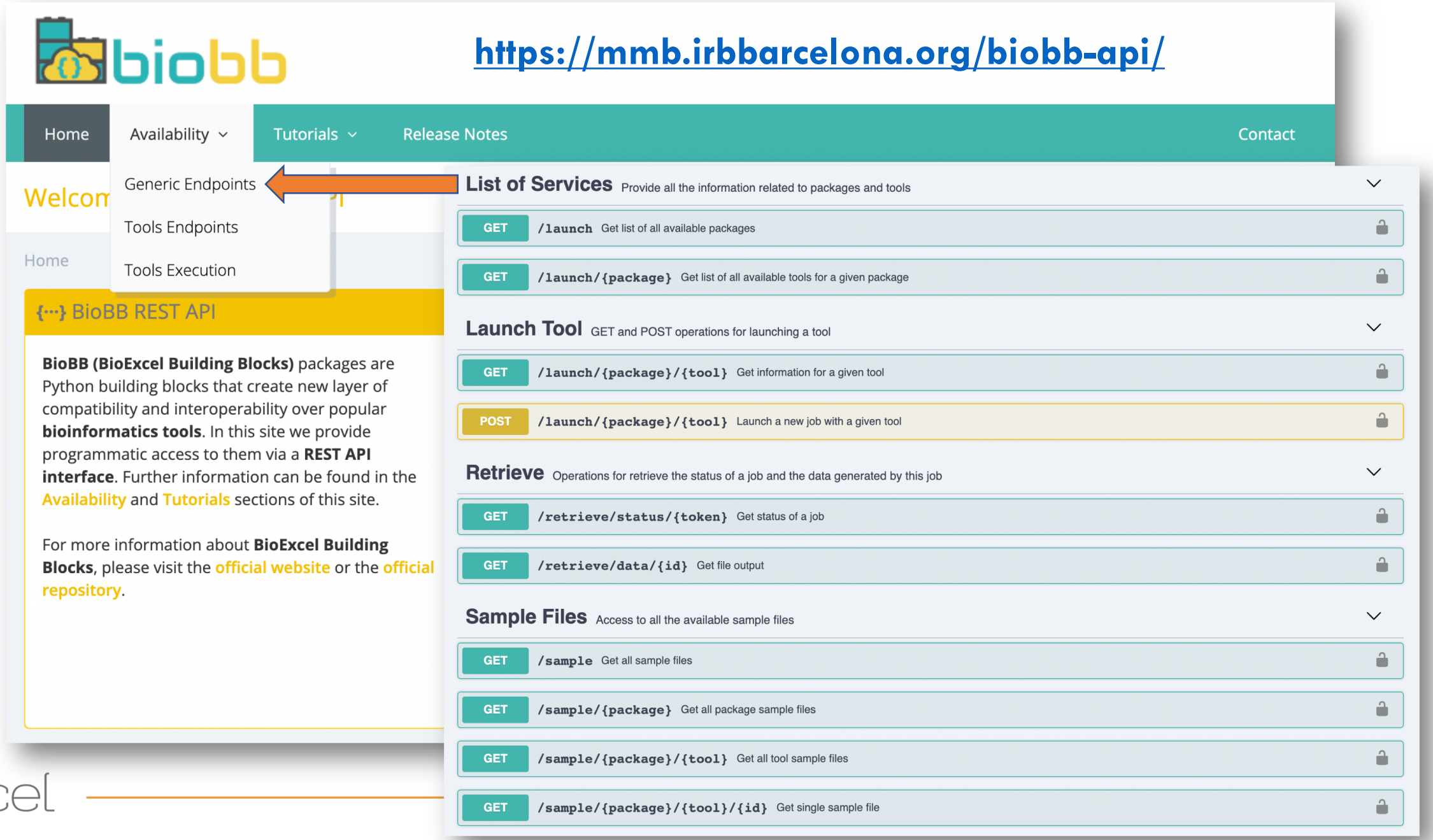
Home

{···} BioBB REST API

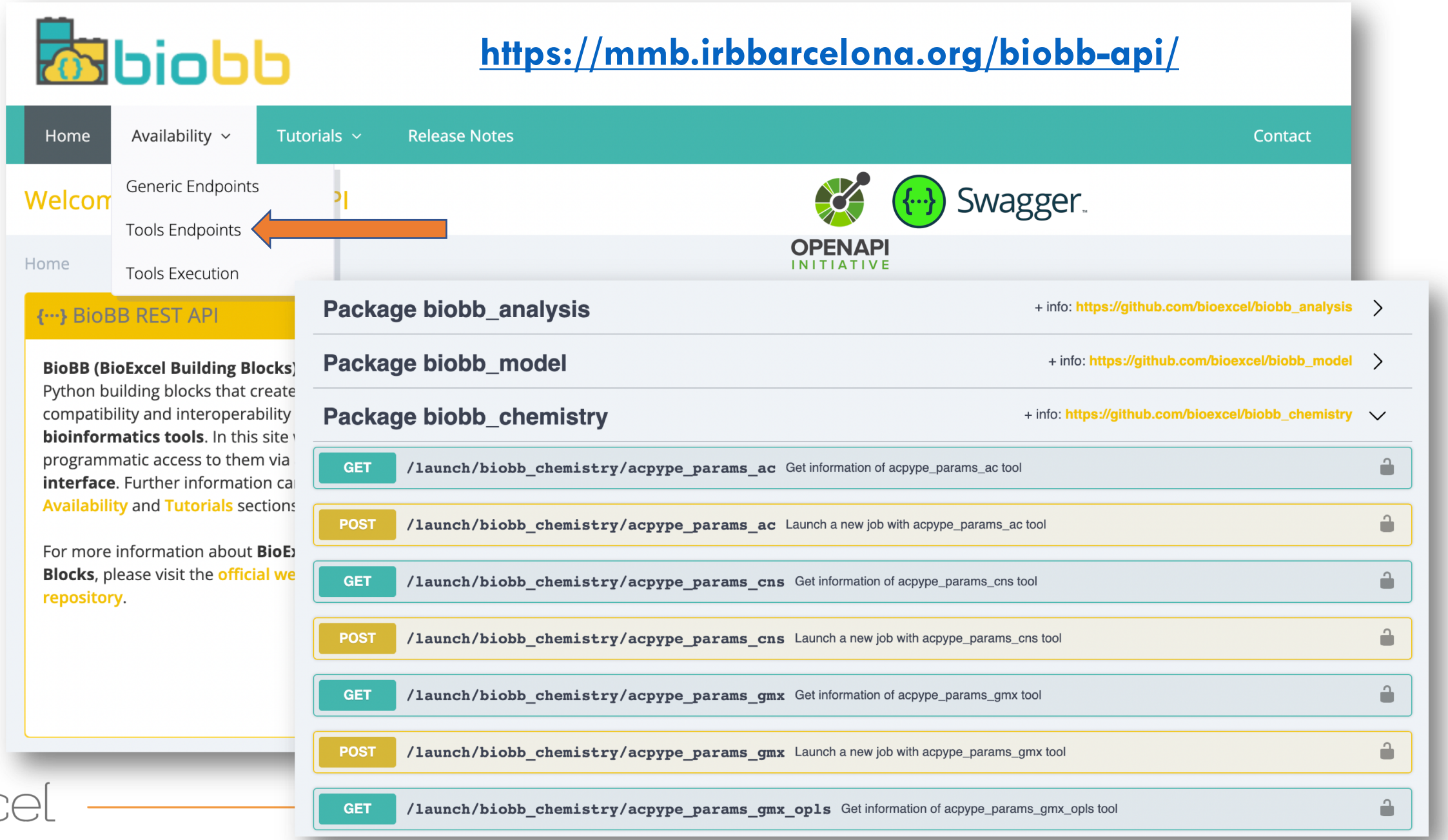
BioBB (BioExcel Building Blocks) packages are Python building blocks that create new layer of compatibility and interoperability over popular **bioinformatics tools**. In this site we provide programmatic access to them via a **REST API interface**. Further information can be found in the Availability and Tutorials sections of this site.

For more information about **BioExcel Building Blocks**, please visit the official website or the official repository.



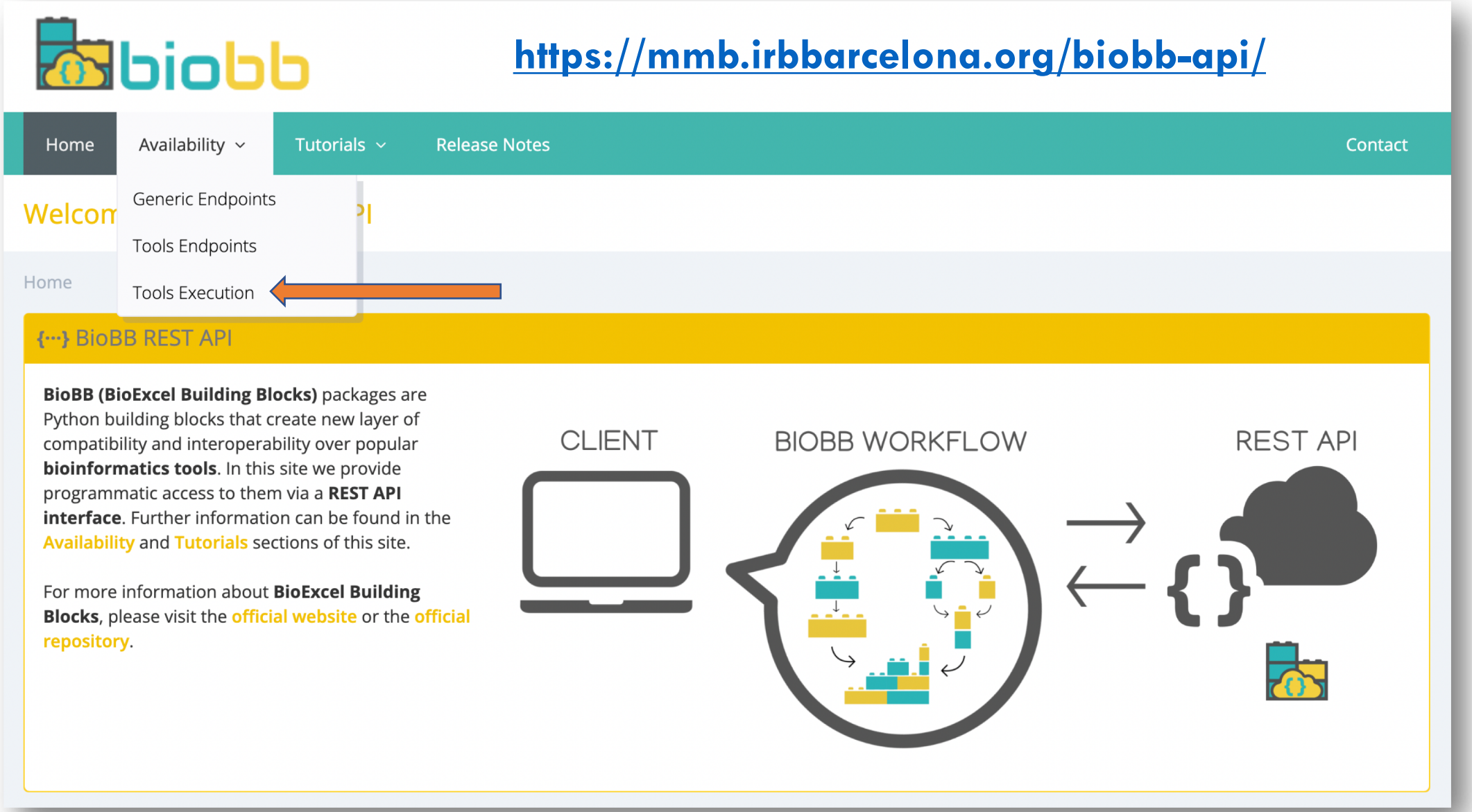

biobb
https://mmb.irbbarcelona.org/biobb-api/
Home
Availability
Tutorials
Release Notes
Contact
Generic Endpoints
Tools Endpoints
Tools Execution
Home
{···} BioBB REST API
BioBB (BioExcel Building Blocks) packages are Python building blocks that create new layer of compatibility and interoperability over popular bioinformatics tools. In this site we provide programmatic access to them via a REST API interface. Further information can be found in the Availability and Tutorials sections of this site.
For more information about BioExcel Building Blocks, please visit the official website or the official repository.
List of Services Provide all the information related to packages and tools
GET /launch Get list of all available packages
GET /launch/{package} Get list of all available tools for a given package
Launch Tool GET and POST operations for launching a tool
GET /launch/{package}/{tool} Get information for a given tool
POST /launch/{package}/{tool} Launch a new job with a given tool
Retrieve Operations for retrieve the status of a job and the data generated by this job
GET /retrieve/status/{token} Get status of a job
GET /retrieve/data/{id} Get file output
Sample Files Access to all the available sample files
GET /sample Get all sample files
GET /sample/{package} Get all package sample files
GET /sample/{package}/{tool} Get all tool sample files
GET /sample/{package}/{tool}/{id} Get single sample file

bioexcel


biobb
https://mmb.irbbarcelona.org/biobb-api/
Home
Availability
Tutorials
Release Notes
Contact
Generic Endpoints
Tools Endpoints
Tools Execution
Swagger
OPENAPI
INITIATIVE
Home
{···} BioBB REST API
BioBB (BioExcel Building Blocks)
Python building blocks that create
compatibility and interoperability
bioinformatics tools. In this site
programmatic access to them via
interface. Further information ca
Availability and Tutorials section
For more information about BioE
Blocks, please visit the official we
repository.
Package biobb_analysis
+ info: https://github.com/bioexcel/biobb_analysis
Package biobb_model
+ info: https://github.com/bioexcel/biobb_model
Package biobb_chemistry
+ info: https://github.com/bioexcel/biobb_chemistry
GET /launch/biobb_chemistry/acpype_params_ac Get information of acpype_params_ac tool
POST /launch/biobb_chemistry/acpype_params_ac Launch a new job with acpype_params_ac tool
GET /launch/biobb_chemistry/acpype_params_cns Get information of acpype_params_cns tool
POST /launch/biobb_chemistry/acpype_params_cns Launch a new job with acpype_params_cns tool
GET /launch/biobb_chemistry/acpype_params_gmx Get information of acpype_params_gmx tool
POST /launch/biobb_chemistry/acpype_params_gmx Launch a new job with acpype_params_gmx tool
GET /launch/biobb_chemistry/acpype_params_gmx_opls Get information of acpype_params_gmx_opls tool

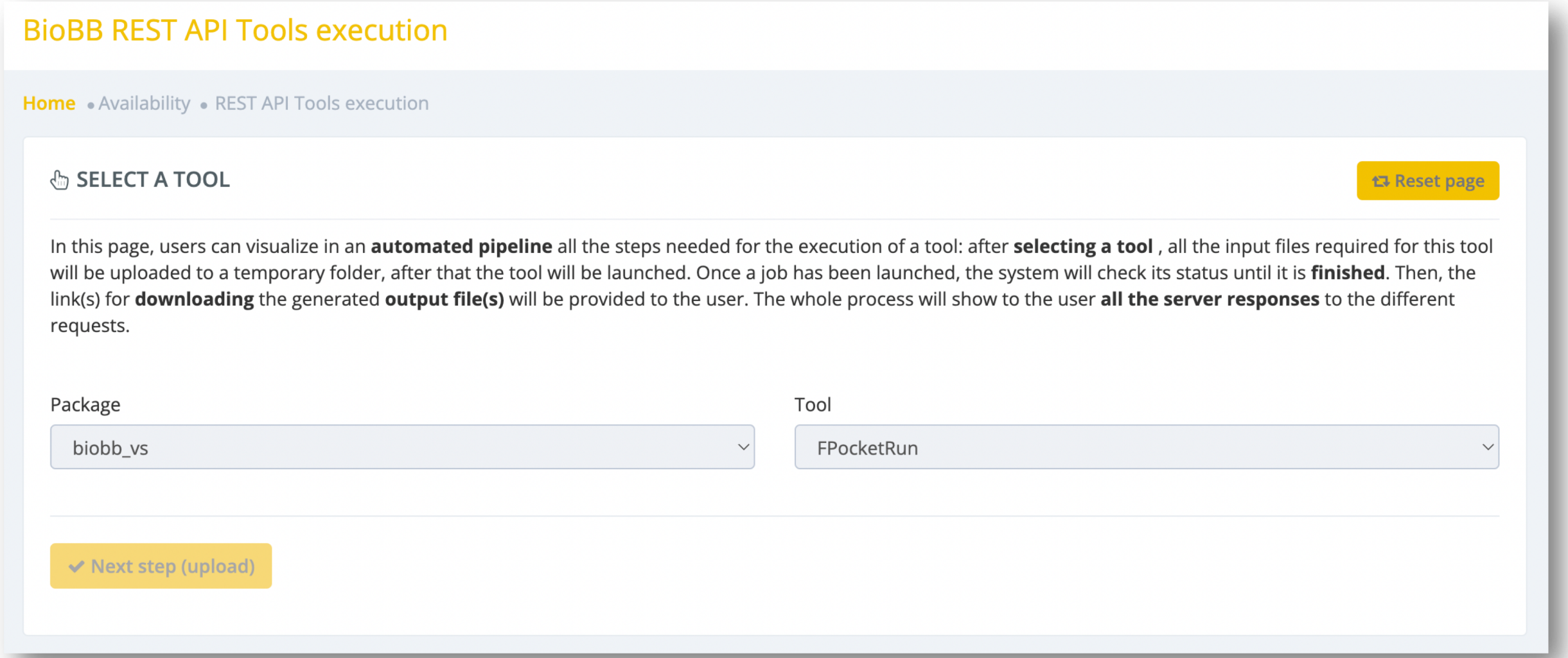
bioexcel

https://mmb.irbbarcelona.org/biobb-api/



– Step 1: Select Tool (Building Block)

*Run example: find **pockets** in a **protein** (with **fpocket**)*



– Step 2: Upload Data (Files)

*Run example: find **pockets** in a **protein** (with **fpocket**)*

UPLOAD DATA

First off, check the documentation for the selected **biobb_vs.fpocket_run** tool clicking here.

Users must upload the data for executing **biobb_vs.fpocket_run**. If you have no data, all the sample files for **biobb_vs.fpocket_run** tool can be downloaded clicking here.

All the fields marked with an * are mandatory.

```
{
  "properties": {
    "min_radius": 3,
    "max_radius": 6,
    "num_spheres": 35,
    "sort_by": "druggability_score"
  }
}
```

Configuration file for the fpock_run tool ⓘ

config_fpocket_run.json | Change | Remove

Path to the PDB structure where the binding site is to be found * ⓘ

7apu.pdb | Change | Remove

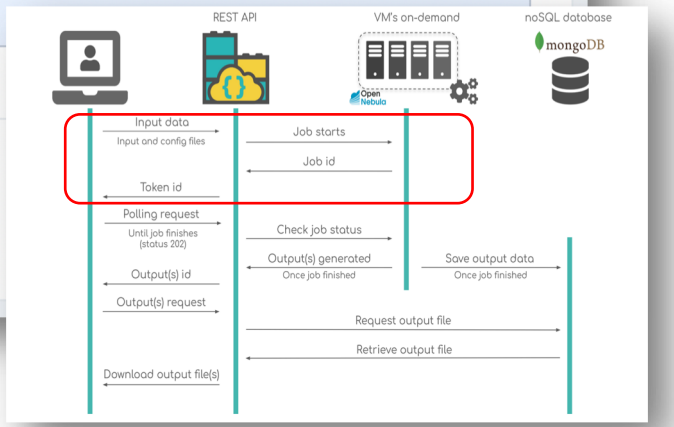
Path to all the pockets found by fpocket in the input_pdb_path structure * ⓘ

pockets_AH.zip

Path to the JSON summary file * ⓘ

pockets_AH.json

Next step (launch)



– Step 3: Check Status (polling)

*Run example: find **pockets** in a **protein** (with **fpocket**)*

CHECK STATUS

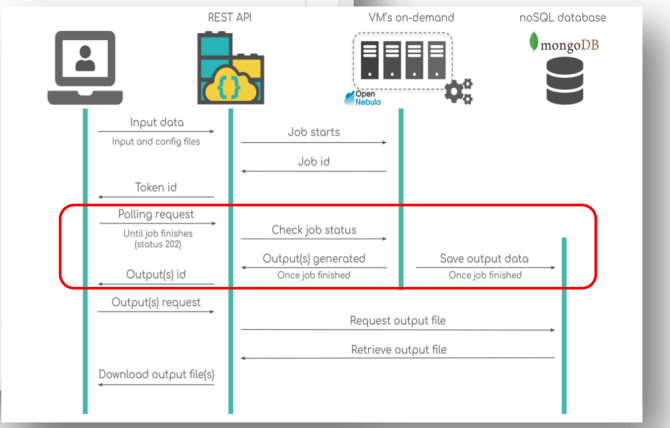
Checking the status of the job with the token given in the previous step. This endpoint will be automatically called while the response status is **202**. Once the response status is **200**, the output files are ready for downloading.

Endpoint

```
https://mmb.irbbarcelona.org/biobb-api/rest/v1/retrieve/status/7f78b371243e53164591b1717fb056e9052d81e4aad98dc1cd1112f40436d8cd234
```

JSON response


```
{
    "code": 200,
    "state": "FINISHED",
    "message": "The requested job has finished successfully, please go to /retrieve/data/{id} for each output_files.",
    "output_files": [
        {
            "id": "645cfc63e32014.74825855",
            "name": "pockets_AH.zip",
            "size": 200445,
            "mimetype": "application/zip"
        },
        {
            "id": "645cfc63e789a3.69986544",
            "name": "pockets_AH.json",
            "size": 7943,
            "mimetype": "application/json"
        }
    ],
    "expiration": "May 17, 2023 22:00 GMT+0000"
}
```

– Step 4: Retrieve Data

*Run example: find **pockets** in a **protein** (with **fpocket**)*

RETRIEVE DATA

Below there is a list with all the retrieve data endpoints and direct download link for each output file generated by the **biobb_vs.fpocket_run** tool.

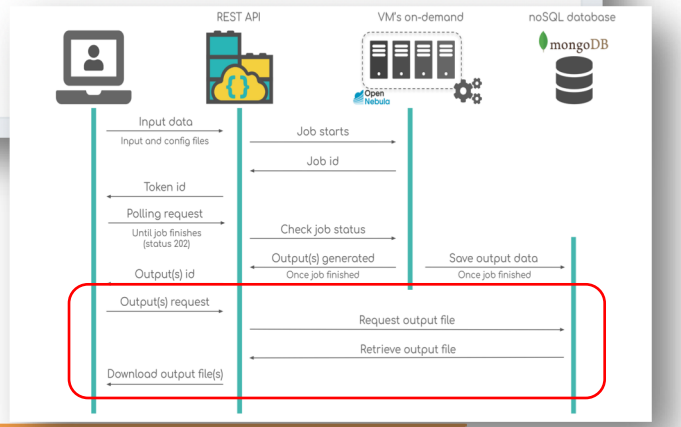
Output(s)

```
https://mmb.irbbarcelona.org/biobb-api/rest/v1/retrieve/data/645cfc63e32014.74825855
```

Download pockets_AH.zip (195.75 KB)

```
https://mmb.irbbarcelona.org/biobb-api/rest/v1/retrieve/data/645cfc63e789a3.69986544
```

Download pockets_AH.json (7.76 KB)



Fetching PDB structure

Downloading **PDB structure** with the **protein molecule** from the RCSB PDB database.
Alternatively, a **PDB file** can be used as starting structure.

BioBB REST API end points used:

- PDB from **biobb_io.api.pdb**

```
# Downloading desired PDB file

# Create properties dict and inputs/outputs
downloaded_pdb = pdbCode + '.pdb'
prop = {
    'pdb_code': pdbCode
}

# Launch bb on REST API
token = launch_job(url = apiURL + 'launch/biobb_io/pdb',
                   config = prop,
                   output_pdb_path = downloaded_pdb)
```

Python

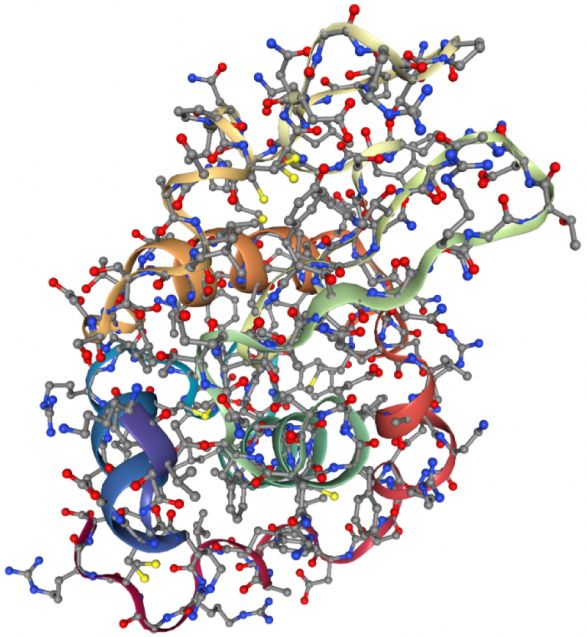
```
# Check job status
out_files = check_job(token, apiURL)
```

```
# Save generated file to disk
retrieve_data(out_files, apiURL)
```

Visualizing 3D structure

Visualizing the downloaded/given **PDB structure** using **NGL**:

```
# Show protein
view = nglview.show_structure_file(downloaded_pdb)
view.add_representation(repr_type='ball+stick', selection='all')
view._remote_call('setSize', target='Widget', args=['','600px'])
view
```



jupyter

Computing Protein Cavities (fpocket)

Computing the **protein cavities** (pockets) using the well-known fpocket tool.
These **cavities** will be then used in the **docking procedure** to try to find the **best region of the protein surface** where the small molecule can **bind**.

Although in this particular example we already know the **binding site** region, as we started from a **protein-ligand complex** structure where the ligand was located in the same **binding site** as **Imatinib** is binding, this is not always the case. In the cases where we do not know these regions, **fpocket** will help us identifying the possible **binding sites** of our **target protein**.

fpocket input parameters, such as **minimum** and **maximum radius** (in Angstroms) the alpha spheres might have in a **binding pocket** can be adjusted (min_radius, max_radius) . Parameters used in this particular example are 3Å for the **minimum radius** and 6Å for the **maximum radius**. The **minimum number of alpha spheres** a pocket must contain in order to figure in the results is also adjusted to 35. See the fpocket manual for more information.

Building Blocks used:

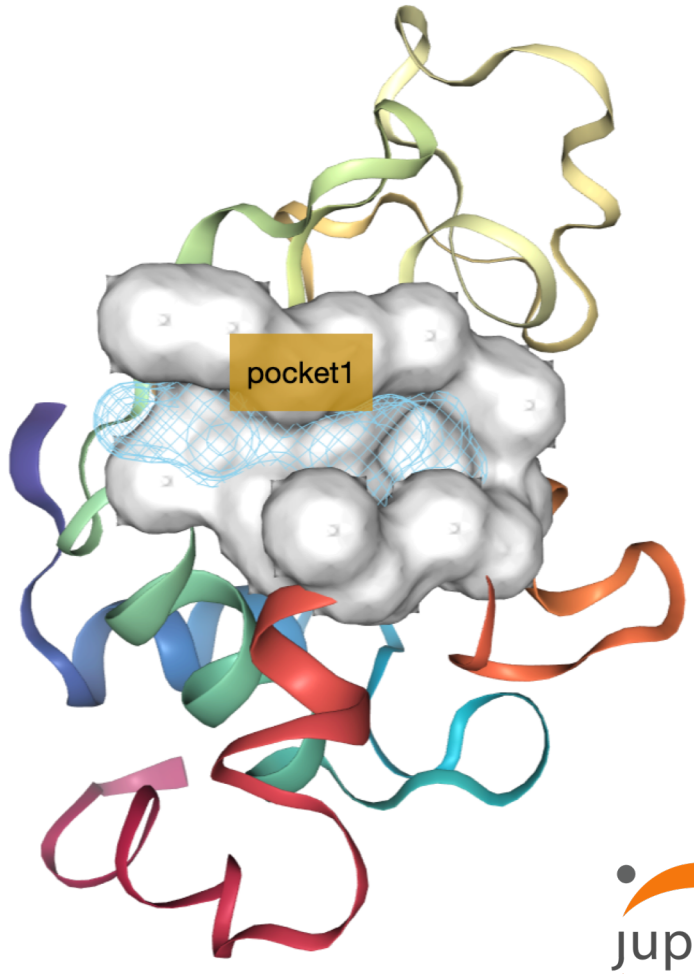
- fpocket_run from **biobb_vs.fpocket.fpocket_run**

In [ ]:
```python
# Finding protein pockets

# Create properties dict and inputs/outputs

out_pockets = "pockets.zip"
out_summary = "pockets.summary.json"

prop = {
    "min_radius": 3,
    "max_radius": 6,
    "num_spheres": 35,
    "sort_by": "druggability_score"

}

# Launch bb on REST API
token = launch_job(url = apiURL + 'launch/biobb_vs/fpocket_run',
                   config = prop,
                   input_pdb_path = downloaded_pdb,
                   output_pockets_zip = out_pockets,
                   output_summary = out_summary)
```

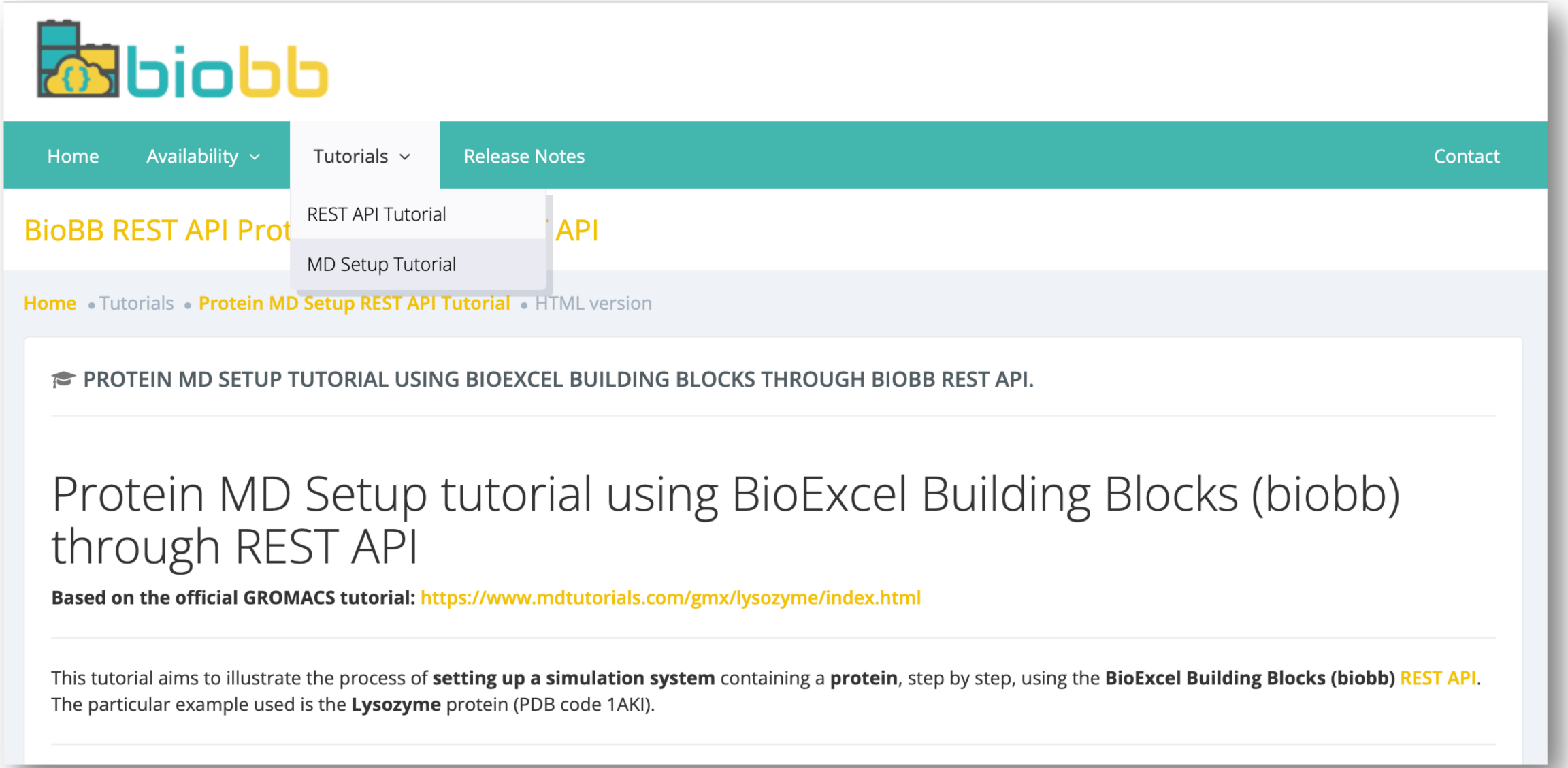
Python

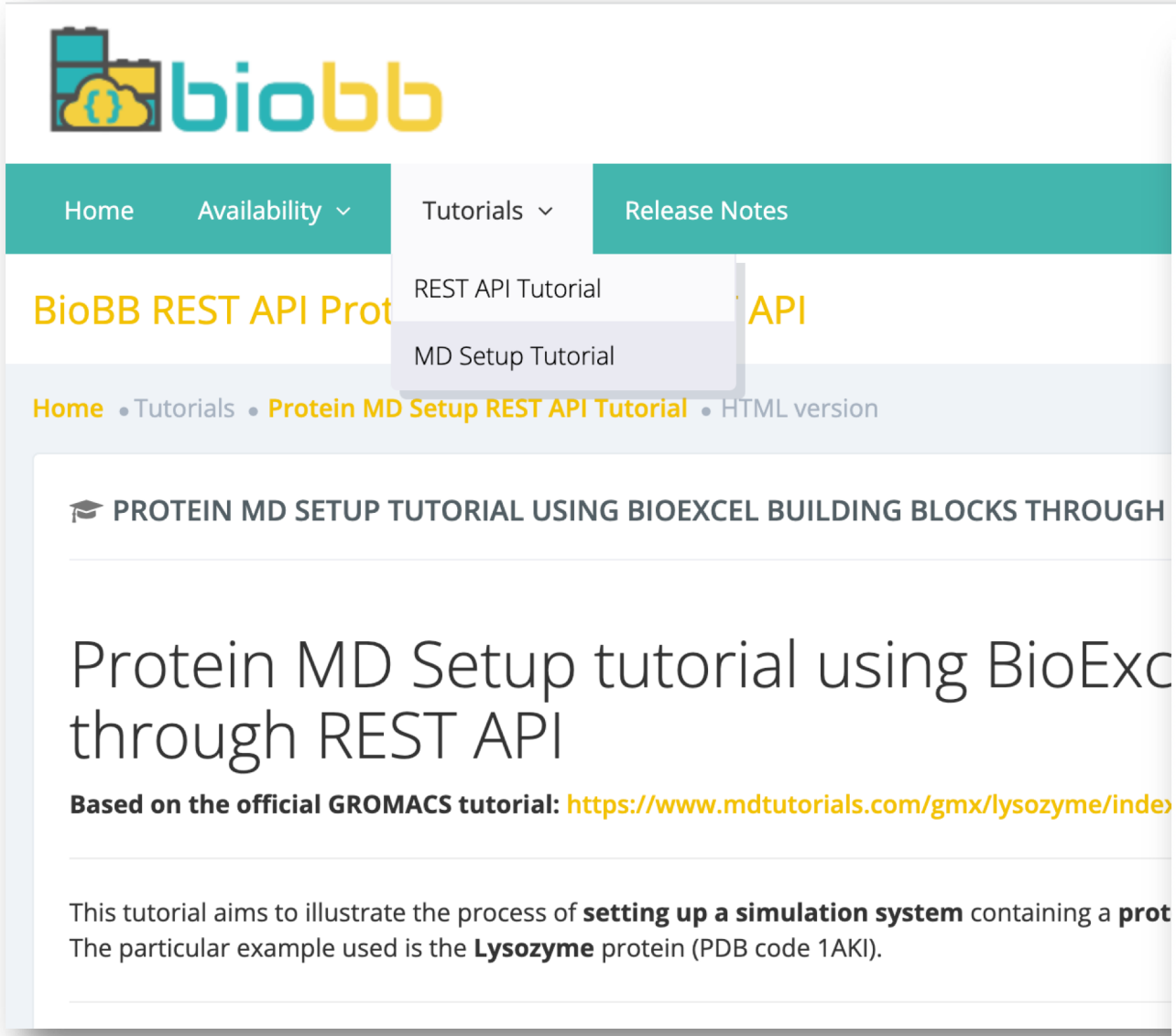
In [ ]:
```python
# Check job status
out_files = check_job(token, apiURL)
```

In [ ]:
```python
# Save generated file to disk
retrieve_data(out_files, apiURL)
```



jupyter


biobb
Home
Availability
Tutorials
Release Notes
Contact
REST API Tutorial
MD Setup Tutorial
BioBB REST API Prot API
Home • Tutorials • Protein MD Setup REST API Tutorial • HTML version
PROTEIN MD SETUP TUTORIAL USING BIOEXCEL BUILDING BLOCKS THROUGH BIOBB REST API.
Protein MD Setup tutorial using BioExcel Building Blocks (biobb) through REST API
Based on the official GROMACS tutorial: https://www.mdtutorials.com/gmx/lysozyme/index.html
This tutorial aims to illustrate the process of setting up a simulation system containing a protein, step by step, using the BioExcel Building Blocks (biobb) REST API.
The particular example used is the Lysozyme protein (PDB code 1AKI).



Pipeline steps

1. Input Parameters
2. Fetching PDB Structure
3. Fix Protein Structure
4. Create Protein System Topology
5. Create Solvent Box
6. Fill the Box with Water Molecules
7. Adding Ions
8. Energetically Minimize the System
9. Equilibrate the System (NVT)
10. Equilibrate the System (NPT)
11. Free Molecular Dynamics Simulation
12. Post-processing and Visualizing Resulting 3D Trajectory
13. Output Files
14. Questions & Comments

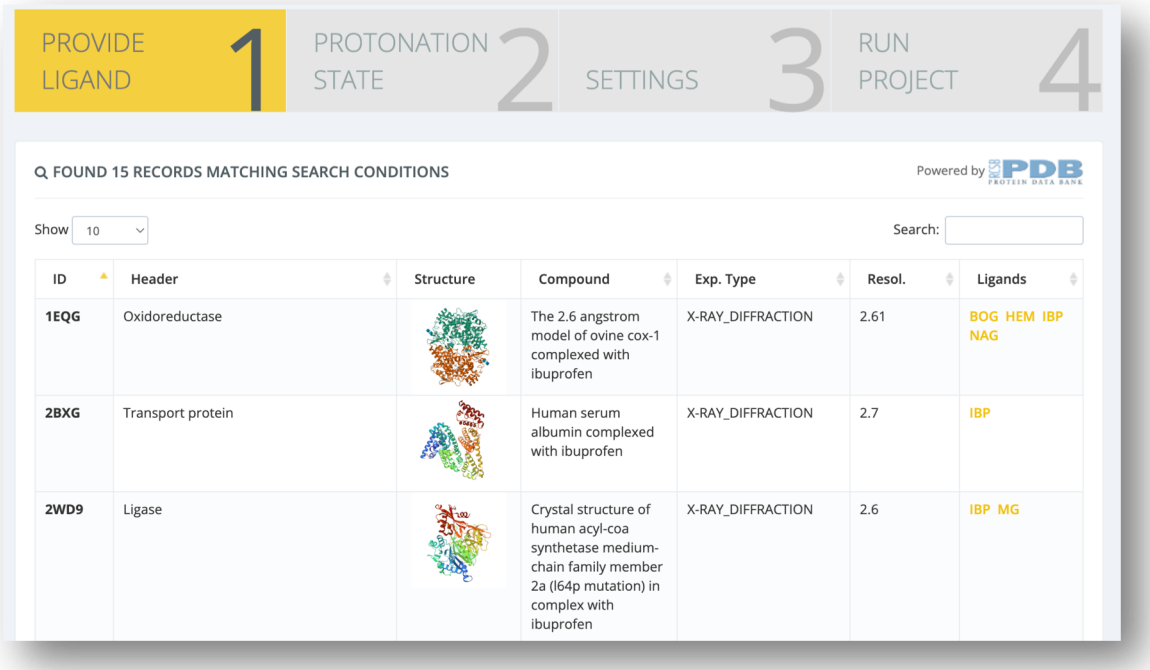
```yaml
name: biobb_MDsetupAPI_tutorial
channels:
  - conda-forge
  - bioconda
dependencies:
  - python
  - nb_conda_kernels
  - plotly
  - nglview
  - simpletraj
  - conda
```

Index

BioBB Wfs | Concept | Implementation | Usage

BioBB Workflows GUI

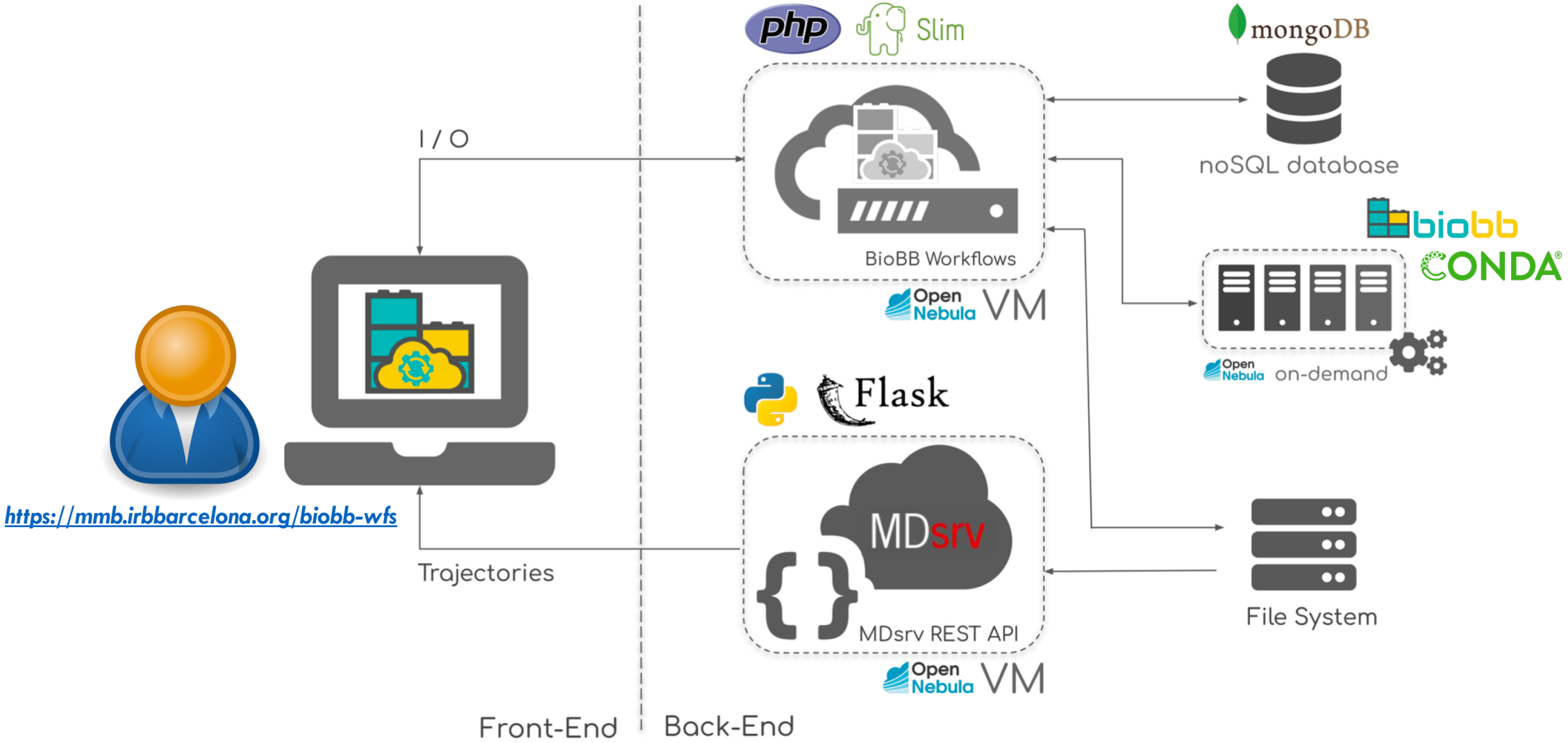
- **Web-based graphical access** to pre-configured **biomolecular simulation workflows**
- **Workflows** are run in the **provider's infrastructure** (with limitations)
  [but can also be downloaded to be run locally]
- **Personal workspace** with user's projects
- **No needs** of any **installation / deployment**
- Connection to external (HPC) **clusters**



Web-based GUI access

BioBB Wfs | Concept | **Implementation** | Usage

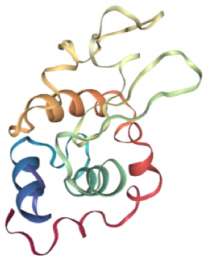
BioBB Workflows GUI



MDsrv: visual sharing and analysis of molecular dynamics simulations
Nucleic Acids Research, Volume 50, W1, 2022, W483–W489

GROMACS PROTEIN MD SETUP

2023.1



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 | Jupyter Notebook | CWL | Python | Galaxy

Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

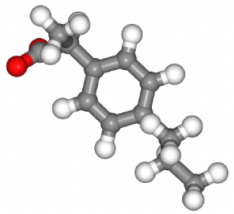
View tutorial | Open Github repository | Open documentation

gmx md protein

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

AUTOMATIC LIGAND PARAMETERIZATION

2023.1



This tutorial aims to illustrate the process of ligand parameterization for a small molecule, step by step, using the BioExcel Building Blocks library (biobb). The particular example used is the Ibuprofen small compound (3-letter code IBP, Drugbank code DB01050), a non-steroidal anti-inflammatory drug (NSAID) derived from propionic acid and it is considered the first of the propionics.

WorkflowHub | Jupyter Notebook | CWL | Python | Galaxy

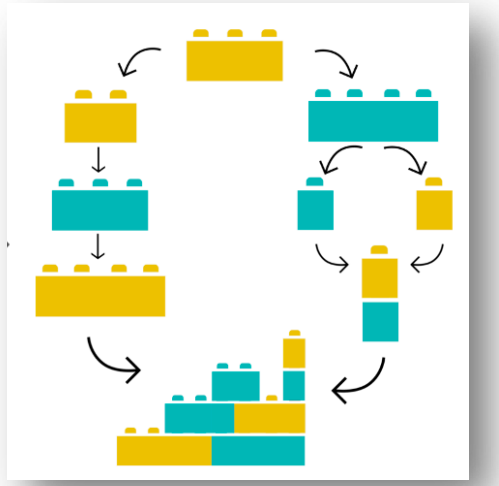
Launch | Jupyter Notebook * | Galaxy | BioBB Workflows

View tutorial | Open Github repository | Open documentation

gmx ligand

(*) Binder for biobb is a small installation and to promote fair use of our resources, one user is allowed to run only one notebook server at a time. Launching a new notebook server should stop the previous one. Users cannot see the notebooks run by other users, but please avoid entering secret data to the notebooks.

- ***MD setup (Protein / DNA)***
  - *Amber*
  - *Gromacs*
  - *Mutation(s)*
- ***Ligand parameterization***
- ***Protein-Ligand Docking***
- ***MD trajectory analyses***
- ***DNA helical parameters***



BioBB Wfs | Concept | Implementation | Usage

BioBB Workflows GUI

biobb

https://mmb.irbbar

Home | Workflows | Create project | Help

Welcome to BioExcel Building Blocks Workflows

Home

BioExcel Building Blocks Workflows

BioBB (BioExcel Building Blocks) packages are Python building blocks that create new layer of compatibility and interoperability over popular **bioinformatics tools**. In this site we provide the ability of running several workflows on top of the **BioBB's**.

Workflows can be run starting from a structure, a DNA sequence, a trajectory, a protein+ligand complex or a small molecule.

For more information about **BioExcel Building Blocks**, please visit the official website or the official repository.

This website is **free and open to all users** and there is no login requirement, though you can use the personal area to perform advanced actions.

BioExcel Building Blocks Workflows list

Home • Workflows

BioExcel Building Blocks Workflows

In this page, users can find a list of all available **BioExcel Building Blocks Workflows**. A list of **main steps** and a direct **launch** button are provided for each workflow.

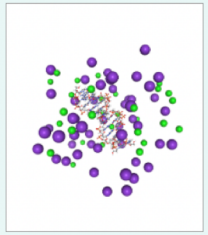
They can also be found in a GitHub repository in all their versions with all the **instructions** and **sample files** needed for **executing** them **at home** in the users' own premises.

Search by tags

Select or write tag(s)

ABC MD SETUP PIPELINE

Schema | Workflow example



This workflow provides a pipeline to setup DNA structures following the recommended guidelines by the Ascona B-DNA Consortium (ABC) members. It follows the work started with the NAFlex tool to offer a single, reproducible pipeline for structure preparation, ensuring reproducibility and coherence between all the members of the consortium.

Workflow main steps

Background information

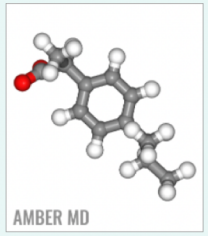
Launch ABC MD Setup workflow

amber md na

python cwl hpc

AUTOMATIC LIGAND PARAMETERIZATION FOR AMBER MD

Schema



This workflow performs an automatic ligand parameterization for a small molecule using GAFF force field, generating parameters compatible with the AMBER MD package.

Workflow main steps

Background information

Launch AMBER Ligand Parameterization workflow

amber ligand

python cwl

bioexcel

BioBB Workflows GUI – Usage

https://mmb.irbbarcelona.org/biobb-wfs/

Sign In | Sign Up

Home | Workflows | Create project | Help | Contact

From Structure
From DNA/RNA Sequence
From Trajectory
From Protein + Ligand
From Small Molecule

Welcome to BioEx Workflows

Home

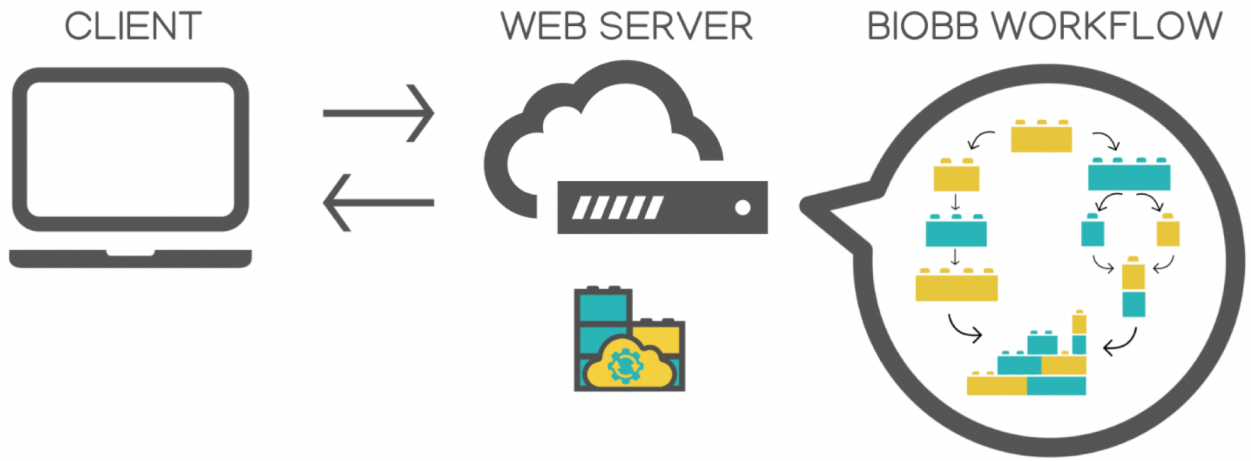
BioExcel Building

BioBB (BioExcel Building Python building blocks that create new layer of compatibility and interoperability over popular **bioinformatics tools**. In this site we provide the ability of running several workflows on top of the **BioBB's**.

Workflows can be run starting from a structure, a DNA sequence, a trajectory, a protein+ligand complex or a small molecule.

For more information about **BioExcel Building Blocks**, please visit the official website or the official repository.

This website is **free and open to all users** and there is no login requirement, though you can use the personal area to perform advanced actions.



Create new project

Home • Create project from structure

Create project from structure

In this category users can **custom**, **configure** and **execute** the following workflows related w
depends on the **structure checking parameters** and the **inputs** provided:

GROMACS Set Up | GROMACS Complex Set Up | AMBER Set Up | AMBER Com

PROVIDE STRUCTURE 1 | CHECK STRUCTURE 2

PROVIDE STRUCTURE

Perform a search on the **Protein Data Bank**.

Insert here the text you want to search in a protein structure.

pyruvate kinase

Search | Reset

Search results for **pyruvate kinase**

Home • Create project from structure • Search results for **pyruvate kinase**

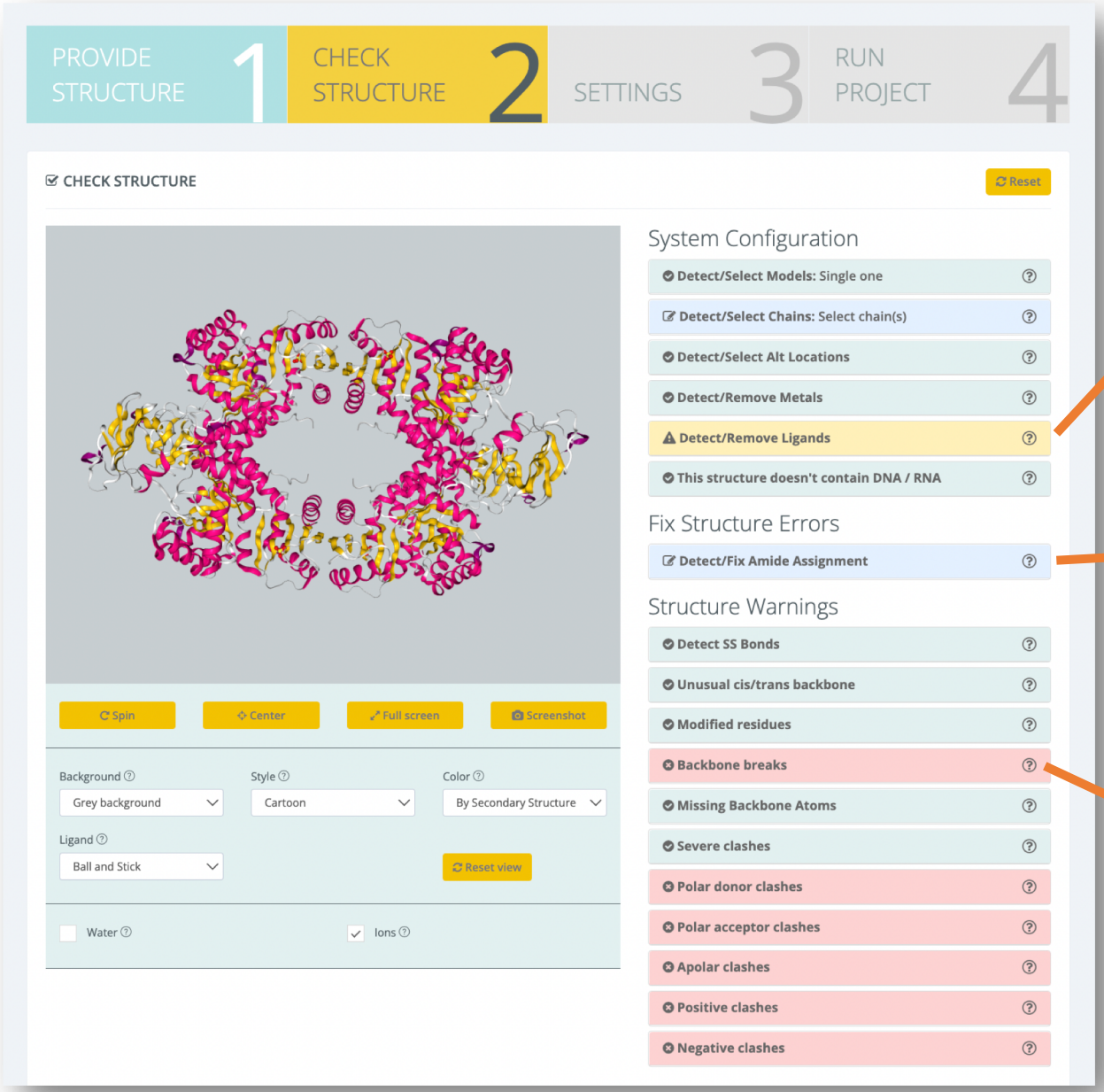
PROVIDE STRUCTURE 1 | CHECK STRUCTURE 2 | SETTINGS 3 | RUN PROJECT 4

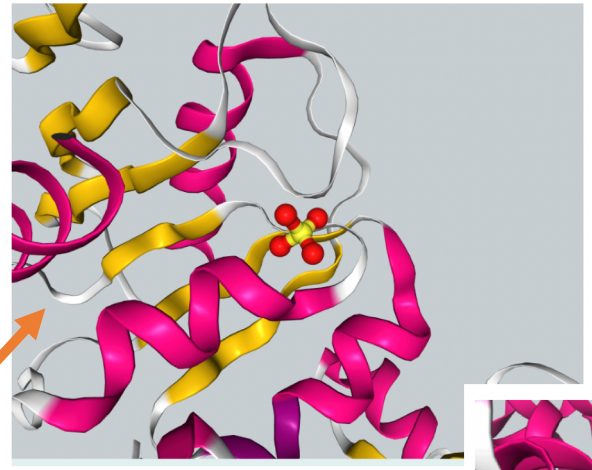
FOUND 150 RECORDS MATCHING SEARCH CONDITIONS

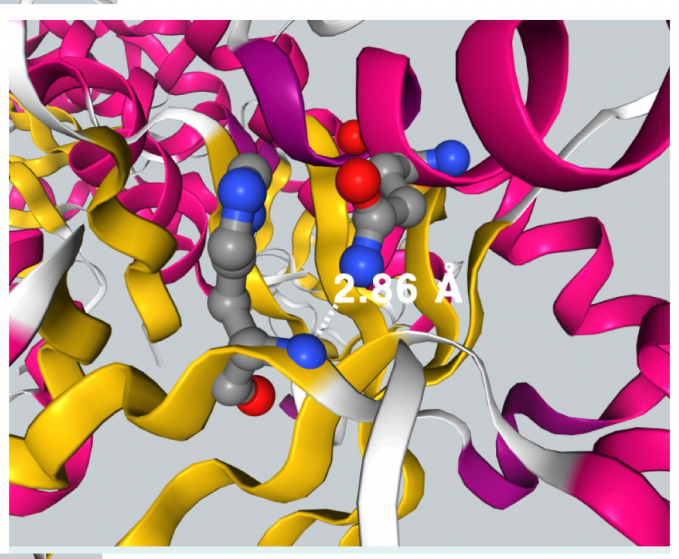
Powered by RCSB PDB PROTEIN DATA BANK

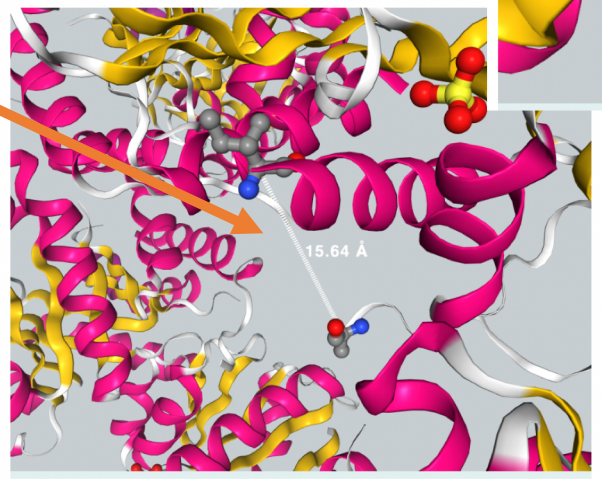
Show 10 Search:

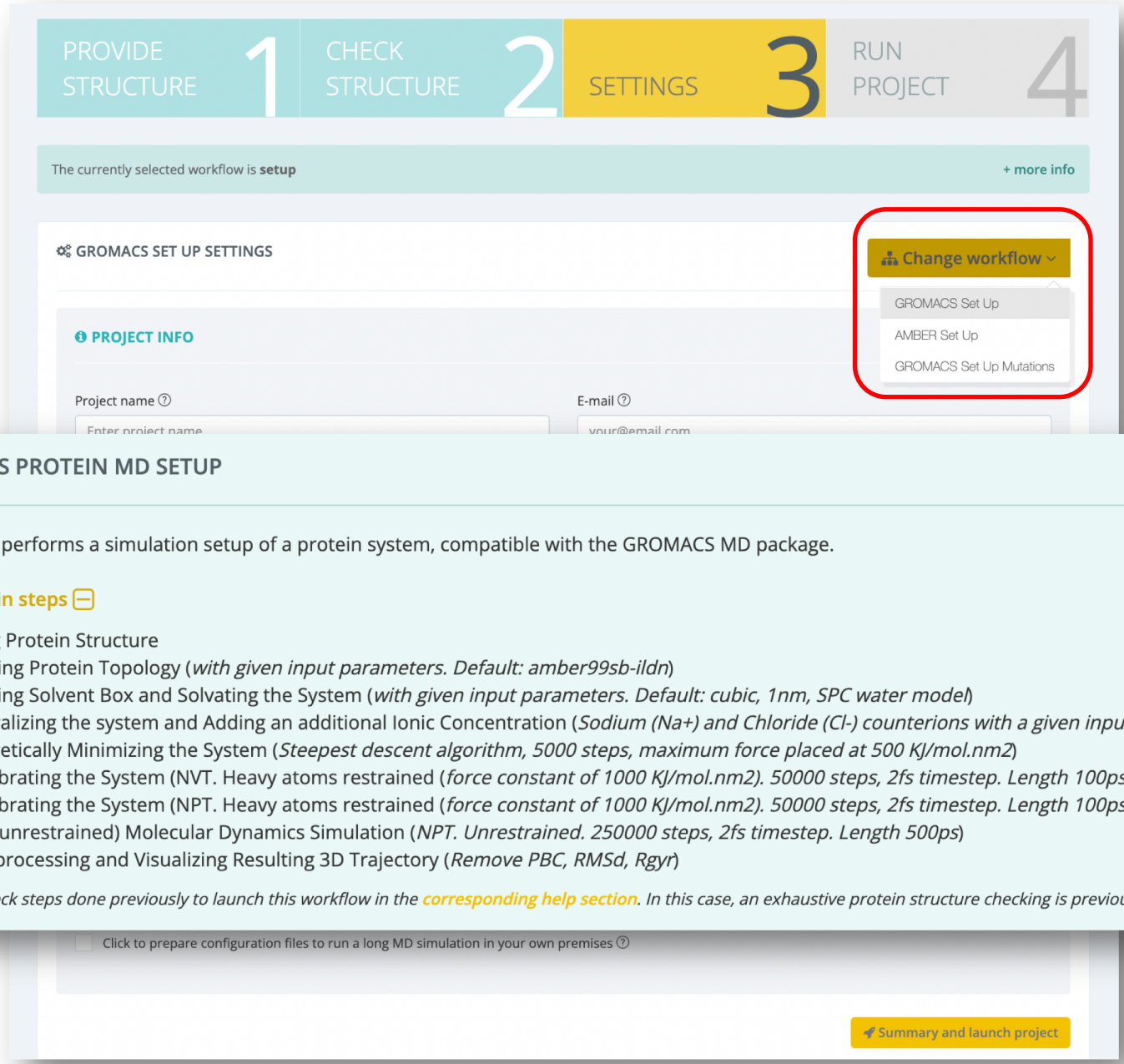
ID	Header	Structure	Compound	Exp. Type	Resol.	Ligands
1A3W	Transferase		**pyruvate kinase** from saccharomyces cerevisiae complexed with fbp, pg, mn2+ and k+	X-RAY_DIFFRACTION	3	FBP K MN PGA
1A3X	Transferase		**pyruvate kinase** from saccharomyces cerevisiae complexed with pg, mn2+ and k+	X-RAY_DIFFRACTION	3	K MN PGA
1A49	Transferase		Bis mg-atp-k-oxalate complex of **pyruvate kinase**	X-RAY_DIFFRACTION	2.1	ATP K MG OXL
1A5U	Transferase		**pyruvate kinase** complex with bis mg-atp-na-oxalate	X-RAY_DIFFRACTION	2.35	ATP MG NA OXL
1AQF	Transferase		**pyruvate kinase** from rabbit muscle with mg, k, and l-phospholactate	X-RAY_DIFFRACTION	2.7	K MG PEQ


PROVIDE STRUCTURE 1
CHECK STRUCTURE 2
SETTINGS 3
RUN PROJECT 4
CHECK STRUCTURE
Reset
Spin
Center
Full screen
Screenshot
Background
Grey background
Style
Cartoon
Color
By Secondary Structure
Ligand
Ball and Stick
Reset view
Water
Ions
System Configuration
Detect/Select Models: Single one
Detect/Select Chains: Select chain(s)
Detect/Select Alt Locations
Detect/Remove Metals
Detect/Remove Ligands
This structure doesn't contain DNA / RNA
Fix Structure Errors
Detect/Fix Amide Assignment
Structure Warnings
Detect SS Bonds
Unusual cis/trans backbone
Modified residues
Backbone breaks
Missing Backbone Atoms
Severe clashes
Polar donor clashes
Polar acceptor clashes
Apolar clashes
Positive clashes
Negative clashes




2.86 Å


15.64 Å



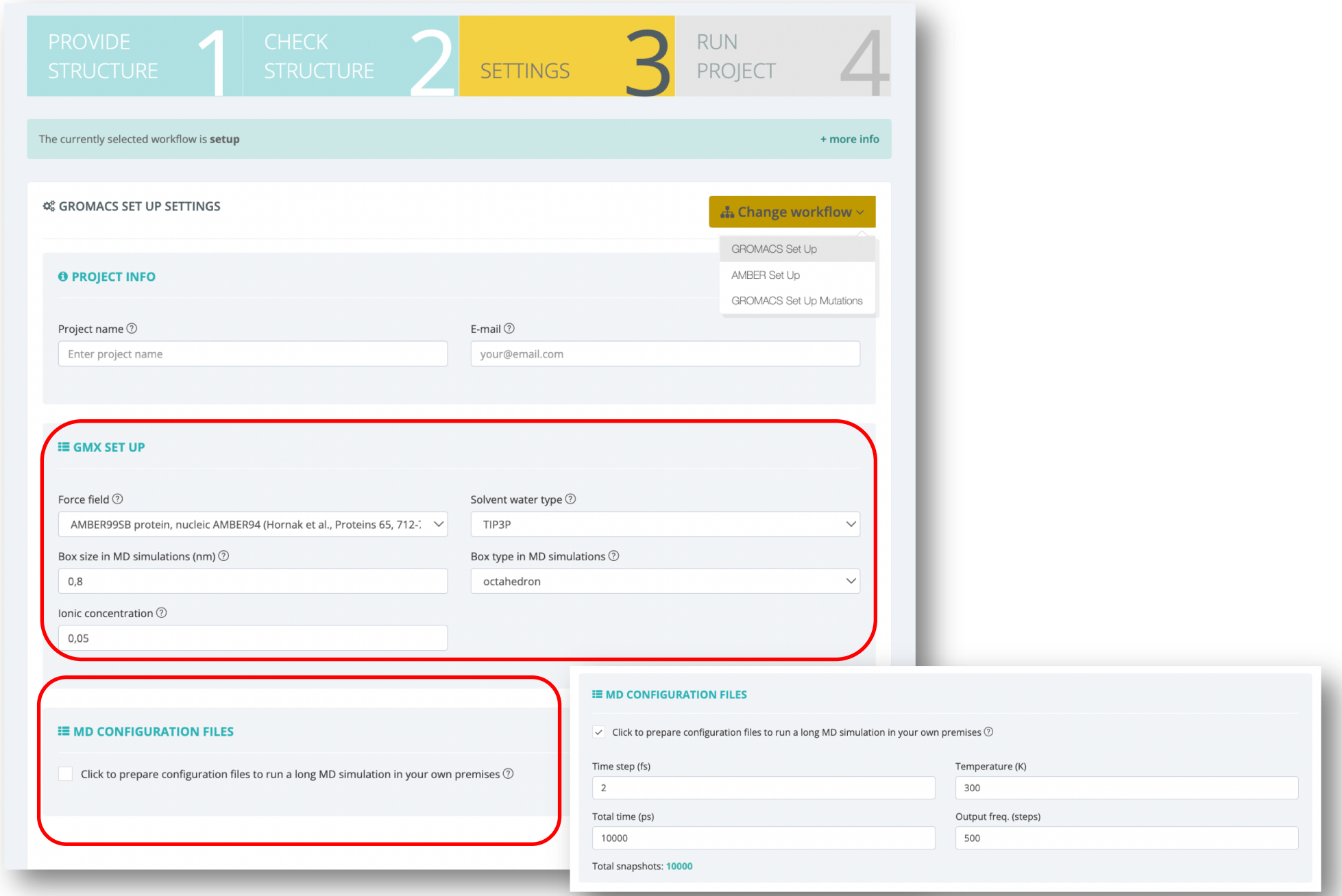
GROMACS PROTEIN MD SETUP

This workflow performs a simulation setup of a protein system, compatible with the GROMACS MD package.

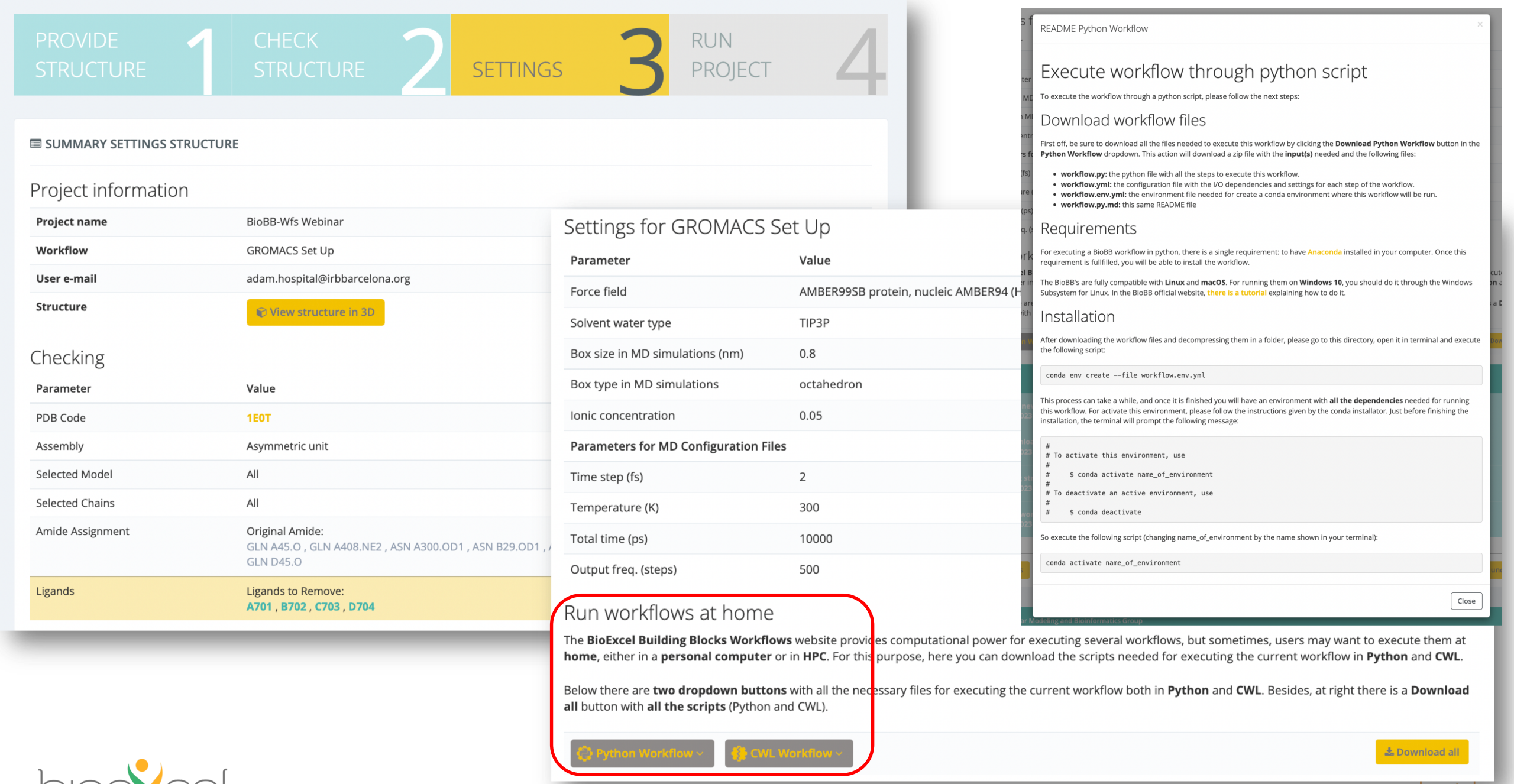
Workflow main steps

1. Fixing Protein Structure
2. Creating Protein Topology (*with given input parameters. Default: amber99sb-ildn*)
3. Creating Solvent Box and Solvating the System (*with given input parameters. Default: cubic, 1nm, SPC water model*)
4. Neutralizing the system and Adding an additional Ionic Concentration (*Sodium (Na+) and Chloride (Cl-) counterions with a given input additional ionic concentration. Default: 50mM.*)
5. Energetically Minimizing the System (*Steepest descent algorithm, 5000 steps, maximum force placed at 500 KJ/mol.nm2*)
6. Equilibrating the System (NVT. Heavy atoms restrained (*force constant of 1000 KJ/mol.nm2). 50000 steps, 2fs timestep. Length 100ps*)
7. Equilibrating the System (NPT. Heavy atoms restrained (*force constant of 1000 KJ/mol.nm2). 50000 steps, 2fs timestep. Length 100ps*)
8. Free (unrestrained) Molecular Dynamics Simulation (*NPT. Unrestrained. 250000 steps, 2fs timestep. Length 500ps*)
9. Post-processing and Visualizing Resulting 3D Trajectory (*Remove PBC, RMSd, Rgyr*)

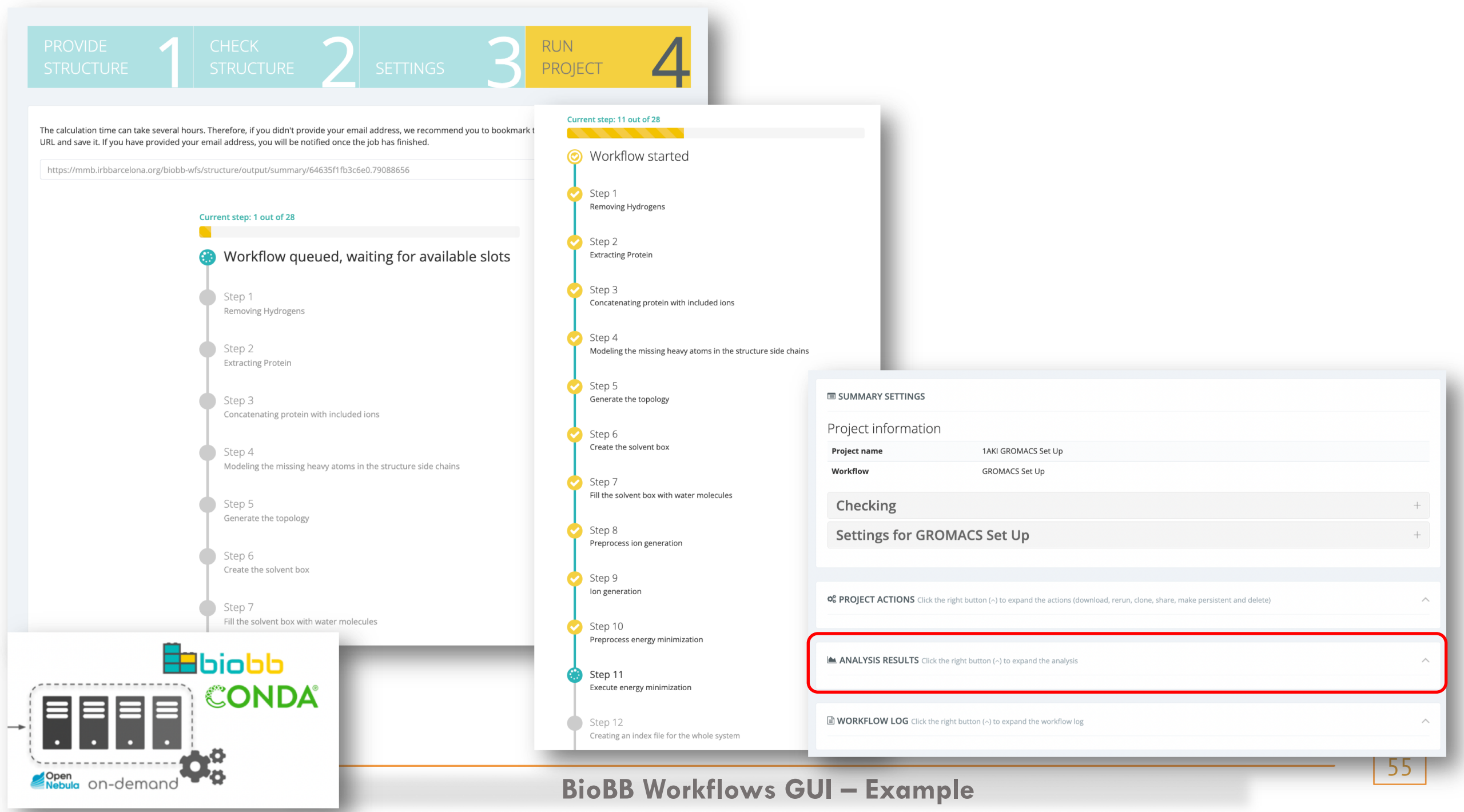
Note: Please check steps done previously to launch this workflow in the corresponding help section. In this case, an exhaustive protein structure checking is previously run by BioBB Workflows.



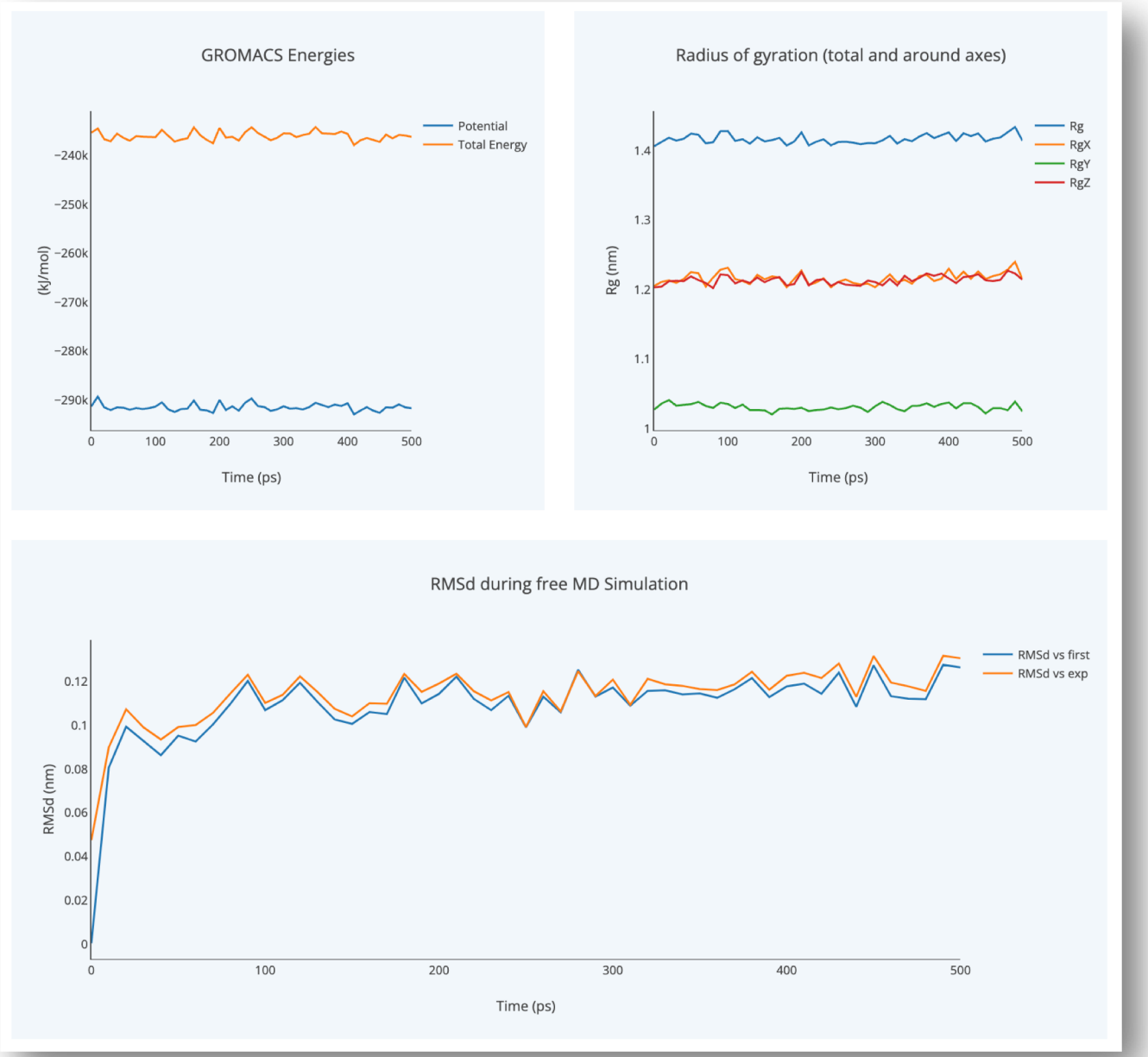
BioBB Workflows GUI – Example



BioBB Workflows GUI – Example

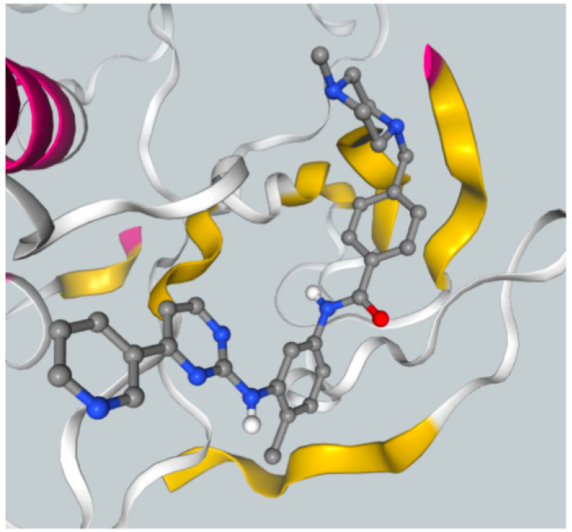


BioBB Workflows GUI – Example



BioBB Workflows GUI – Example

Complex final result



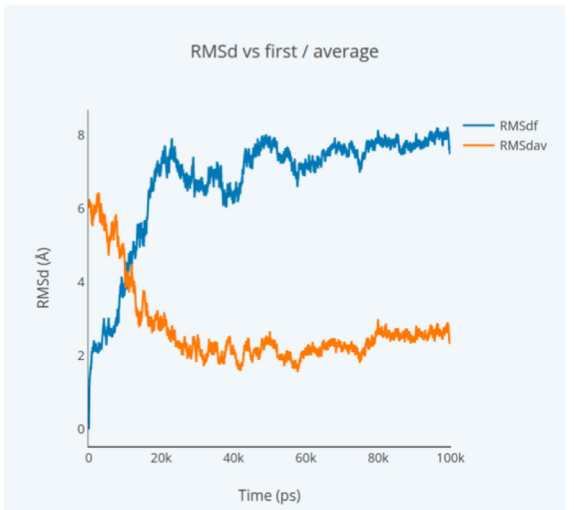
Show 10 Search:

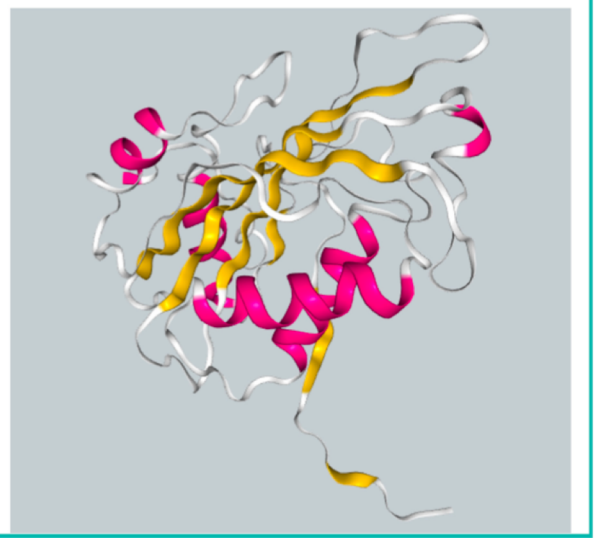
	Mode	Affinity	Dist from best mode RMSd l.b.	RMSd u.b.
☐	1	-6.8	0.000	0.000
☐	2	-6.2	1.051	1.414
☐	3	-5.9	2.262	11.782
☐	4	-5.3	1.831	2.493
☐	5	-4.5	2.698	3.221
☑	6	-4.3	2.449	11.142
☐	7	-4.2	2.551	3.519
☐	8	-4.0	2.856	3.486

Showing 1 to 8 of 8 records

RMSd First / Average vs Trajectory

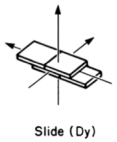
Click into a point in the left plot to see the corresponding frame in the 3D representation at right.



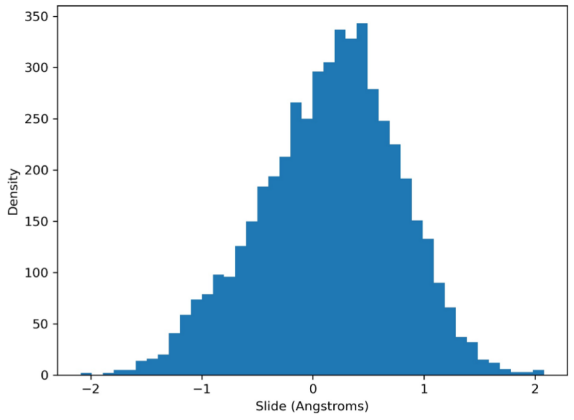


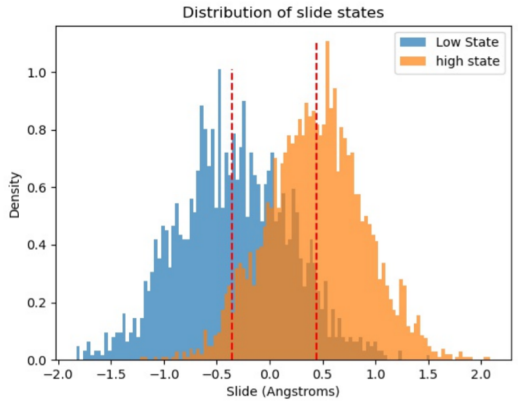
BIMODALITY Click the right button (^) to expand the analysis

Select parameter



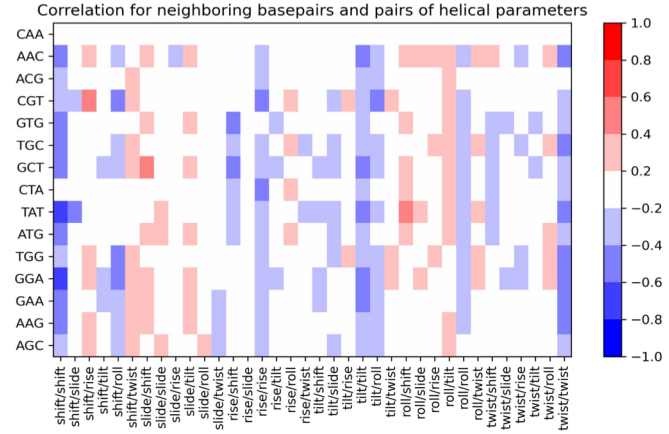
Select pair from sequence

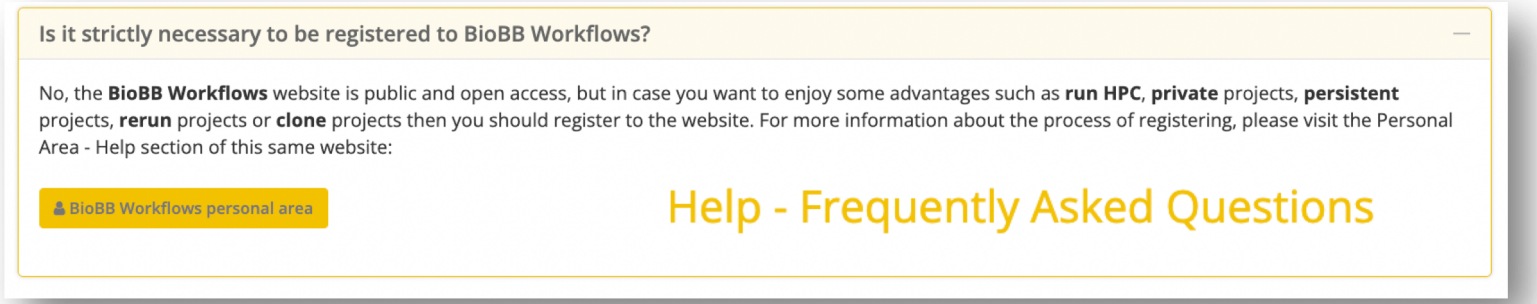


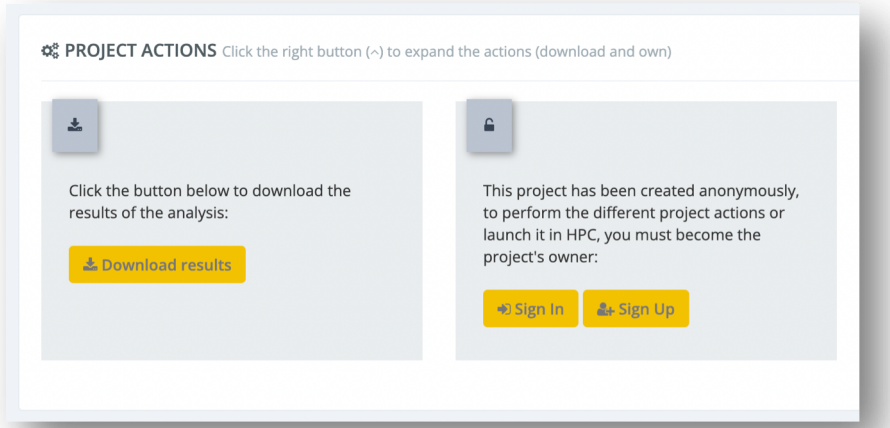


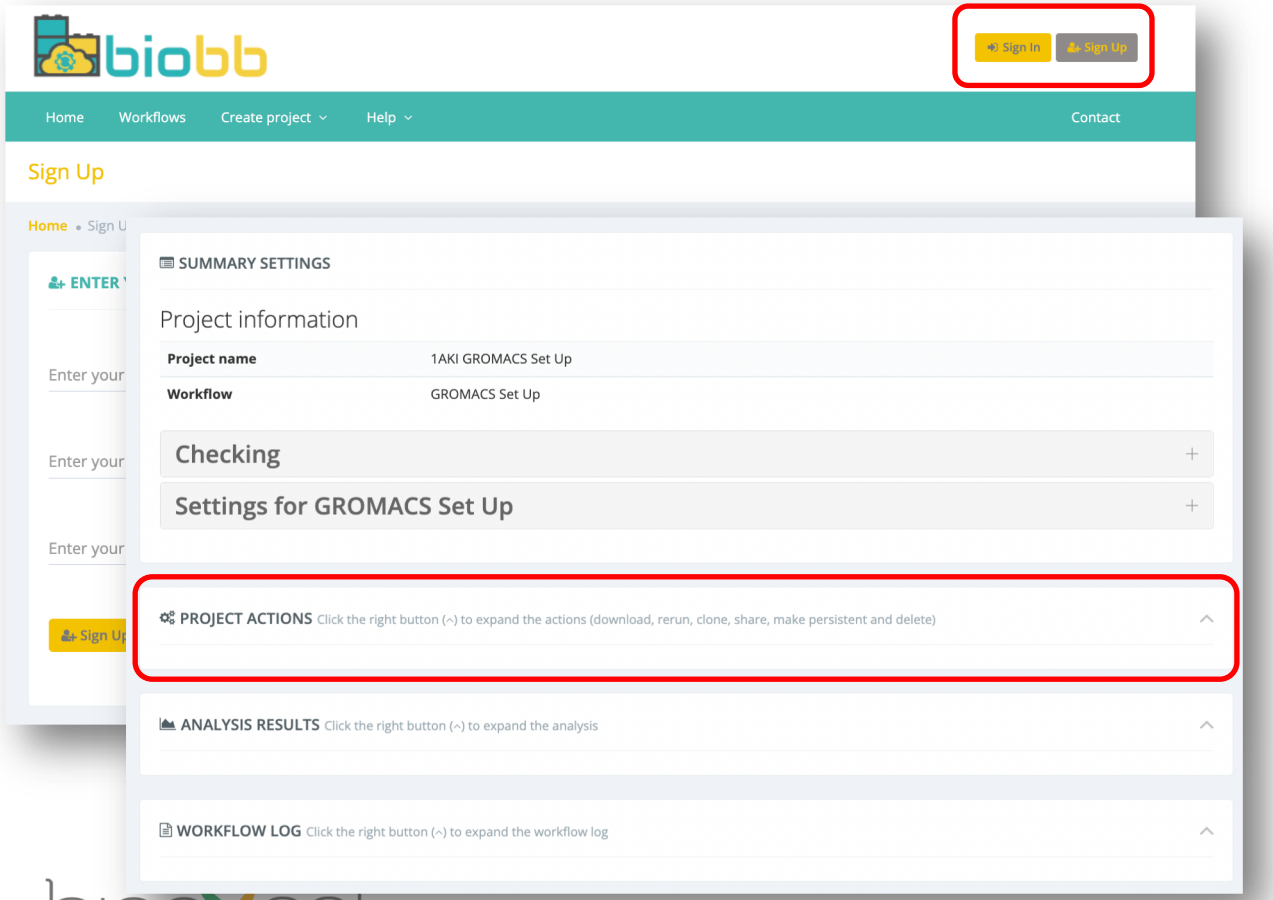
CORRELATIONS Click the right button (^) to expand the analysis

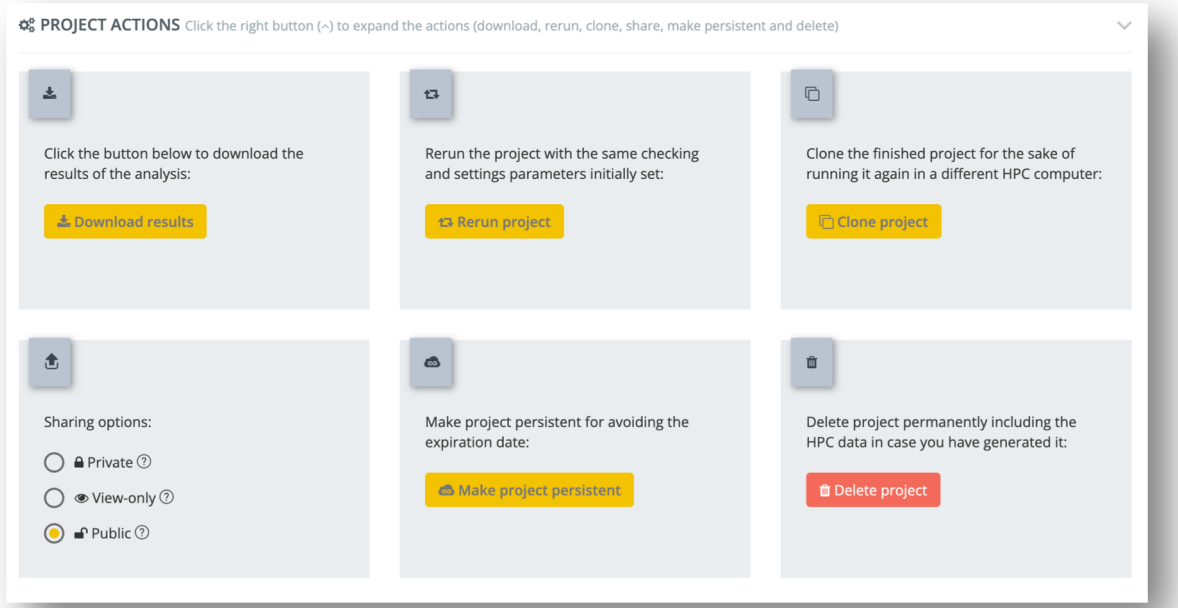
Select correlation type

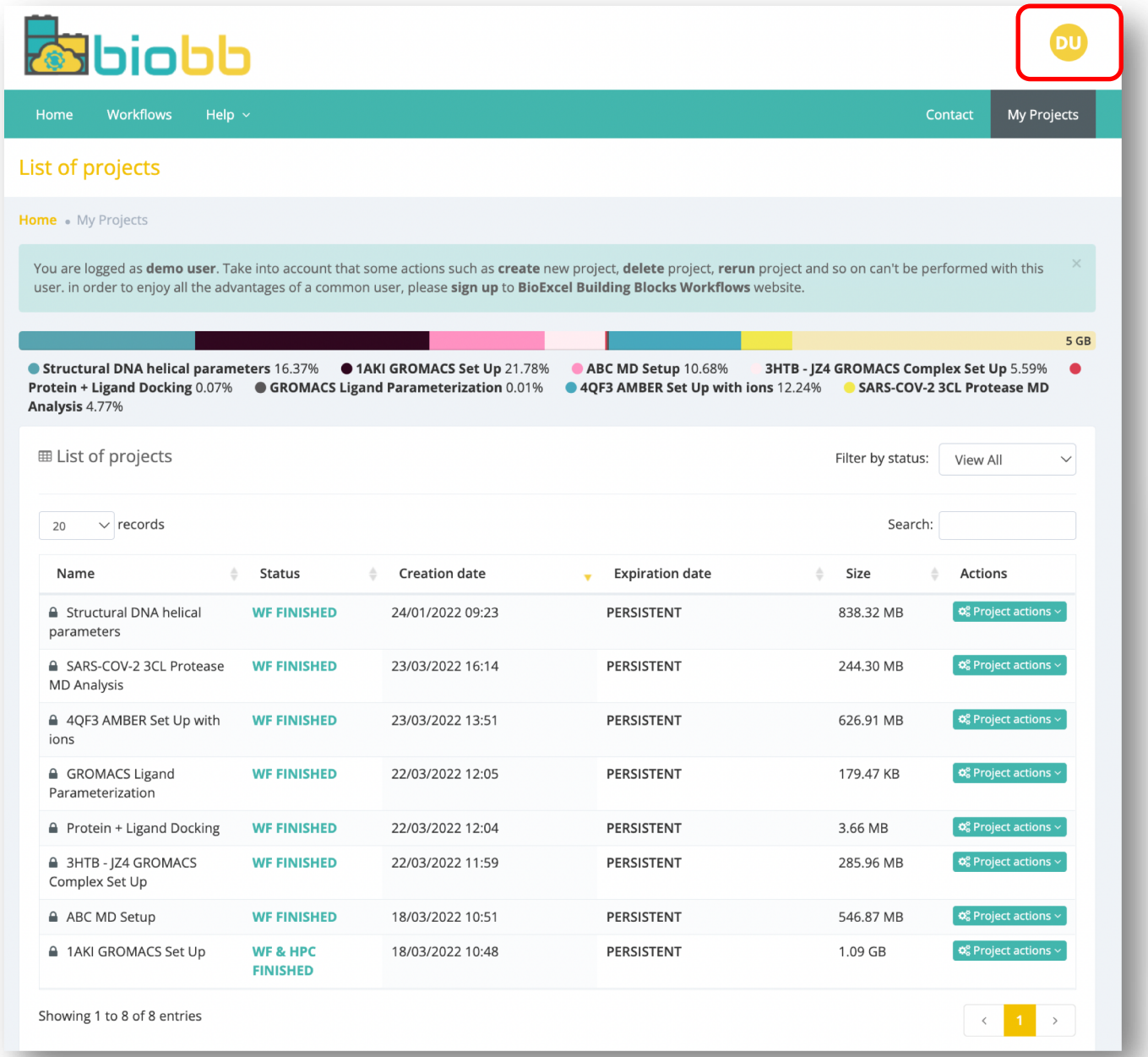


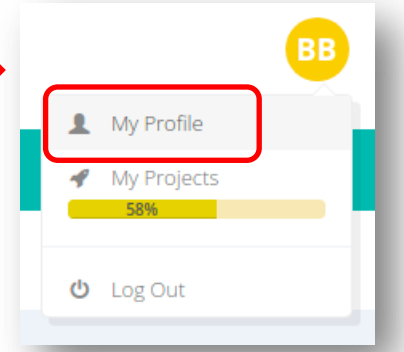

Is it strictly necessary to be registered to BioBB Workflows?
No, the BioBB Workflows website is public and open access, but in case you want to enjoy some advantages such as run HPC, private projects, persistent projects, rerun projects or clone projects then you should register to the website. For more information about the process of registering, please visit the Personal Area - Help section of this same website:
BioBB Workflows personal area
Help - Frequently Asked Questions

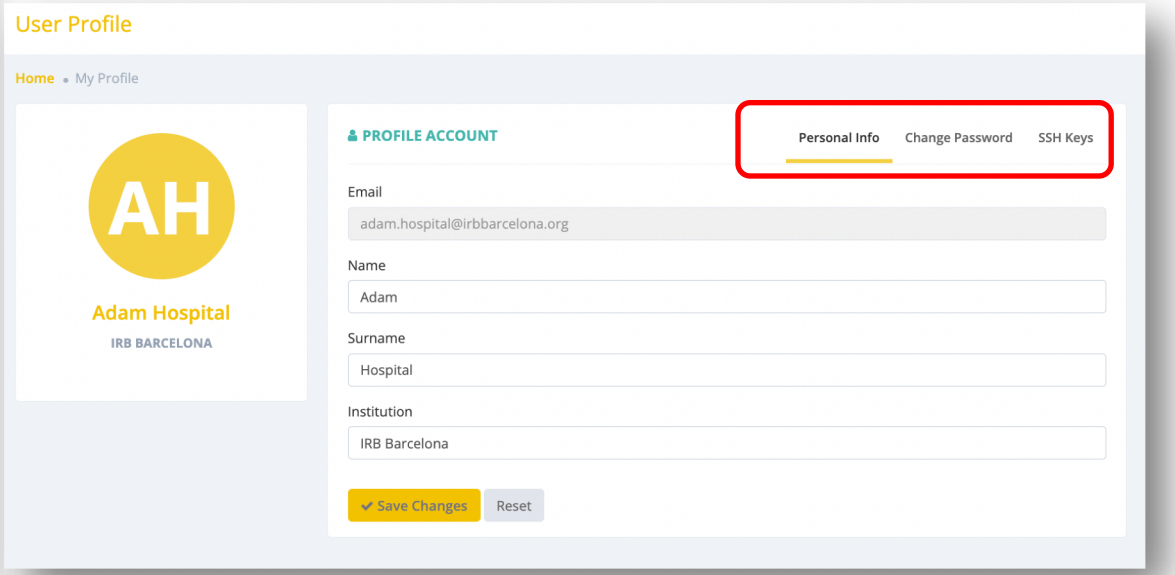

PROJECT ACTIONS Click the right button (^) to expand the actions (download and own)
Click the button below to download the results of the analysis:
Download results
This project has been created anonymously, to perform the different project actions or launch it in HPC, you must become the project's owner:
Sign In
Sign Up

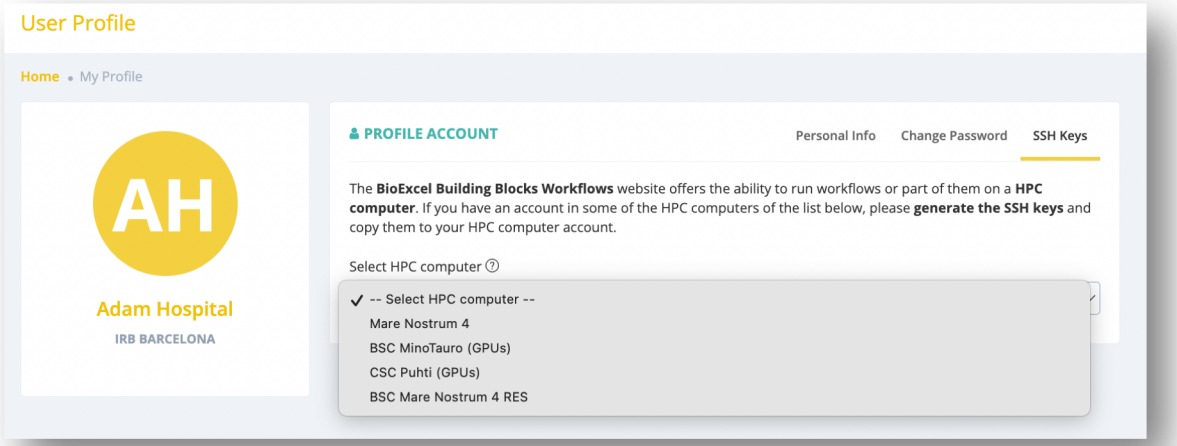

biobb
Sign In
Sign Up
Home
Workflows
Create project
Help
Contact
Sign Up
SUMMARY SETTINGS
Project information
Project name
1AKI GROMACS Set Up
Workflow
GROMACS Set Up
Checking
Settings for GROMACS Set Up
PROJECT ACTIONS Click the right button (^) to expand the actions (download, rerun, clone, share, make persistent and delete)
ANALYSIS RESULTS Click the right button (^) to expand the analysis
WORKFLOW LOG Click the right button (^) to expand the workflow log

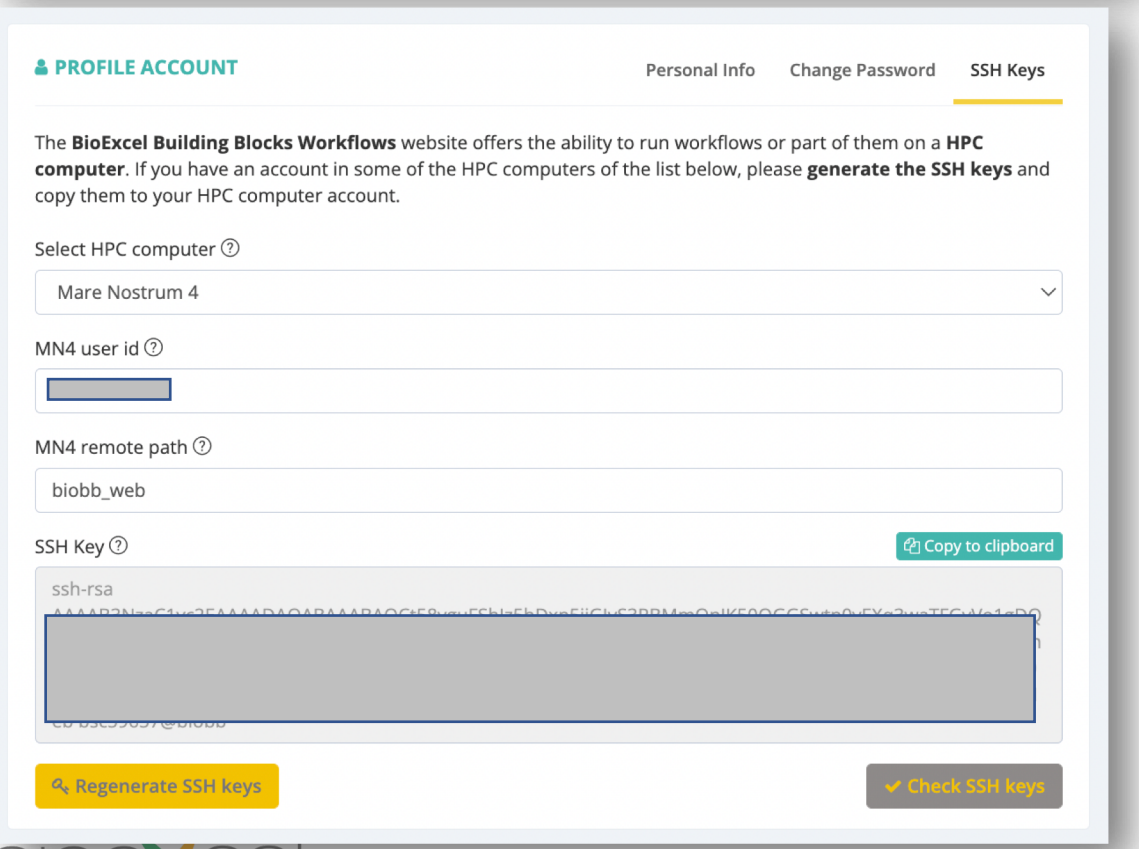

PROJECT ACTIONS Click the right button (^) to expand the actions (download, rerun, clone, share, make persistent and delete)
Click the button below to download the results of the analysis:
Download results
Rerun the project with the same checking and settings parameters initially set:
Rerun project
Clone the finished project for the sake of running it again in a different HPC computer:
Clone project
Sharing options:
Private
View-only
Public
Make project persistent for avoiding the expiration date:
Make project persistent
Delete project permanently including the HPC data in case you have generated it:
Delete project

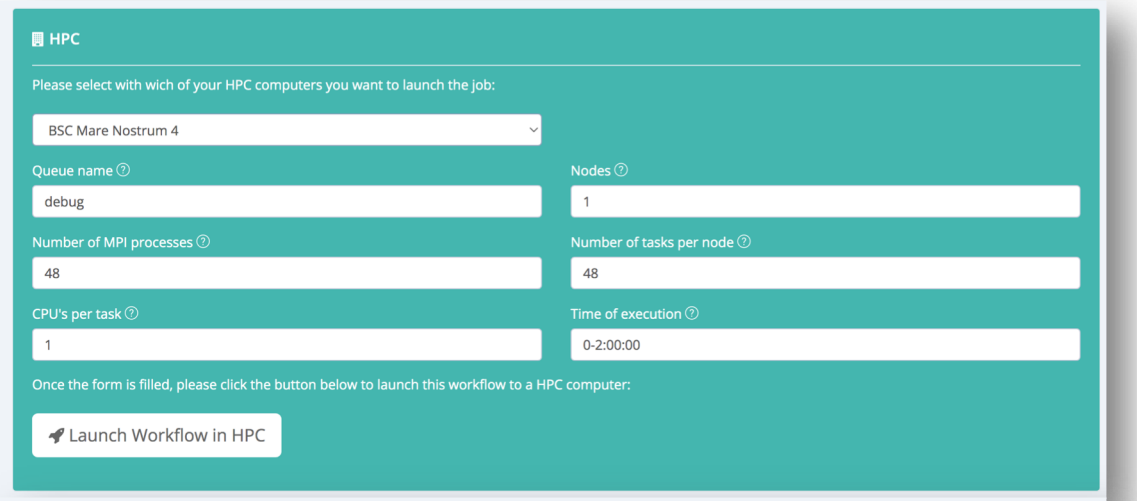

biobb
DU
Home
Workflows
Help
Contact
My Projects
List of projects
Home • My Projects
You are logged as demo user. Take into account that some actions such as create new project, delete project, rerun project and so on can't be performed with this user. in order to enjoy all the advantages of a common user, please sign up to BioExcel Building Blocks Workflows website.
5 GB
Structural DNA helical parameters 16.37%
1AKI GROMACS Set Up 21.78%
ABC MD Setup 10.68%
3HTB - JZ4 GROMACS Complex Set Up 5.59%
Protein + Ligand Docking 0.07%
GROMACS Ligand Parameterization 0.01%
4QF3 AMBER Set Up with ions 12.24%
SARS-COV-2 3CL Protease MD Analysis 4.77%
List of projects
Filter by status:
View All
20 records
Search:
Name
Status
Creation date
Expiration date
Size
Actions
Structural DNA helical parameters
WF FINISHED
24/01/2022 09:23
PERSISTENT
838.32 MB
Project actions
SARS-COV-2 3CL Protease MD Analysis
WF FINISHED
23/03/2022 16:14
PERSISTENT
244.30 MB
Project actions
4QF3 AMBER Set Up with ions
WF FINISHED
23/03/2022 13:51
PERSISTENT
626.91 MB
Project actions
GROMACS Ligand Parameterization
WF FINISHED
22/03/2022 12:05
PERSISTENT
179.47 KB
Project actions
Protein + Ligand Docking
WF FINISHED
22/03/2022 12:04
PERSISTENT
3.66 MB
Project actions
3HTB - JZ4 GROMACS Complex Set Up
WF FINISHED
22/03/2022 11:59
PERSISTENT
285.96 MB
Project actions
ABC MD Setup
WF FINISHED
18/03/2022 10:51
PERSISTENT
546.87 MB
Project actions
1AKI GROMACS Set Up
WF & HPC FINISHED
18/03/2022 10:48
PERSISTENT
1.09 GB
Project actions
Showing 1 to 8 of 8 entries
1

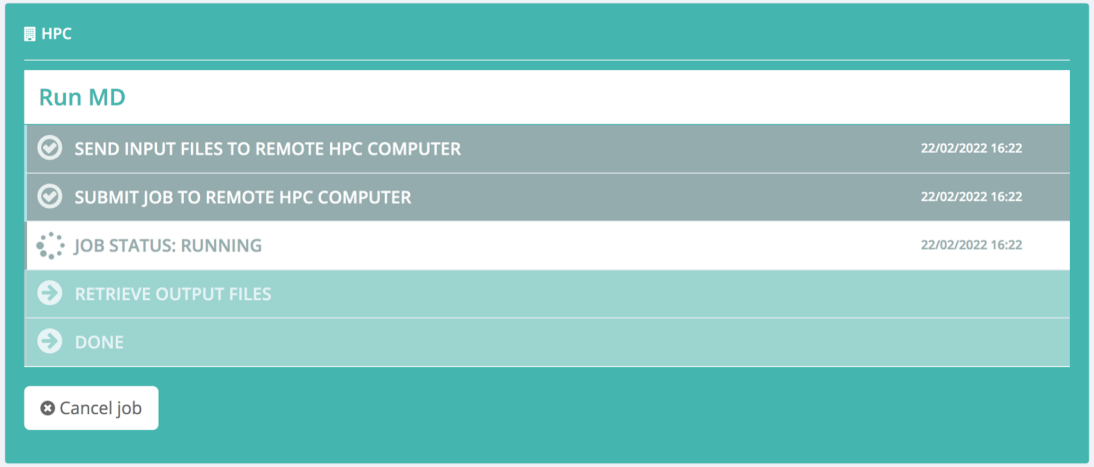

BB
My Profile
My Projects
58%
Log Out

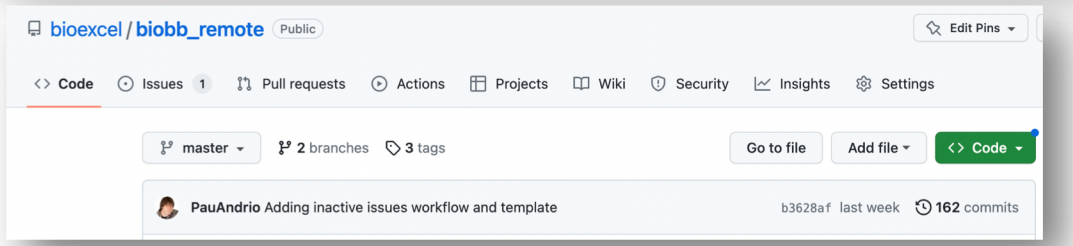

User Profile
Home • My Profile
AH
Adam Hospital
IRB BARCELONA
PROFILE ACCOUNT
Personal Info
Change Password
SSH Keys
Email
adam.hospital@irbbarcelona.org
Name
Adam
Surname
Hospital
Institution
IRB Barcelona
Save Changes
Reset


User Profile
Home • My Profile
AH
Adam Hospital
IRB BARCELONA
PROFILE ACCOUNT
Personal Info
Change Password
SSH Keys
The BioExcel Building Blocks Workflows website offers the ability to run workflows or part of them on a HPC computer. If you have an account in some of the HPC computers of the list below, please generate the SSH keys and copy them to your HPC computer account.
Select HPC computer
-- Select HPC computer --
Mare Nostrum 4
BSC MinoTauro (GPUs)
CSC Puhti (GPUs)
BSC Mare Nostrum 4 RES


PROFILE ACCOUNT
Personal Info
Change Password
SSH Keys
The BioExcel Building Blocks Workflows website offers the ability to run workflows or part of them on a HPC computer. If you have an account in some of the HPC computers of the list below, please generate the SSH keys and copy them to your HPC computer account.
Select HPC computer
Mare Nostrum 4
MN4 user id
MN4 remote path
biobb_web
SSH Key
Copy to clipboard
ssh-rsa
Regenerate SSH keys
Check SSH keys


HPC
Please select with wich of your HPC computers you want to launch the job:
BSC Mare Nostrum 4
Queue name
debug
Nodes
1
Number of MPI processes
48
Number of tasks per node
48
CPU's per task
1
Time of execution
0-2:00:00
Once the form is filled, please click the button below to launch this workflow to a HPC computer:
Launch Workflow in HPC


HPC
Run MD
SEND INPUT FILES TO REMOTE HPC COMPUTER
22/02/2022 16:22
SUBMIT JOB TO REMOTE HPC COMPUTER
22/02/2022 16:22
JOB STATUS: RUNNING
22/02/2022 16:22
RETRIEVE OUTPUT FILES
DONE
Cancel job


bioexcel / biobb_remote
Public
Edit Pins
Code
Issues 1
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
master
2 branches
3 tags
Go to file
Add file
Code
PauAndrio Adding inactive issues workflow and template
b3628af last week
162 commits

Index

Summary

BioBB REST-API

- **Remote** and **programmatic access** to biomolecular simulation processes
- Processes are run in the **provider's infrastructure** (with limitations)
- **No needs** of any **installation** / **deployment**

https://mmb.irbbarcelona.org/biobb-api/

BioBB Workflows GUI

- **Web-based graphical access** to pre-configured **biomolecular simulation workflows**
- **Workflows** are run in the **provider's infrastructure** (with limitations)
  [but can also be downloaded to be run locally]
- **Personal workspace** with user's projects
- **No needs** of any **installation** / **deployment**
- Connection to external **clusters**

https://mmb.irbbarcelona.org/biobb-wfs/

Acknowledgments

Prof. Modesto Orozco

Prof. Josep Lluís Gelpí

RESEARCH GROUP
Molecular Modelling and Bioinformatics
Mechanisms of Disease

IRB BARCELONA
INSTITUTE FOR RESEARCH IN BIOMEDICINE

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Genís Bayarri

Pau Andrio

Questions / Suggestions

JOURNAL ARTICLE

BioExcel Building Blocks REST API (BioBB REST API), programmatic access to interoperable biomolecular simulation tools Get access

Genís Bayarri, Pau Andrio, Adam Hospital, Modesto Orozco, Josep Lluís Gelpí

Bioinformatics, Volume 38, Issue 12, 15 June 2022, Pages 3302–3303,
https://doi.org/10.1093/bioinformatics/btac316
Published: 11 May 2022 **Article history**

JOURNAL ARTICLE

BioExcel Building Blocks Workflows (BioBB-Wfs), an integrated web-based platform for biomolecular simulations

Genís Bayarri, Pau Andrio, Adam Hospital, Modesto Orozco, Josep Lluís Gelpí

Nucleic Acids Research, Volume 50, Issue W1, 5 July 2022, Pages W99–W107,
https://doi.org/10.1093/nar/gkac380
Published: 26 May 2022 **Article history**

- What **functionality/workflow** would you like to see **integrated** in these tools? [*Add*]
- I would consider **using** them only if …… [*Change*]
- I find them **useful / not useful** because …… [*Feedback*]

adam.hospital@irbbarcelona.org