

---

**biobb**<sub>*i*</sub>*oDocumentation*

***Release 3.0.0***

**Bioexcel Project**

**Aug 07, 2020**



---

## Contents

---

<b>1</b>	<b>Contents</b>	<b>3</b>
<b>2</b>	<b>Indices and tables</b>	<b>17</b>
<b>3</b>	<b>Github repository.</b>	<b>19</b>
	<b>Python Module Index</b>	<b>21</b>
	<b>Index</b>	<b>23</b>







## 1.1 biobb\_io

### 1.1.1 Introduction

Biobb\_io is the Biobb module collection to fetch data to be consumed by the rest of the Biobb building blocks. Biobb (BioExcel building blocks) packages are Python building blocks that create new layer of compatibility and interoperability over popular bioinformatics tools. The latest documentation of this package can be found in our readthedocs site: [latest API documentation](#).

### 1.1.2 Version

v3.0.0 2020.1

### 1.1.3 Installation

Using PIP:

- Installation:

```
pip install "biobb_io==3.0.0"
```

- Usage: [Python API documentation](#)

Using ANACONDA:

- Installation:

```
conda install -c bioconda "biobb_io==3.0.0"
```

- Usage: With conda installation BioBBs can be used with the [Python API documentation](#) and the [Command Line documentation](#)

Using DOCKER:

- Installation:

```
docker pull quay.io/biocontainers/biobb_io:3.0.0--py_0
```

- Usage:

```
docker run quay.io/biocontainers/biobb_io:3.0.0--py_0 <command>
```

The command list and specification can be found at the [Command Line documentation](#).

Using SINGULARITY:

**MacOS users:** it's strongly recommended to avoid Singularity and use **Docker** as containerization system.

- Installation:

```
singularity pull --name biobb_io.sif shub://bioexcel/biobb_io
```

- Usage:

```
singularity exec biobb_io.sif <command>
```

The command list and specification can be found at the [Command Line documentation](#).

## 1.1.4 Copyright & Licensing

This software has been developed in the MMB group at the BSC & IRB for the European BioExcel, funded by the European Commission (EU H2020 823830, EU H2020 675728).

- (c) 2015-2020 [Barcelona Supercomputing Center](#)
- (c) 2015-2020 [Institute for Research in Biomedicine](#)

Licensed under the [Apache License 2.0](#), see the file LICENSE for details.





## 1.2 biobb\_io

### 1.2.1 configuration package

#### Submodules

##### api.ligand module

Module containing the Ligand class and the command line interface.

**class** `api.ligand.Ligand` (*output\_pdb\_path*, *properties=None*, *\*\*kwargs*)

Bases: `object`

Wrapper class for the PDB REST API. This class is a wrapper for the [MMB PDB mirror](#).

##### Parameters

- **output\_pdb\_path** (*str*) – Path to the output PDB ligand file. File type: output. [Sample file](#). Accepted formats: pdb.
- **properties** (*dic*) –
  - **ligand\_code** (*str*) - (“12D”) RSCB PDB ligand code.
  - **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
  - **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

**launch** () → int

Writes the PDB file content of the first `ligand_code` to `output_pdb_path`.

`api.ligand.main` ()

Command line interface.

##### api.pdb module

Module containing the Pdb class and the command line interface.

**class** `api.pdb.Pdb` (*output\_pdb\_path*, *properties=None*, *\*\*kwargs*)

Bases: `object`

Wrapper class for the PDB REST API. This class is a wrapper for the [PDB download page](#).

##### Parameters

- **output\_pdb\_path** (*str*) – Path to the output PDB file. File type: output. [Sample file](#). Accepted formats: pdb.
- **properties** (*dic*) –
  - **pdb\_code** (*str*) - (“1ubq”) RSCB PDB code.
  - **filter** (*str*) - ([“ATOM”, “MODEL”, “ENDMDL”]) Array of groups to be kept. If value is None or False no filter will be applied. All the possible values are defined in the official PDB specification (<http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>)
  - **url** (*str*) - (“<https://files.rcsb.org/download/>”) URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API (“<http://mmb.irbbarcelona.org/api/pdb/>”).
  - **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
  - **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

**launch** () → int

Writes the PDB file content of the first `pdb_code` to `output_pdb_path`.

`api.pdb.main` ()

Command line interface.

## api.pdb\_variants module

MmbPdbVariants Module

**class** `api.pdb_variants.MmbPdbVariants` (`output_mutations_list_txt`, `properties=None`,  
`**kwargs`)

Bases: `object`

Wrapper class for the MMB group UNIPROT REST API. This class is a wrapper for the [UNIPROT](#) mirror of the [MMB PDB mirror](#).

### Parameters

- **output\_mutations\_list\_txt** (`str`) – Path to the TXT file containing an ASCII comma separated values of the mutations. File type: output. Accepted formats: txt.
- **properties** (`dic`) –
  - **pdb\_code** (`str`): (“2vgb”) RSCB PDB four letter code. ie: “2ki5”.
  - **url** (`str`) - (“<https://files.rcsb.org/download/>”) URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API (“<http://mmb.irbbarcelona.org/api/pdb/>”).
  - **remove\_tmp** (`bool`) - (True) [WF property] Remove temporal files.
  - **restart** (`bool`) - (False) [WF property] Do not execute if output files exist.

**launch** () → int

Writes the variants of the selected `pdb_code` to `output_mutations_list_txt`

`api.pdb_variants.main` ()

## api.pdb\_cluster\_zip module

MmbPdbClusterZip Module

**class** `api.pdb_cluster_zip.MmbPdbClusterZip` (`output_pdb_zip_path`, `properties=None`,  
`**kwargs`)

Bases: `object`

Wrapper class for the [MMB PDB mirror](#).

### Parameters

- **output\_pdb\_zip\_path** (`str`) – Path to the ZIP or PDB file containing the output PDB files. File type: output. [Sample file](#). Accepted formats: pdb, zip.
- **properties** (`dic`) –
  - **pdb\_code** (`str`) - (“2vgb”) RSCB PDB code. ie: “2VGB”
  - **filter** (`str`) - ([“ATOM”, “MODEL”, “ENDMDL”]) Array of groups to be kept. If value is None or False no filter will be applied. All the possible values are defined in the official PDB specification (<http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>)
  - **cluster** (`str`) - (90) Cluster number for the `biobb_io.api.MmbPdb.get_pdb_cluster_zip` () method.

- **url** (*str*) - (“https://files.rcsb.org/download/”) URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API (“http://mmb.irbbarcelona.org/api/pdb”).
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

**launch** () → int

Writes each PDB file content of each `pdb_code` in the cluster to a `pdb_file` then creates a `zip_file` output `put_pdb_zip_path`.

`api.pdb_cluster_zip.main()`

## api.drugbank module

Module containing the Drugbank class and the command line interface.

**class** `api.drugbank.Drugbank` (*output\_sdf\_path*, *properties=None*, *\*\*kwargs*)

Bases: `object`

Download a component in SDF format from the [Drugbank](#).

### Parameters

- **output\_sdf\_path** (*str*) – Path to the output SDF component file. File type: output. [Sample file](#). Accepted formats: `sdf`.
- **properties** (*dic*) –
  - **drugbank\_id** (*str*) - (“DB00530”) Drugbank component id.
  - **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
  - **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

**launch** () → int

Writes the SDF content of the first `drugbank_id` to `output_sdf_path`.

`api.drugbank.main()`

Command line interface.

## 1.3 BioBB IO Command Line Help

Generic usage:

```
biobb_command [-h] [-c CONFIG] --in_file <in_file> --out_file <out_file>
```

### 1.3.1 ligand

This class is a wrapper for the MMB PDB mirror (<http://mmb.irbbarcelona.org/api/>)

#### Get help

```
ligand -h
```

```
usage: ligand.py [-h] [-c CONFIG] -o OUTPUT_PDB_PATH

Wrapper for the PDB ('http://www.rcsb.org/pdb/home/home.do') mirror of the MMB group,
↳ REST API ('http://mmbl.irbbarcelona.org/api/') for additional help in the
↳ commandline usage please check ('https://biobb-io.readthedocs.io/en/latest/command_
↳ line.html')

optional arguments:
  -h, --help            show this help message and exit
  -c CONFIG, --config CONFIG
                        This file can be a YAML file, JSON file or JSON string

required arguments:
  -o OUTPUT_PDB_PATH, --output_pdb_path OUTPUT_PDB_PATH
                        Path to the output PDB ligand file.
```

## I / O Arguments

Syntax: input\_argument (datatype) : Definition

Config input / output arguments for this building block:

- **output\_pdb\_path** (*str*): Path to the output PDB ligand file. File type: output. [Sample file](#). Accepted formats: pdb.

## Config

Syntax: input\_parameter (datatype) - (default\_value) Definition

Config parameters for this building block:

- **ligand\_code** (*str*) - ("12D") RSCB PDB ligand code. ie: "12D"
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

## Default config

```
ligand -o 12D.pdb
```

## YAML

### File config

```
properties:
  ligand_code: 12D
```

```
ligand -c conf.yml -o 12D.pdb
```

## JSON

### String config

```
pdb -c '{"ligand_code":"12D"}' -o 12D.pdb
```

### File config

```
{
  "properties": {
    "ligand_code":"12D"
  }
}
```

```
ligand -c conf.json -o 12D.pdb
```

## 1.3.2 pdb

Download PDB files from RCSB PDB API

### Get help

```
pdb -h
```

```
usage: pdb.py [-h] [-c CONFIG] -o OUTPUT_PDB_PATH
```

Wrapper **for** the PDB (<http://www.rcsb.org/pdb/home/home.do>) mirror of the MMB group,   
 ↳ REST API (<http://mmb.irbbarcelona.org/api/>) **for** additional help **in** the   
 ↳ commandline usage please check ([https://biobb-io.readthedocs.io/en/latest/command\\_line.html](https://biobb-io.readthedocs.io/en/latest/command_line.html))

optional arguments:

```
-h, --help          show this help message and exit
-c CONFIG, --config CONFIG
                    This file can be a YAML file, JSON file or JSON string
```

required arguments:

```
-o OUTPUT_PDB_PATH, --output_pdb_path OUTPUT_PDB_PATH
                    Output file name
```

### I / O Arguments

Syntax: input\_argument (datatype) : Definition

Config input / output arguments for this building block:

- **output\_pdb\_path** (*str*): Path to the output PDB file. File type: output. Sample file. Accepted formats: pdb.

## Config

Syntax: `input_parameter (datatype) - (default_value) Definition`

Config parameters for this building block:

- **pdb\_code** (*str*) - ('1ubq') RSCB PDB code. ie: "2VGB"
- **filter** (*str*) - (["ATOM", "MODEL", "ENDMDL"]) Array of groups to be kept. If value is None or False no filter will be applied. All the possible values are defined in the official PDB specification (<http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>)
- **url** (*str*) - ("https://files.rcsb.org/download/") URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API ("http://mmb.irbbarcelona.org/api/pdb")
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

## Default config

```
pdb -o laki.pdb
```

## YAML

### File config

```
properties:  
  pdb_code: laki  
  filter: ["ATOM"]
```

```
pdb -c conf.yml -o laki.pdb
```

## JSON

### String config

```
pdb -c '{"pdb_code":"laki", "filter":["ATOM"]}' -o laki.pdb
```

### File config

```
{  
  "properties": {  
    "pdb_code": "laki",  
    "filter": ["ATOM"]  
  }  
}
```

```
pdb -c conf.json -o laki.pdb
```

### 1.3.3 pdb\_variants

Retrieve variants from UNIPROT mapped to the selected PDB

#### Get help

```
pdb_variants -h
```

```
usage: pdb_variants.py [-h] [-c CONFIG] -o OUTPUT_MUTATIONS_LIST_TXT

Wrapper for the PDB Variants (http://www.rcsb.org/pdb/home/home.do) mirror of the MMB
↳group REST API (http://mmb.irbbarcelona.org/api/) for additional help in the
↳commandline usage please check ('https://biobb-io.readthedocs.io/en/latest/command\_line.html')

optional arguments:
  -h, --help            show this help message and exit
  -c CONFIG, --config CONFIG
                        This file can be a YAML file, JSON file or JSON string

required arguments:
  -o OUTPUT_MUTATIONS_LIST_TXT, --output_mutations_list_txt OUTPUT_MUTATIONS_LIST_TXT
                        Output variants list text file name
```

#### I / O Arguments

Syntax: `input_argument (datatype) : Definition`

Config input / output arguments for this building block:

- **output\_mutations\_list\_txt** (*str*): Path to the TXT file containing an ASCII comma separated values of the mutations. File type: output. Accepted formats: txt.

#### Config

Syntax: `input_parameter (datatype) - (default_value) Definition`

Config parameters for this building block:

- **pdb\_code** (*str*): (“2vgb”) RSCB PDB four letter code. ie: “2ki5”.
- **url** (*str*) - (“https://files.rcsb.org/download/”) URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API (“http://mmb.irbbarcelona.org/api/pdb/”).
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

#### Default config

```
pdb_variants -o mutations.txt
```

## YAML

### File config

```
properties:
  pdb_code: 2src
```

```
pdb_variants -c conf.yml -o mutations.txt
```

## JSON

### String config

```
pdb_variants -c '{"pdb_code":"2src"}' -o mutations.txt
```

### File config

```
{
  "properties": {
    "pdb_code": "2src"
  }
}
```

```
pdb_variants -c conf.json -o mutations.txt
```

---

### 1.3.4 pdb\_cluster\_zip

Download the selected similarity cluster of the selected PDB

#### Get help

```
pdb_cluster_zip -h
```

```
usage: pdb_cluster_zip.py [-h] [-c CONFIG] -o OUTPUT_PDB_ZIP_PATH
```

```
Wrapper for the PDB Cluster (http://www.rcsb.org/pdb/home/home.do) mirror of the MMB_
↳group REST API (http://mmb.irbbarcelona.org/api/) for additional help in the_
↳commandline usage please check ('https://biobb-io.readthedocs.io/en/latest/command\_
↳line.html')
```

optional arguments:

```
-h, --help          show this help message and exit
-c CONFIG, --config CONFIG
                    This file can be a YAML file, JSON file or JSON string
```

required arguments:

(continues on next page)



(continued from previous page)

```
-o OUTPUT_PDB_ZIP_PATH, --output_pdb_zip_path OUTPUT_PDB_ZIP_PATH
    Output ZIP file name
```

## I / O Arguments

Syntax: `input_argument (datatype) : Definition`

Config input / output arguments for this building block:

- **output\_pdb\_zip\_path** (*str*): Path to the ZIP or PDB file containing the output PDB files. File type: output. Sample file. Accepted formats: pdb, zip.

## Config

Syntax: `input_parameter (datatype) - (default_value) Definition`

Config parameters for this building block:

- **pdb\_code** (*str*) - ('2vgb') RSCB PDB code. ie: "2VGB"
- **filter** (*str*) - (["ATOM", "MODEL", "ENDMDL"]) Array of groups to be kept. If value is None or False no filter will be applied. All the possible values are defined in the official PDB specification (<http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>)
- **cluster** (*str*) - (90) Cluster number for the `:meth:biobb_io.api.MmbPdb.get_pdb_cluster_zip` method.
- **url** (*str*) - ("https://files.rcsb.org/download/") URL of the PDB REST API. Another option for this parameter is the MMB PDB mirror API ("http://mmb.irbbarcelona.org/api/pdb/").
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

## Default config

```
pdb_cluster_zip -o 2vgb_cluster90.zip
```

## YAML

### File config

```
properties:
  pdb_code: 2src
  cluster: 95
```

```
pdb_cluster_zip -c conf.yml -o 2src_cluster95.zip
```

## JSON

### String config

```
pdb_cluster_zip -c '{"pdb_code":"2src", "cluster":95}' -o 2src_cluster95.zip
```

### File config

```
{
  "properties": {
    "pdb_code":"2src",
    "cluster":95
  }
}
```

```
pdb_cluster_zip -c conf.json -o 2src_cluster95.zip
```

## 1.3.5 drugbank

Download a component in SDF format from the [Drugbank](https://www.drugbank.ca/).

### Get help

```
drugbank -h
```

```
usage: drugbank [-h] [-c CONFIG] -o OUTPUT_SDF_PATH
```

Download a component **in** SDF format **from the Drugbank** (<https://www.drugbank.ca/>).

optional arguments:

```
-h, --help          show this help message and exit
-c CONFIG, --config CONFIG
                    This file can be a YAML file, JSON file or JSON string
```

required arguments:

```
-o OUTPUT_SDF_PATH, --output_sdf_path OUTPUT_SDF_PATH
                    Path to the output SDF component file. Accepted formats: sdf.
```

### I / O Arguments

Syntax: `input_argument (datatype) : Definition`

Config input / output arguments for this building block:

- **output\_sdf\_path** (*str*): Path to the output PDB ligand file. File type: output. [Sample file](#). Accepted formats: sdf.

## Config

Syntax: `input_parameter (datatype) - (default_value) Definition`

Config parameters for this building block:

- **drugbank\_id** (*str*) - ("DB00530") Drugbank component id.
- **remove\_tmp** (*bool*) - (True) [WF property] Remove temporal files.
- **restart** (*bool*) - (False) [WF property] Do not execute if output files exist.

### Default config

```
drugbank -o DB00530.sdf
```

## YAML

### File config

```
properties:
  drugbank_id : DB00530
```

```
drugbank -c conf.yml -o output.sdf
```

## JSON

### String config

```
pdb -c '{"drugbank_id":"DB00530"}' -o output.sdf
```

### File config

```
{
  "properties": {
    "drugbank_id": "DB00530"
  }
}
```

```
ligand -c conf.json -o output.sdf
```

## 1.4 Biobb IO changelog

### 1.4.1 What's new in version 3.0.0?

In version 3.0.0 Python has been updated to version 3.7. Big changes in the documentation style and content. Finally a new conda installation recipe has been introduced.

## New features

- Update to Python 3.7 (general)
- New conda installer (installation)
- Adding type hinting for easier usage (API)
- Deprecating `os.path` in favour of `pathlib.path` (modules)
- New command line documentation (documentation)

## Other changes

- New documentation styles (documentation)

## CHAPTER 2

---

### Indices and tables

---

- `genindex`
- `modindex`
- `search`



## CHAPTER 3

---

Github repository.

---





**a**

`api.drugbank`, 7

`api.ligand`, 5

`api.pdb`, 5

`api.pdb_cluster_zip`, 6

`api.pdb_variants`, 6



## A

`api.drugbank` (module), 7  
`api.ligand` (module), 5  
`api.pdb` (module), 5  
`api.pdb_cluster_zip` (module), 6  
`api.pdb_variants` (module), 6

## D

`Drugbank` (class in `api.drugbank`), 7

## L

`launch()` (`api.drugbank.Drugbank` method), 7  
`launch()` (`api.ligand.Ligand` method), 5  
`launch()` (`api.pdb.Pdb` method), 5  
`launch()` (`api.pdb_cluster_zip.MmbPdbClusterZip` method), 7  
`launch()` (`api.pdb_variants.MmbPdbVariants` method), 6  
`Ligand` (class in `api.ligand`), 5

## M

`main()` (in module `api.drugbank`), 7  
`main()` (in module `api.ligand`), 5  
`main()` (in module `api.pdb`), 6  
`main()` (in module `api.pdb_cluster_zip`), 7  
`main()` (in module `api.pdb_variants`), 6  
`MmbPdbClusterZip` (class in `api.pdb_cluster_zip`), 6  
`MmbPdbVariants` (class in `api.pdb_variants`), 6

## P

`Pdb` (class in `api.pdb`), 5