EMDB

EM Standards Workshop 2024

Kyle L. Morris
EMDB Team Leader

kyle@ebi.ac.uk

EMBL-EBI

EMDB
Electron Microscopy Data Bank

WORLDWIDE
wwPDB
PROTEIN DATA BANK

PDBe

EMPIAR

Sanja

Jack

Amudha

Lucas

Neli

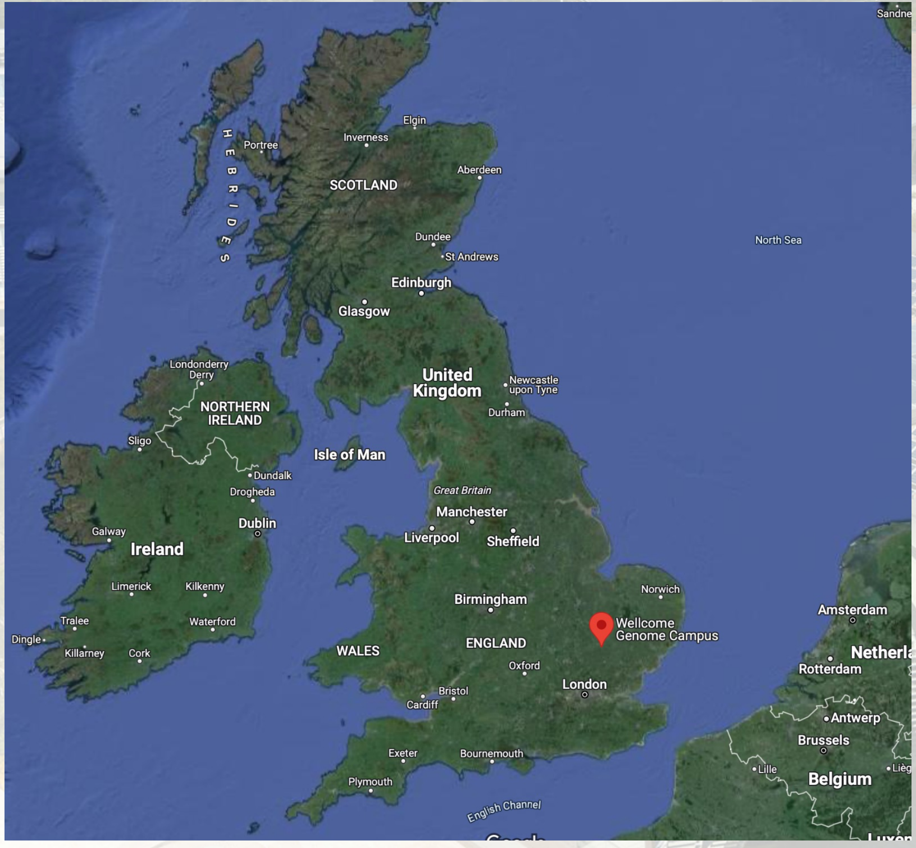
Zhe

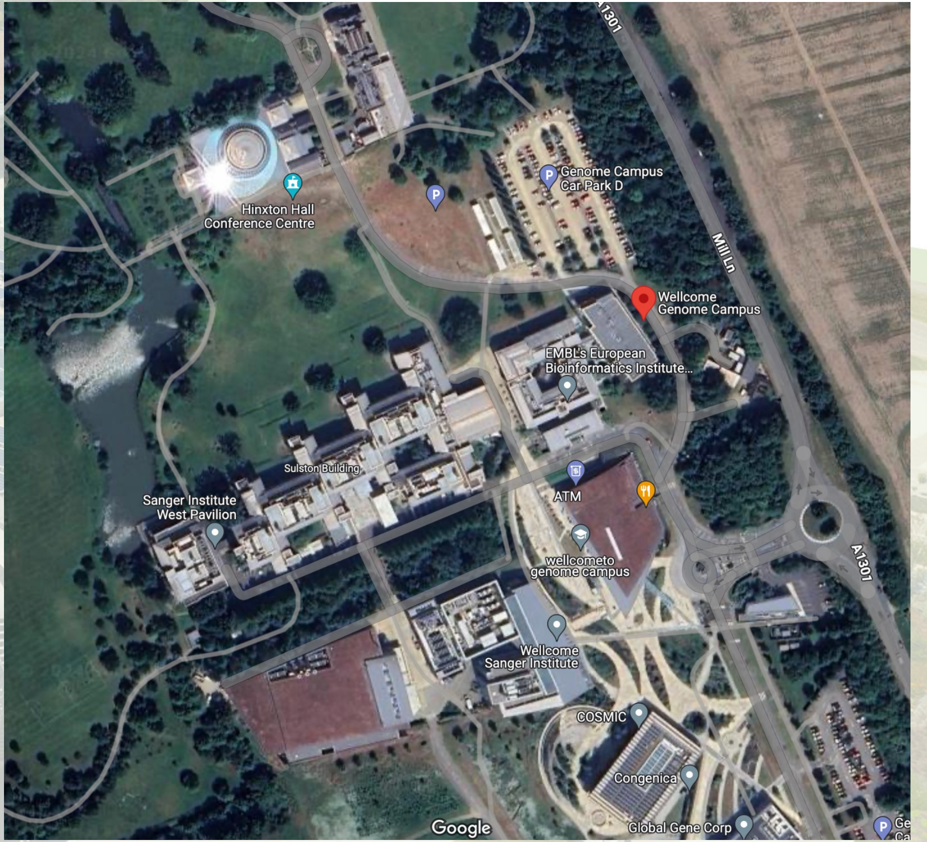
Osman

BioImage Archive

EMBL-EBI

wellcome


Sandne
Elgin
Portree
Inverness
H E B R I D E S
Aberdeen
SCOTLAND
Dundee
St Andrews
North Sea
Edinburgh
Glasgow
Londonderry
Derry
United
Kingdom
Newcastle
upon Tyne
NORTHERN
IRELAND
Durham
Sligo
Isle of Man
Dundalk
Drogheda
Great Britain
Manchester
Dublin
Liverpool
Sheffield
Galway
Ireland
Limerick
Kilkenny
Norwich
Birmingham
Amsterdam
Tralee
Waterford
Wellcome
Genome Campus
Dingle
Killarney
Cork
WALES
ENGLAND
Netherla
Rotterdam
Oxford
London
Bristol
Cardiff
Antwerp
Brussels
Exeter
Bournemouth
Lille
Belgium
Liège
Plymouth
English Channel


A1301
Genome Campus
Car Park D
Hinxton Hall
Conference Centre
Mill Ln
Wellcome
Genome Campus
EMBL's European
Bioinformatics Institute...
Sulston Building
ATM
Sanger Institute
West Pavilion
wellcometo
genome campus
A1301
Wellcome
Sanger Institute
COSMIC
Congenica
Google
Global Gene Corp
Ge

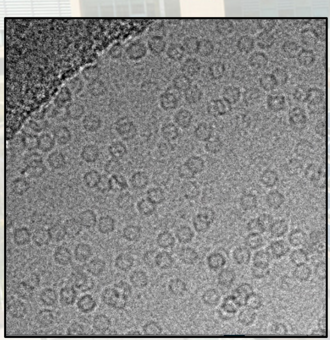
"Don't mind Ashley. After looking through a microscope all day, anything large startles him."

I come from:
- experimental training
- doing microscopy
- doing structural biology

- data producer
- facility staff
- data consumer
- data steward

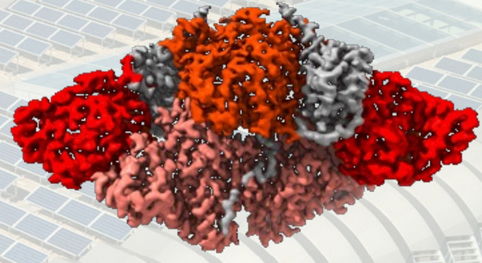
I'm here to listen

Raw



EMPIAR

Derived data



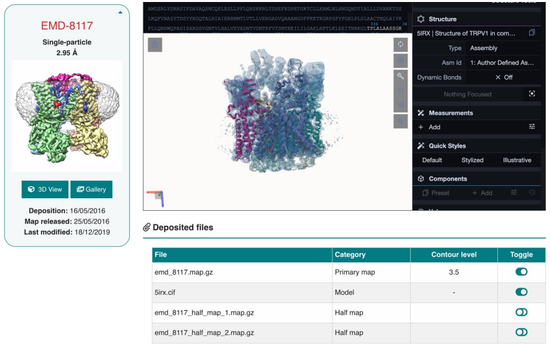
EMDB

Electron Microscopy Data Bank

Derived data

PDBe

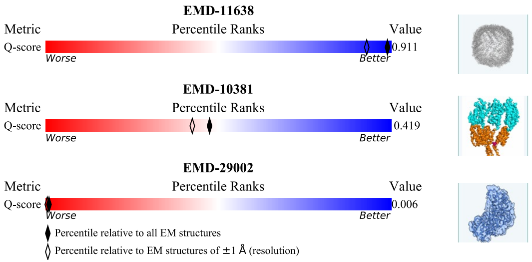
EMDB: What’s new

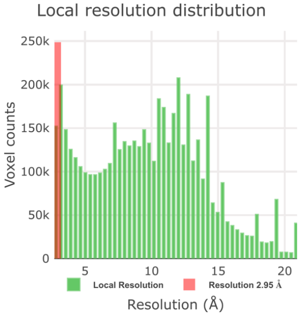


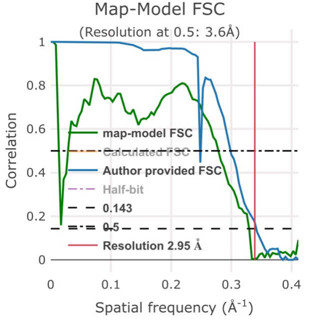
Deposition: 16/05/2016
Map released: 25/05/2016
Last modified: 18/12/2019

Deposited files

File	Category	Contour level	Toggle
emd_8117.map.gz	Primary map	3.5	●
5irx.cif	Model	-	●
emd_8117_half_map_1.map.gz	Half map		○
emd_8117_half_map_2.map.gz	Half map		○







Mol*

Modern molecular viewer

Explore entry maps

View validation

Human readable validation

Evaluation bars

Map-model metric summary

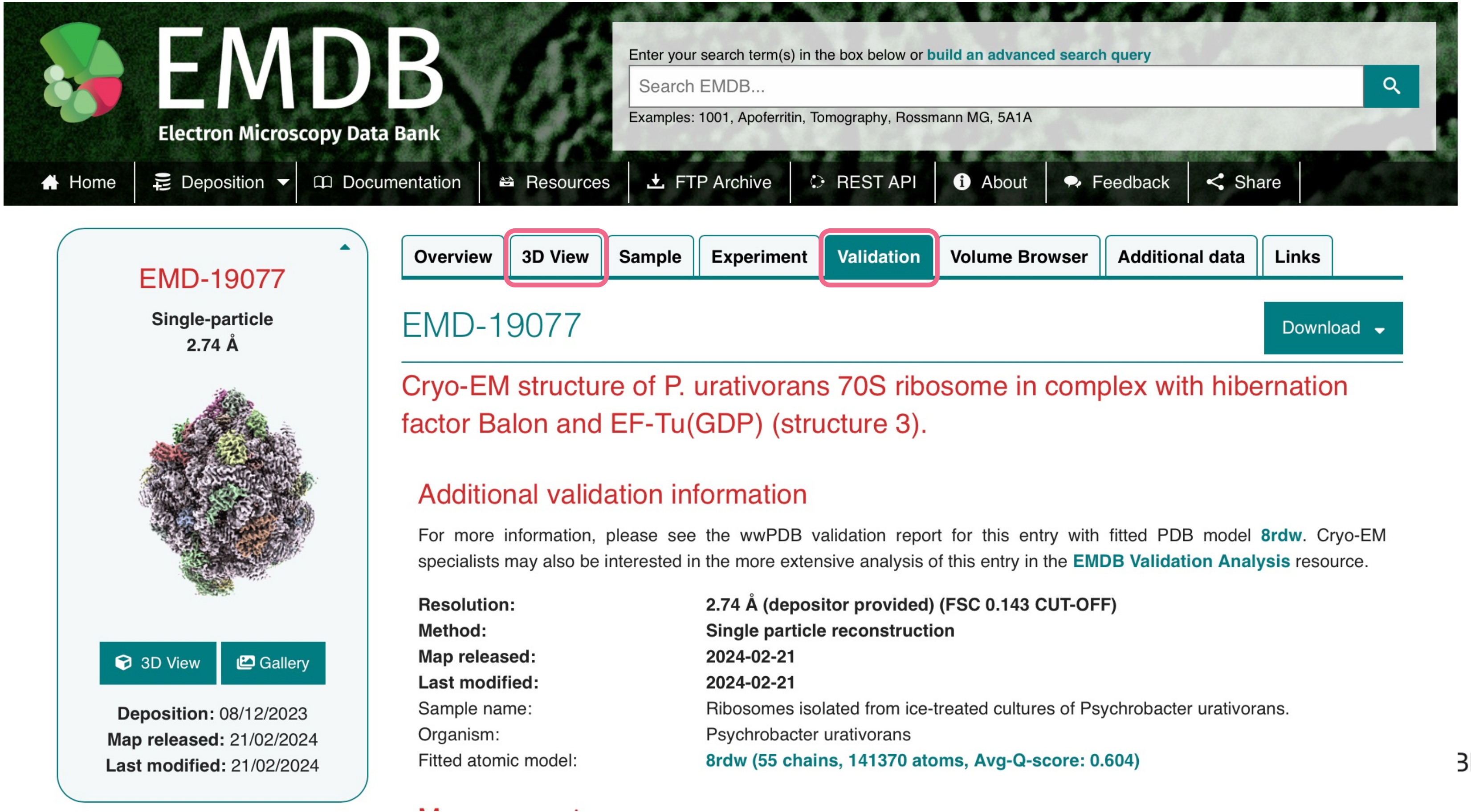
Critically needed validation

Phase randomisation FSC

Map-model FSC

Local resolution

EMDB: What’s new



Enter your search term(s) in the box below or **build an advanced search query**

Search EMDB...

Examples: 1001, Apoferritin, Tomography, Rossmann MG, 5A1A

EMD-19077

Single-particle
2.74 Å

Deposition: 08/12/2023
Map released: 21/02/2024
Last modified: 21/02/2024

EMD-19077

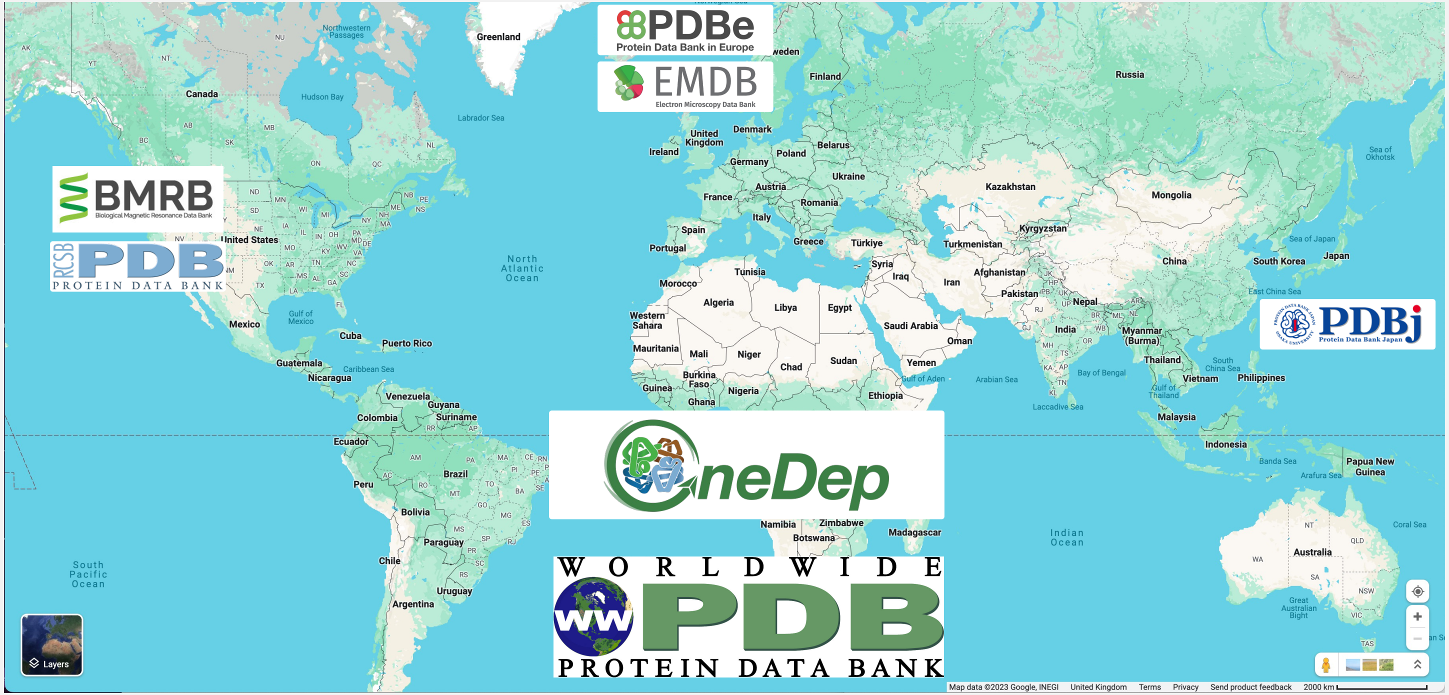
Cryo-EM structure of P. urativorans 70S ribosome in complex with hibernation factor Balon and EF-Tu(GDP) (structure 3).

Additional validation information

For more information, please see the wwPDB validation report for this entry with fitted PDB model **8rdw**. Cryo-EM specialists may also be interested in the more extensive analysis of this entry in the **EMDB Validation Analysis** resource.

Resolution:	**2.74 Å (depositor provided) (FSC 0.143 CUT-OFF)**
Method:	**Single particle reconstruction**
Map released:	**2024-02-21**
Last modified:	**2024-02-21**
Sample name:	Ribosomes isolated from ice-treated cultures of Psychrobacter urativorans.
Organism:	Psychrobacter urativorans
Fitted atomic model:	**8rdw (55 chains, 141370 atoms, Avg-Q-score: 0.604)**

Data and metadata comes through OneDep



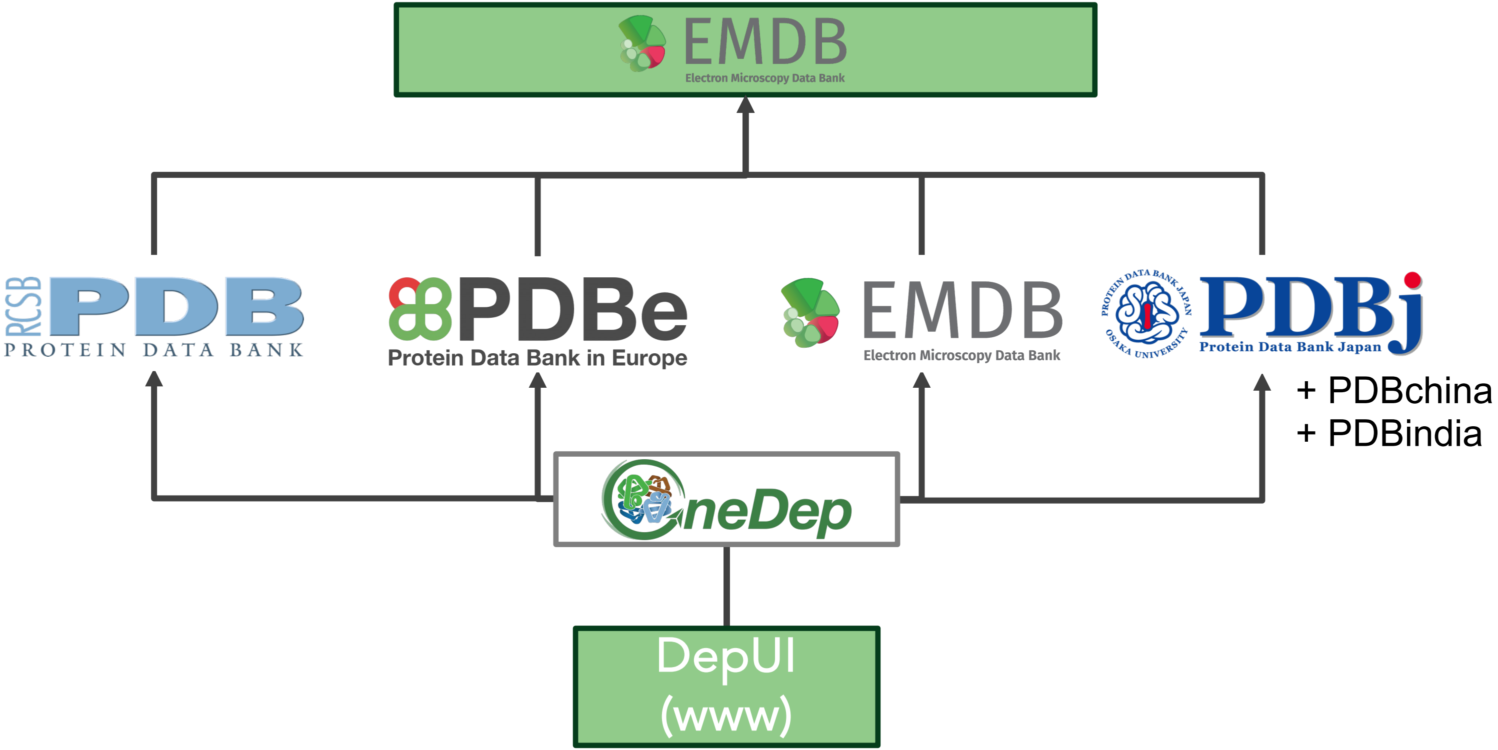
Data and metadata comes through OneDep

Archive

Curation

Processing

Deposition



+ PDBchina
+ PDBindia

DepUI
(www)

Metadata exists in XML and mmCIF data models

Website

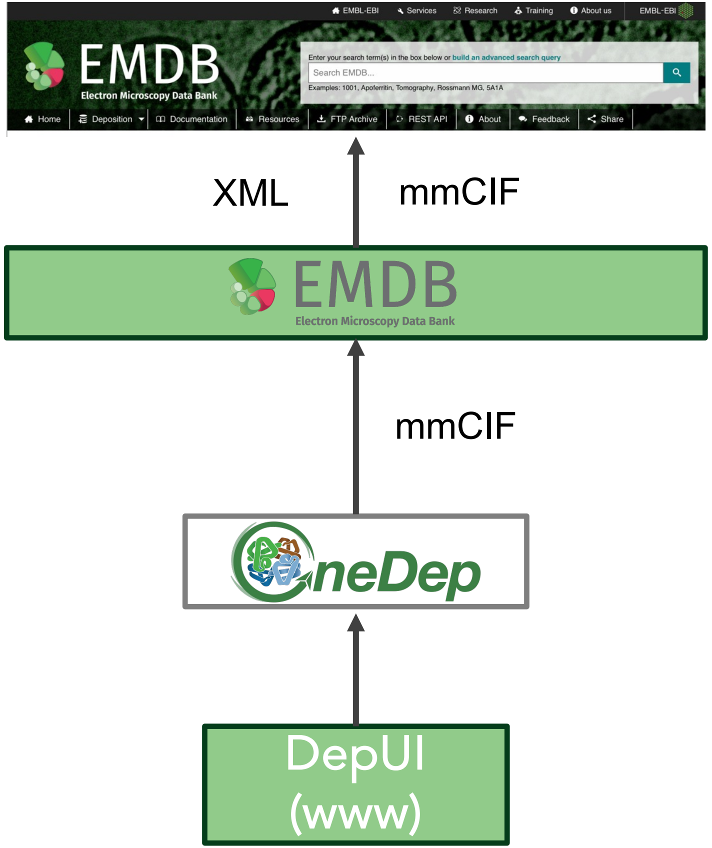
Metadata created

Archive

Metadata created

Processing

Deposition



Enter your search term(s) in the box below or build an advanced search query

Examples: 1001, Apoferritin, Tomography, Rossmann MG, 5A1A

XML

mmCIF

EMDB
Electron Microscopy Data Bank

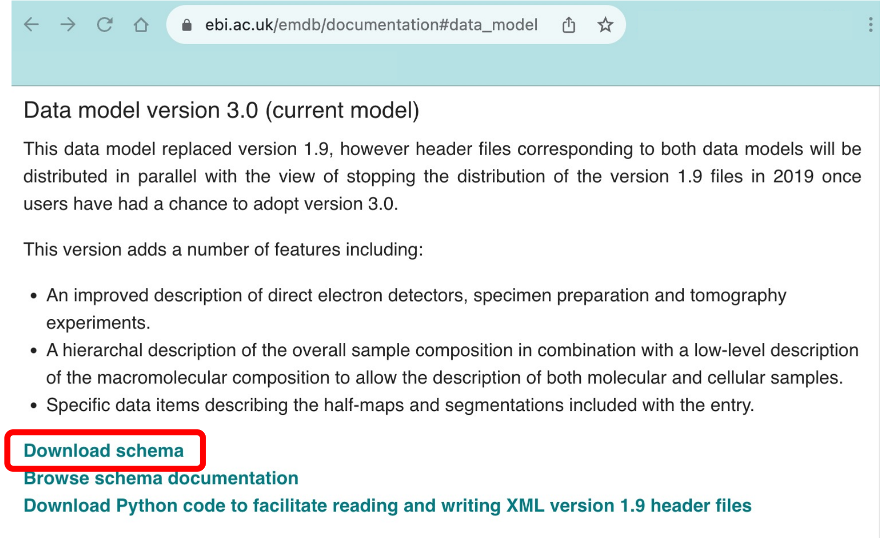
mmCIF

OneDep

DepUI
(www)

EMBL-EBI

EMDB data model



Data model version 3.0 (current model)

This data model replaced version 1.9, however header files corresponding to both data models will be distributed in parallel with the view of stopping the distribution of the version 1.9 files in 2019 once users have had a chance to adopt version 3.0.

This version adds a number of features including:

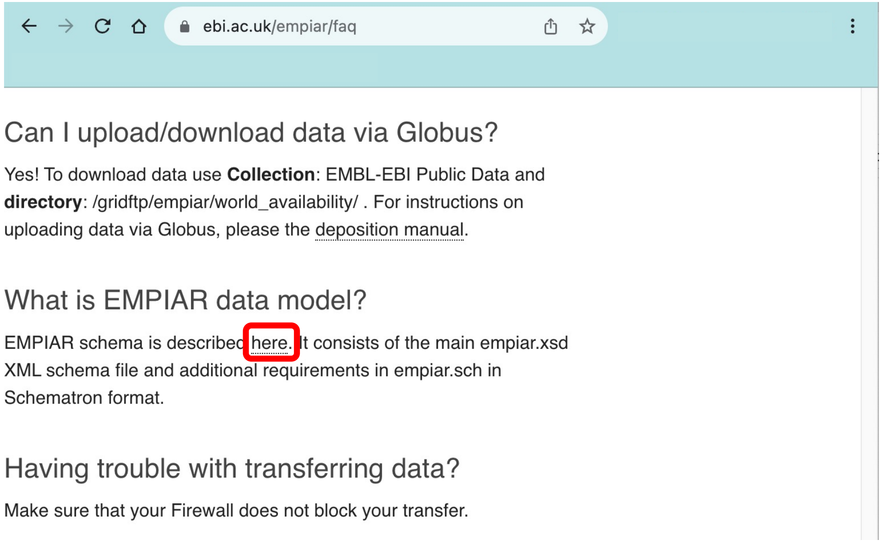
- An improved description of direct electron detectors, specimen preparation and tomography experiments.
- A hierarchal description of the overall sample composition in combination with a low-level description of the macromolecular composition to allow the description of both molecular and cellular samples.
- Specific data items describing the half-maps and segmentations included with the entry.

Download schema
Browse schema documentation
Download Python code to facilitate reading and writing XML version 1.9 header files

https://www.ebi.ac.uk/emdb/documentation#data_model

https://ftp.ebi.ac.uk/pub/databases/emdb/doc/XML-schemas/emdb-schemas/v3/current_v3/emdb.xsd

EMPIAR data model



Can I upload/download data via Globus?

Yes! To download data use **Collection**: EMBL-EBI Public Data and **directory**: /gridftp/empiar/world_availability/ . For instructions on uploading data via Globus, please the deposition manual.

What is EMPIAR data model?

EMPIAR schema is described here. It consists of the main empiar.xsd XML schema file and additional requirements in empiar.sch in Schematron format.

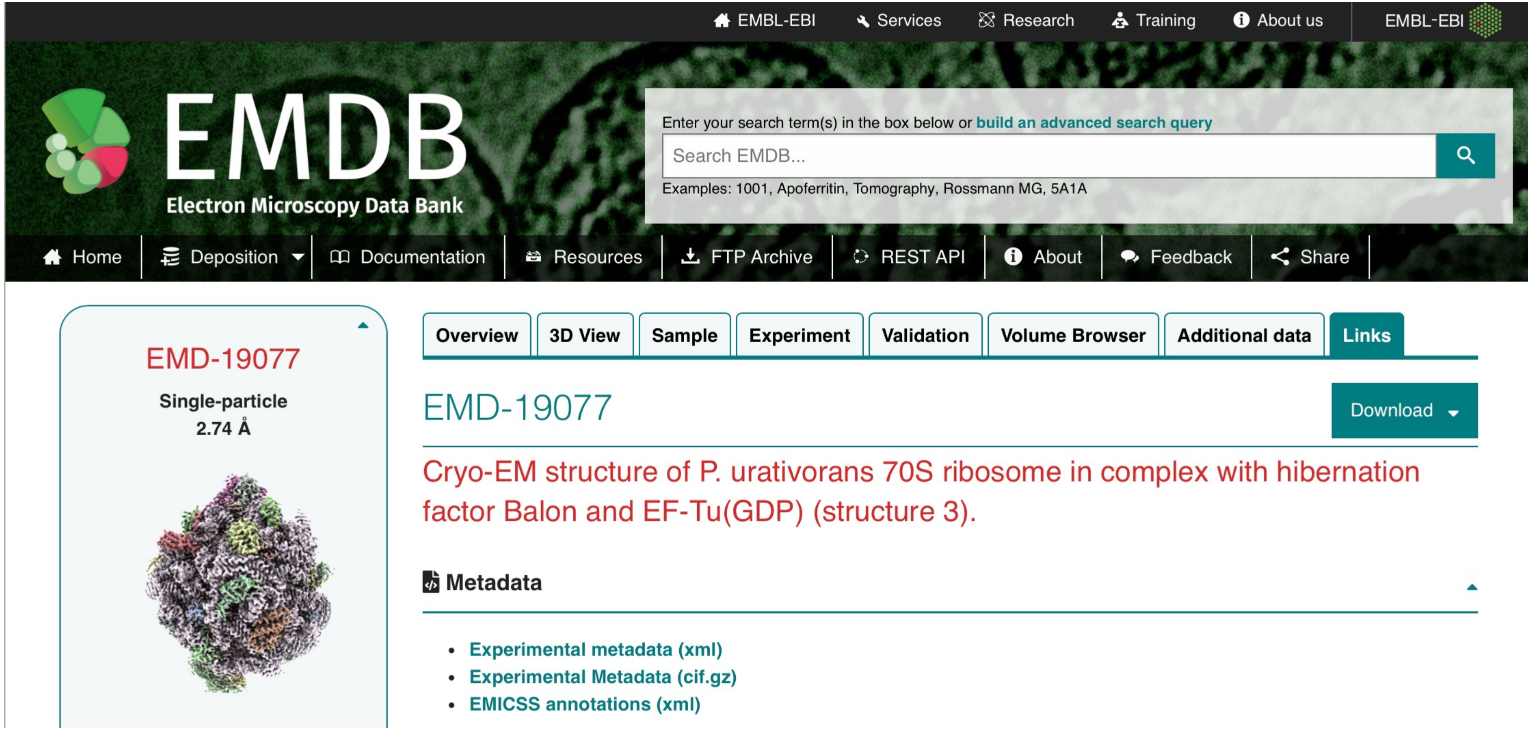
Having trouble with transferring data?

Make sure that your Firewall does not block your transfer.

https://www.ebi.ac.uk/empiar/faq

https://ftp.ebi.ac.uk/pub/databases/emtest/empiar/schema/empiar.xsd

Metadata exists in XML and mmCIF data models



Enter your search term(s) in the box below or build an advanced search query

Search EMDB...

Examples: 1001, Apoferritin, Tomography, Rossmann MG, 5A1A

EMD-19077

Single-particle
2.74 Å

EMD-19077

Cryo-EM structure of P. urativorans 70S ribosome in complex with hibernation factor Balon and EF-Tu(GDP) (structure 3).

Metadata

- Experimental metadata (xml)
- Experimental Metadata (cif.gz)
- EMICSS annotations (xml)

Data model schema

- XML schema:
  https://www.ebi.ac.uk/emdb/documentation#version30
- Browse XML schema:
  http://ftp.ebi.ac.uk/pub/databases/emdb/doc/XML-schemas/emdb-schemas/v3/current_v3/doc/Untitled.html

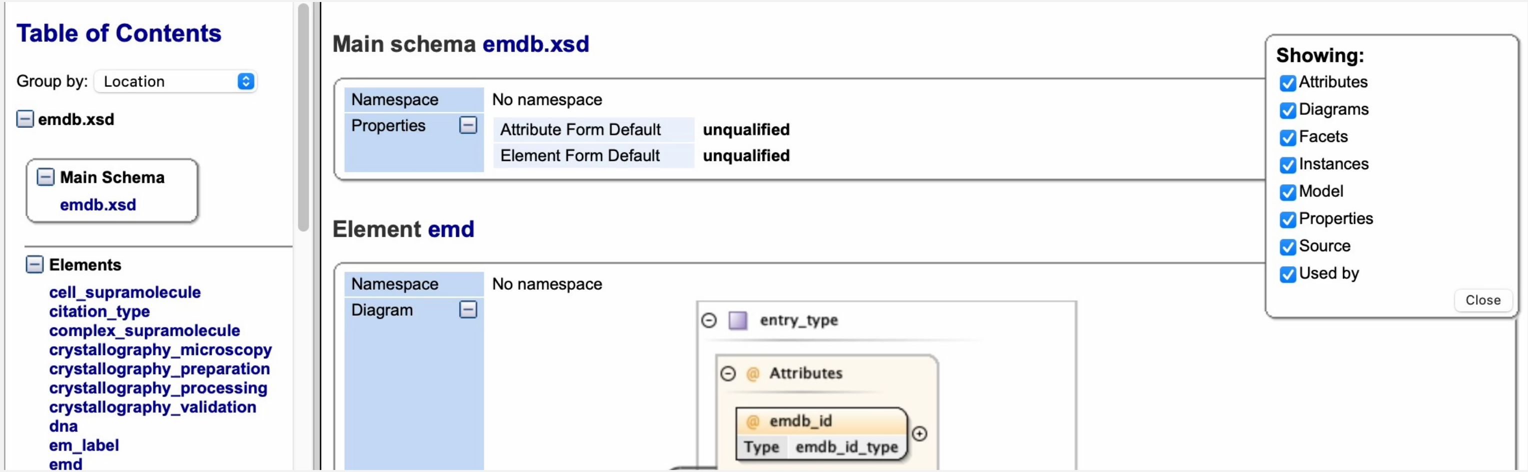


Table of Contents

Group by: Location

⊟ **emdb.xsd**

⊟ **Main Schema**
emdb.xsd

⊟ **Elements**
cell_supramolecule
citation_type
complex_supramolecule
crystallography_microscopy
crystallography_preparation
crystallography_processing
crystallography_validation
dna
em_label
emd

Main schema emdb.xsd

Namespace	No namespace	
Properties ⊟	Attribute Form Default	**unqualified**
	Element Form Default	**unqualified**

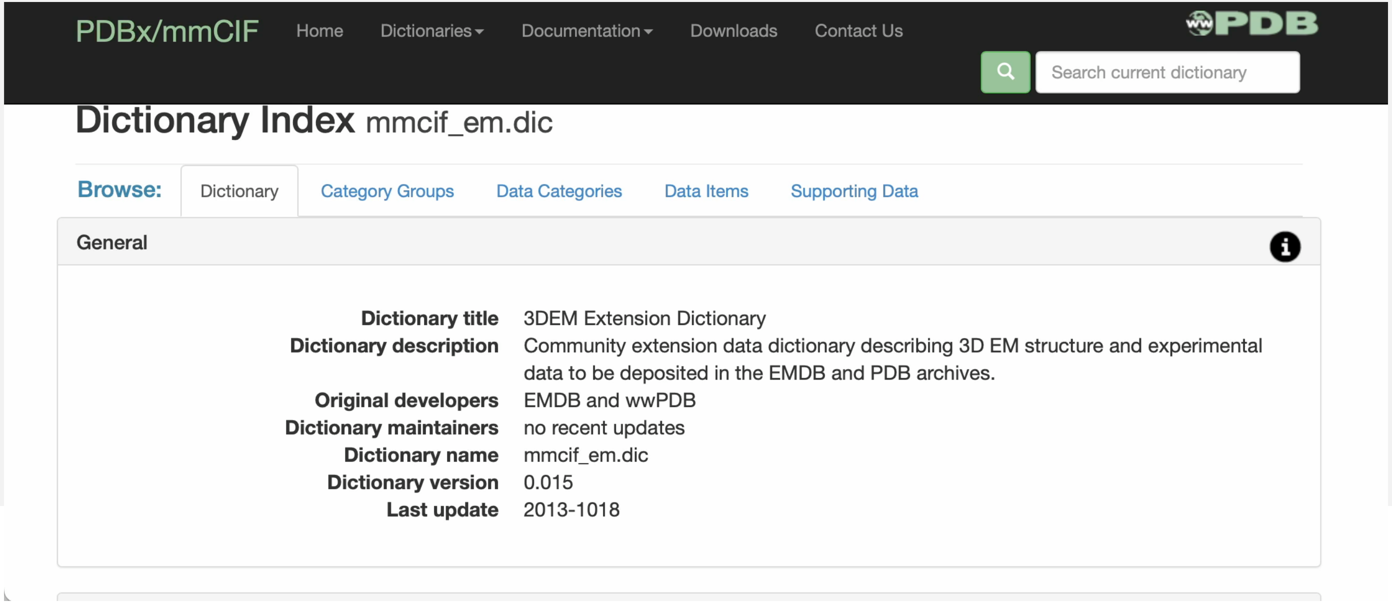
Element emd

Namespace	No namespace
Diagram ⊟	

Showing:
☑ Attributes
☑ Diagrams
☑ Facets
☑ Instances
☑ Model
☑ Properties
☑ Source
☑ Used by

Data model schema

- PDBx/mmCIF dictionary
  https://mmcif.wwpdb.org/dictionaries/ascii/mmcif_em.dic
- Browse mmCIF dictionary:
  https://mmcif.wwpdb.org/dictionaries/mmcif_em.dic/Index/index.html



Dictionary Index mmcif_em.dic

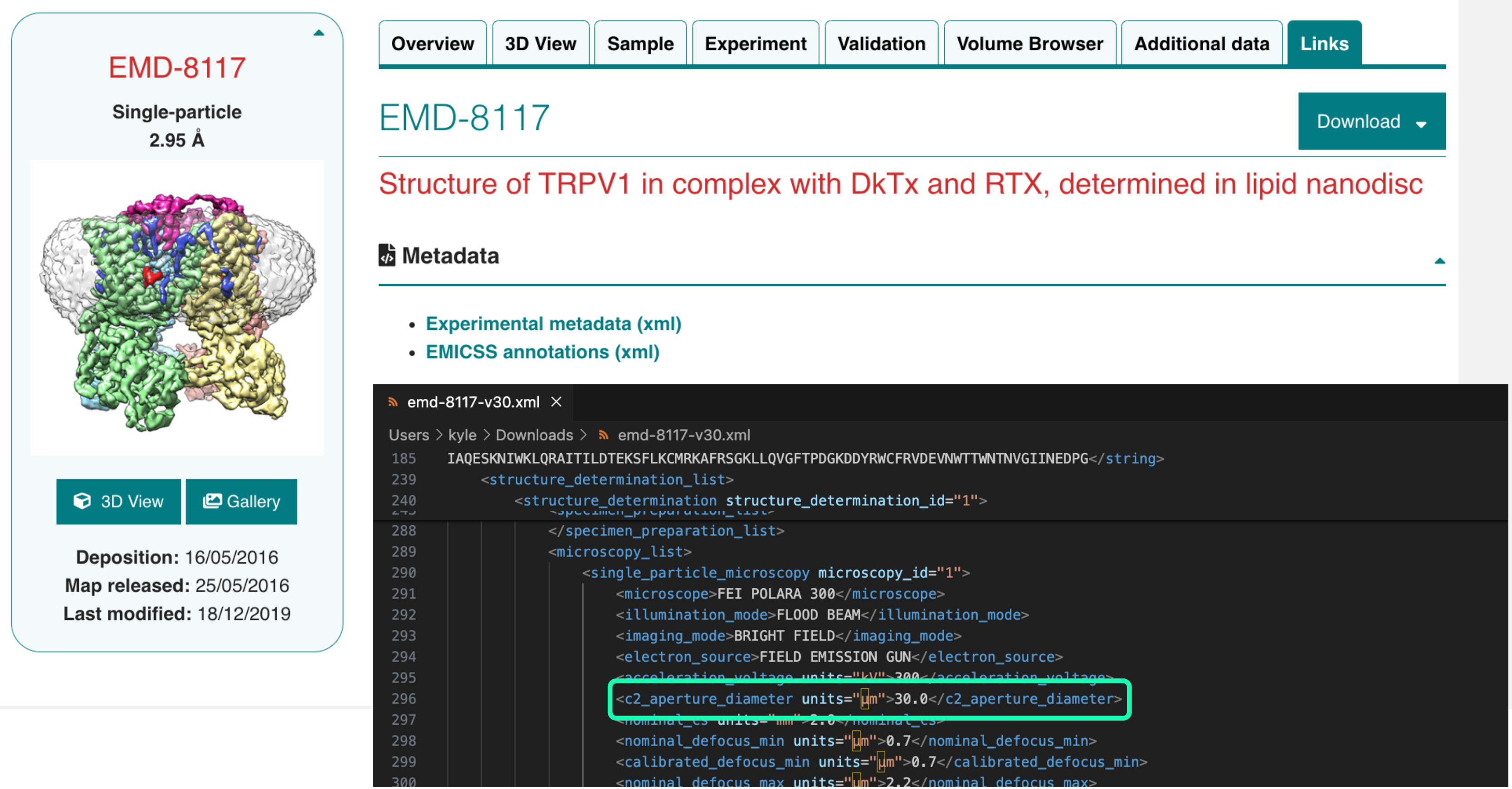
General

Dictionary title	3DEM Extension Dictionary
Dictionary description	Community extension data dictionary describing 3D EM structure and experimental data to be deposited in the EMDB and PDB archives.
Original developers	EMDB and wwPDB
Dictionary maintainers	no recent updates
Dictionary name	mmcif_em.dic
Dictionary version	0.015
Last update	2013-1018

Data model

XML 3.0	mmCIF
microscope	_em_imaging.microscope_model
illumination_mode	_em_imaging.illumination_mode
imaging_mode	_em_imaging.mode
acceleration_voltage	_em_imaging.accelerating_voltage
c2_aperture_diameter	_em_imaging.c2_aperture_diameter
nominal_cs	_em_imaging.nominal_cs
nominal_defocus_min	_em_imaging.nominal_defocus_min
nominal_defocus_max	_em_imaging.nominal_defocus_max

Data model



Deposition: 16/05/2016
Map released: 25/05/2016
Last modified: 18/12/2019

EMD-8117

Structure of TRPV1 in complex with DkTx and RTX, determined in lipid nanodisc

Metadata

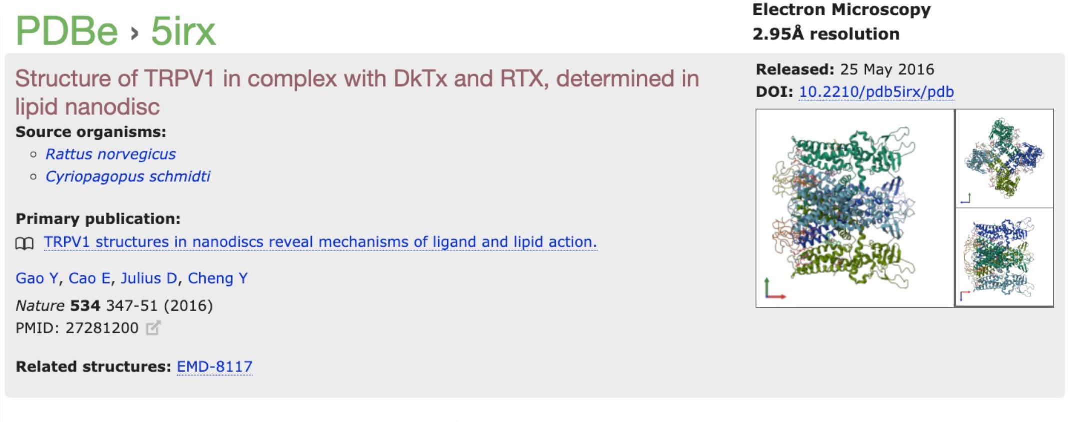
- Experimental metadata (xml)
- EMICSS annotations (xml)

emd-8117-v30.xml

Users > kyle > Downloads > emd-8117-v30.xml

```
IAQESKNIWKLQRAITILDTEKSFLKCMRKAFRSGKLLQVGFTPDGKDDYRWCFRVDEVNWTTWNTNVGIINEDPG</string>
    <structure_determination_list>
        <structure_determination structure_determination_id="1">
            </specimen_preparation_list>
            <microscopy_list>
                <single_particle_microscopy microscopy_id="1">
                    <microscope>FEI POLARA 300</microscope>
                    <illumination_mode>FLOOD BEAM</illumination_mode>
                    <imaging_mode>BRIGHT FIELD</imaging_mode>
                    <electron_source>FIELD EMISSION GUN</electron_source>
                    <acceleration_voltage units="kV">300</acceleration_voltage>
                    <c2_aperture_diameter units="µm">30.0</c2_aperture_diameter>
                    <nominal_cs units="mm">2.0</nominal_cs>
                    <nominal_defocus_min units="µm">0.7</nominal_defocus_min>
                    <calibrated_defocus_min units="µm">0.7</calibrated_defocus_min>
                    <nominal_defocus_max units="µm">2.2</nominal_defocus_max>
```

Data model



PDBe › 5irx

Electron Microscopy
2.95Å resolution

Structure of TRPV1 in complex with DkTx and RTX, determined in lipid nanodisc

Source organisms:

- *Rattus norvegicus*
- *Cyriopagopus schmidti*

Primary publication:

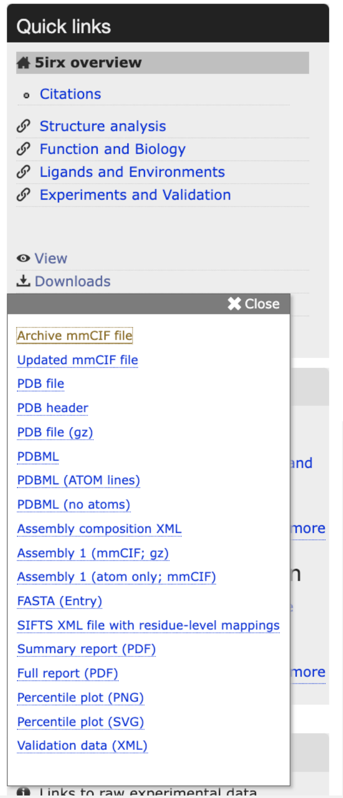
TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action.

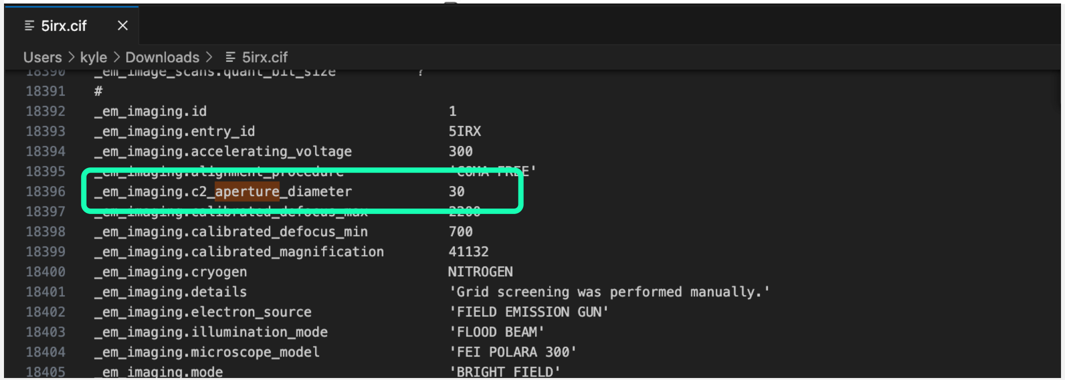
Gao Y, Cao E, Julius D, Cheng Y

Nature **534** 347-51 (2016)
PMID: 27281200

Related structures: EMD-8117

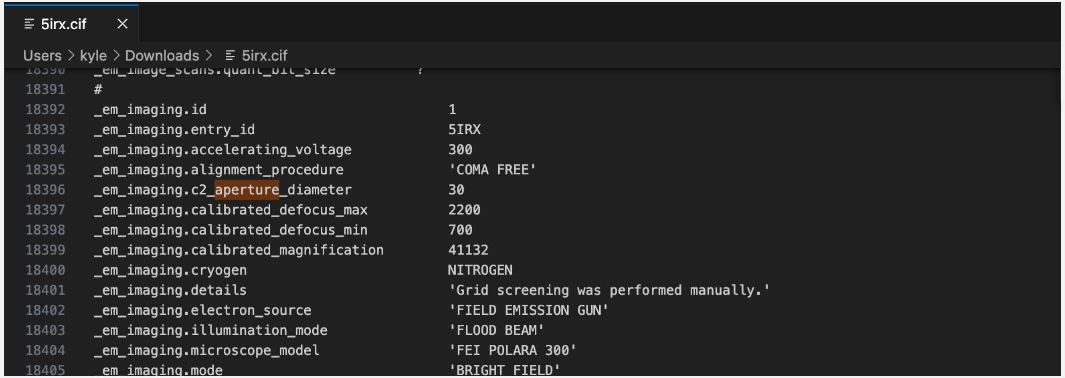
Released: 25 May 2016
DOI: 10.2210/pdb5irx/pdb



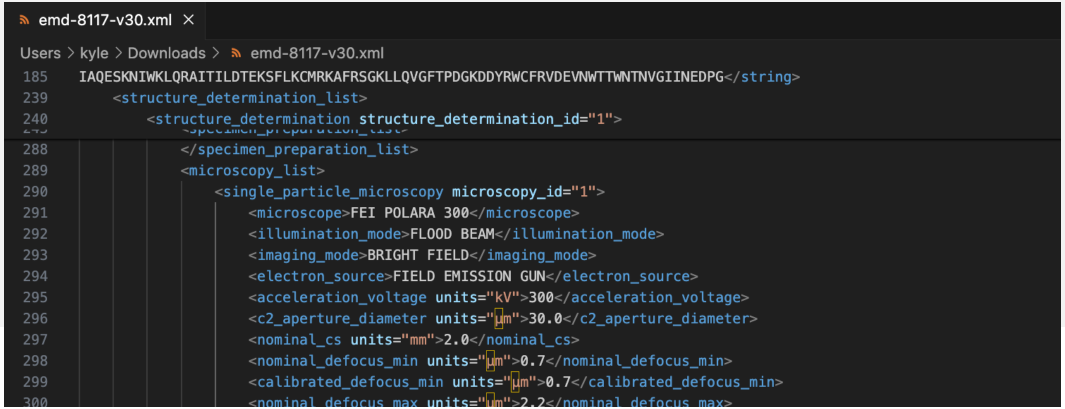


```
_em_image_scans.quant_bit_size           ?
#
_em_imaging.id                           1
_em_imaging.entry_id                     5IRX
_em_imaging.accelerating_voltage         300
_em_imaging.alignment_procedure          'COMA FREE'
_em_imaging.c2_aperture_diameter         30
_em_imaging.calibrated_defocus_max       2200
_em_imaging.calibrated_defocus_min       700
_em_imaging.calibrated_magnification     41132
_em_imaging.cryogen                      NITROGEN
_em_imaging.details                      'Grid screening was performed manually.'
_em_imaging.electron_source              'FIELD EMISSION GUN'
_em_imaging.illumination_mode            'FLOOD BEAM'
_em_imaging.microscope_model             'FEI POLARA 300'
_em_imaging.mode                         'BRIGHT FIELD'
```

Data model



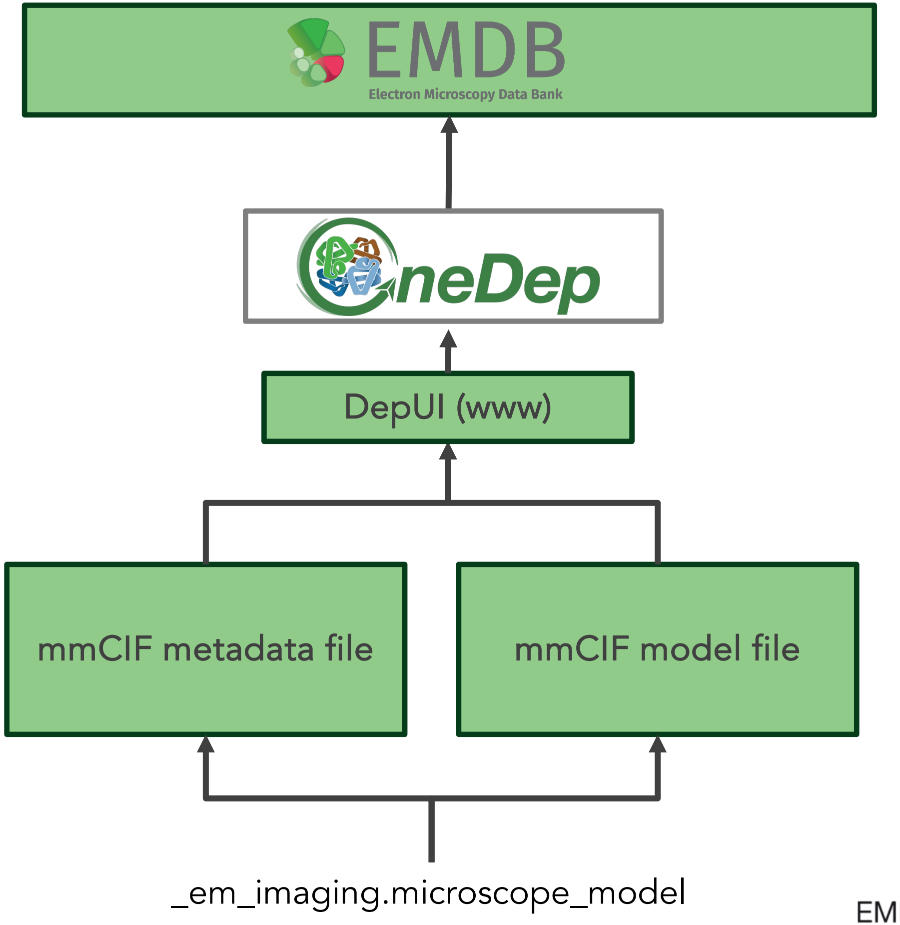
```
#
_em_imaging.id                              1
_em_imaging.entry_id                        5IRX
_em_imaging.accelerating_voltage            300
_em_imaging.alignment_procedure             'COMA FREE'
_em_imaging.c2_aperture_diameter            30
_em_imaging.calibrated_defocus_max          2200
_em_imaging.calibrated_defocus_min          700
_em_imaging.calibrated_magnification        41132
_em_imaging.cryogen                         NITROGEN
_em_imaging.details                         'Grid screening was performed manually.'
_em_imaging.electron_source                 'FIELD EMISSION GUN'
_em_imaging.illumination_mode               'FLOOD BEAM'
_em_imaging.microscope_model                'FEI POLARA 300'
_em_imaging.mode                            'BRIGHT FIELD'
```



```
IAQESKNIWKLQRAITILDTEKSFLKCMRKAFRSGKLLQVGFTPDGKDDYRWCFRVDEVNWTTWNTNVGIINEDPG</string>
    <structure_determination_list>
        <structure_determination structure_determination_id="1">
            </specimen_preparation_list>
            <microscopy_list>
                <single_particle_microscopy microscopy_id="1">
                    <microscope>FEI POLARA 300</microscope>
                    <illumination_mode>FLOOD BEAM</illumination_mode>
                    <imaging_mode>BRIGHT FIELD</imaging_mode>
                    <electron_source>FIELD EMISSION GUN</electron_source>
                    <acceleration_voltage units="kV">300</acceleration_voltage>
                    <c2_aperture_diameter units="µm">30.0</c2_aperture_diameter>
                    <nominal_cs units="mm">2.0</nominal_cs>
                    <nominal_defocus_min units="µm">0.7</nominal_defocus_min>
                    <calibrated_defocus_min units="µm">0.7</calibrated_defocus_min>
                    <nominal_defocus_max units="µm">2.2</nominal_defocus_max>
```

Data in

Archive



EMDB
Electron Microscopy Data Bank

OneDep

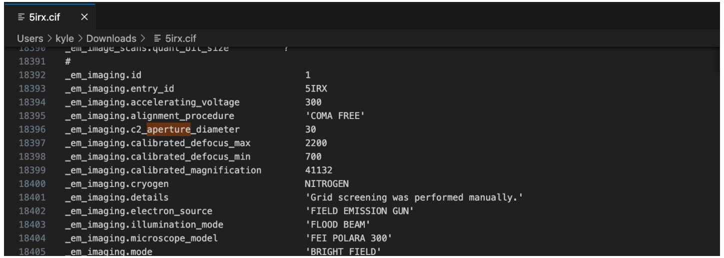
DepUI (www)

mmCIF metadata file

mmCIF model file

_em_imaging.microscope_model

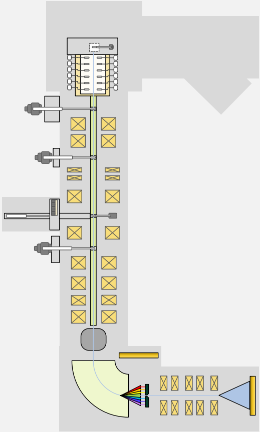
Metadata capture



```
_em_image_scans.quant_bit_size          ?
#
_em_imaging.id                           1
_em_imaging.entry_id                     5IRX
_em_imaging.accelerating_voltage         300
_em_imaging.alignment_procedure          'COMA FREE'
_em_imaging.c2_aperture_diameter         30
_em_imaging.calibrated_defocus_max       2200
_em_imaging.calibrated_defocus_min       700
_em_imaging.calibrated_magnification     41132
_em_imaging.cryogen                      NITROGEN
_em_imaging.details                      'Grid screening was performed manually.'
_em_imaging.electron_source              'FIELD EMISSION GUN'
_em_imaging.illumination_mode            'FLOOD BEAM'
_em_imaging.microscope_model             'FEI POLARA 300'
_em_imaging.mode                         'BRIGHT FIELD'
```

Depositor

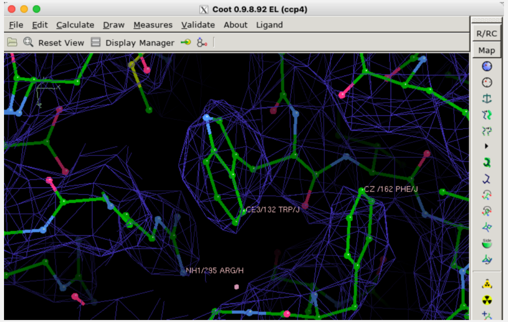
mmCIF models with properly populated metadata would already make depositors lives easier



cryoSPARC

SCIPION

CCP-EM



instrument

processing

modelling

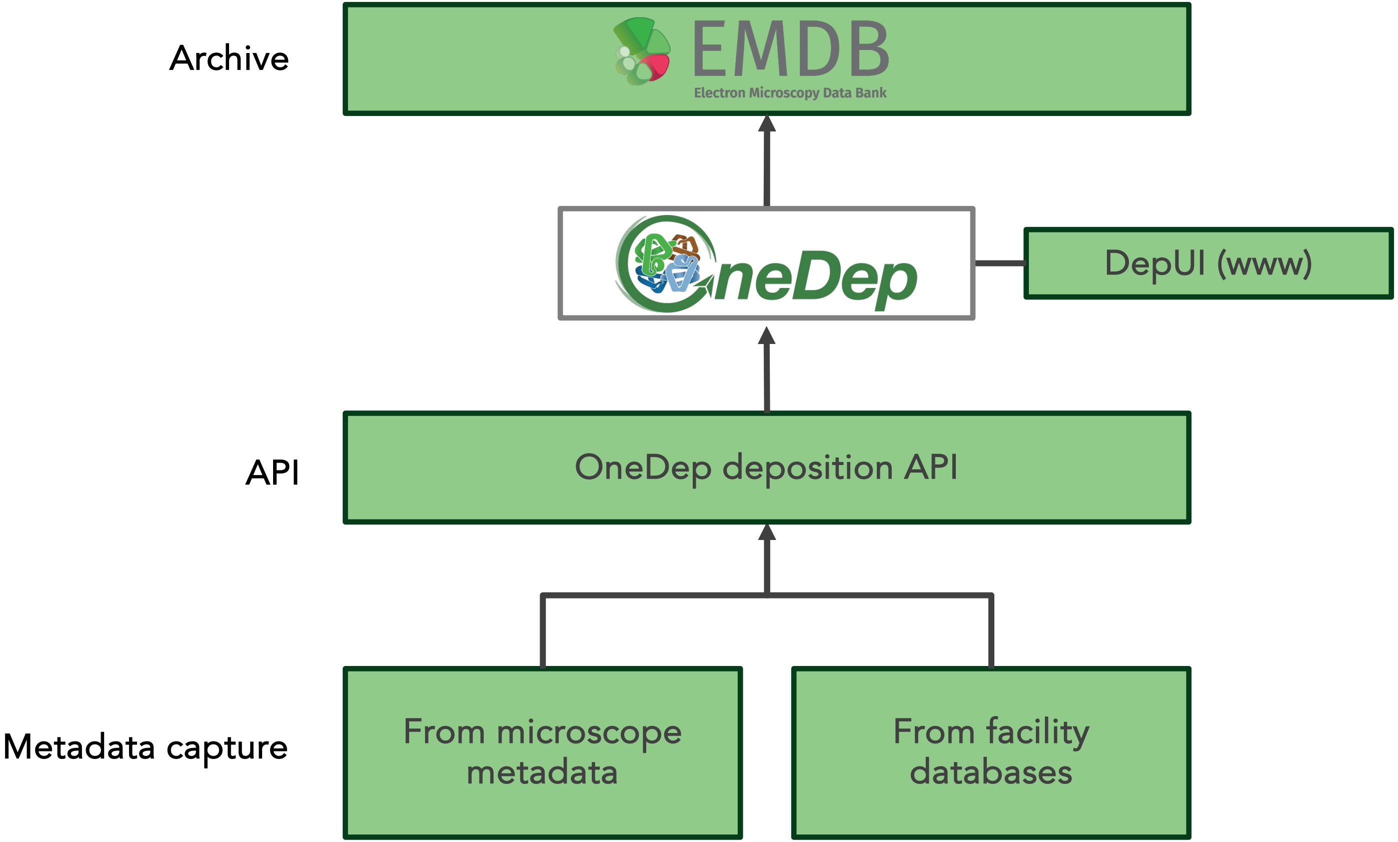
mmCIF model
+instrument metadata
+processing metadata
+modelling metadata

Data in

Weslley (PDBe)

Lucas (EMDB)

Neli (EMDB)



Archive

EMDB

Electron Microscopy Data Bank

OneDep

DepUI (www)

API

OneDep deposition API

Metadata capture

From microscope metadata

From facility databases

EMBL-EBI

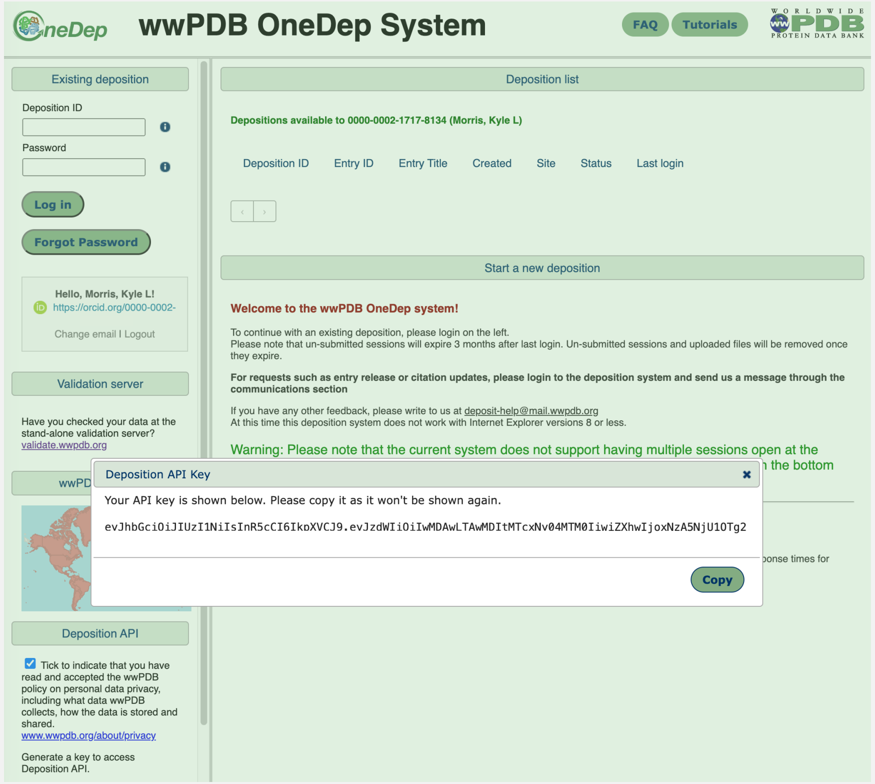
Data in API

OneDep deposition test server

API keys can be generated

They are linked to your ORCID

Developers may already begin to test deposition to this test server and we would appreciate your feedback



Depositions available to 0000-0002-1717-8134 (Morris, Kyle L)

Hello, Morris, Kyle L!
https://orcid.org/0000-0002-

Have you checked your data at the stand-alone validation server?
validate.wwpdb.org

☑ Tick to indicate that you have read and accepted the wwPDB policy on personal data privacy, including what data wwPDB collects, how the data is stored and shared.
www.wwpdb.org/about/privacy

Generate a key to access Deposition API.

Welcome to the wwPDB OneDep system!

To continue with an existing deposition, please login on the left.
Please note that un-submitted sessions will expire 3 months after last login. Un-submitted sessions and uploaded files will be removed once they expire.

For requests such as entry release or citation updates, please login to the deposition system and send us a message through the communications section

If you have any other feedback, please write to us at deposit-help@mail.wwpdb.org
At this time this deposition system does not work with Internet Explorer versions 8 or less.

Warning: Please note that the current system does not support having multiple sessions open at the

the bottom

oonse times for

Your API key is shown below. Please copy it as it won't be shown again.

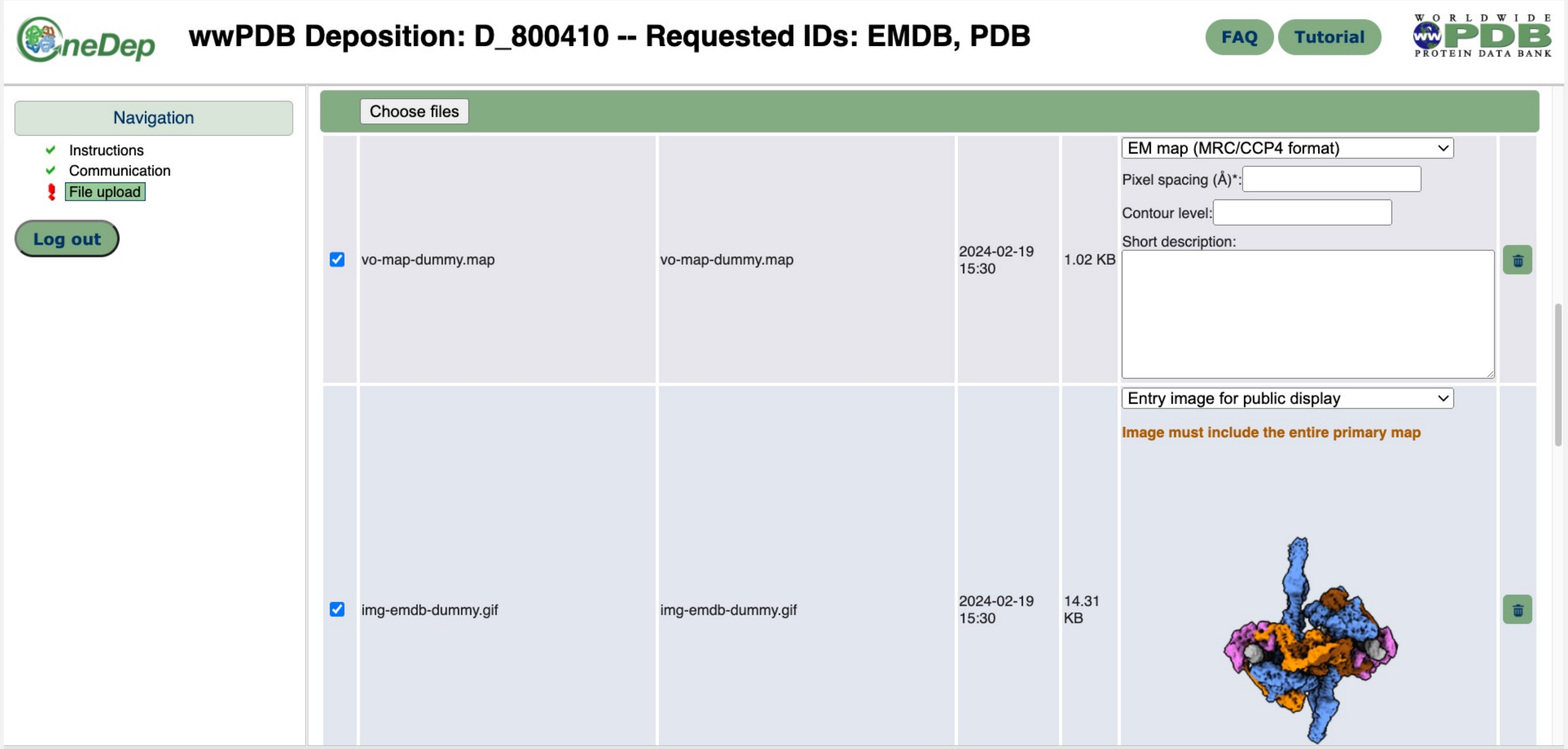
```
evJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.evJzdWIiOiIwMDAwLTAwMDItMTcxNv04MTM0IiwiZXhwIjoxNzA5NjU1OTg2
```

Data in API

- Deposition test server accepting deposition via API
  - https://onedep-depui-test.wwpdb.org/deposition
- Wrapper and CLI for API
  - https://github.com/wwPDB/py-onedep_deposition/blob/master/docs/cli.md
- Low level API documentation on apiary.io
  - https://depapi1.docs.apiary.io/

- EMDB are testing file upload via API
- Providing metadata via specific end points
- File upload should be able to interpret and populate deposition from file upload of mmCIF

Data in API – test server

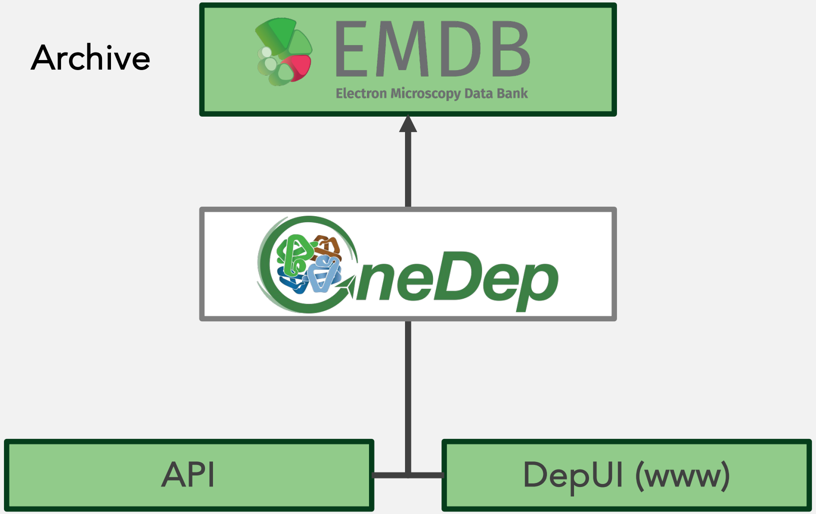


wwPDB Deposition: D_800410 -- Requested IDs: EMDB, PDB

☑	vo-map-dummy.map	vo-map-dummy.map	2024-02-19 15:30	1.02 KB	EM map (MRC/CCP4 format) Pixel spacing (Å)*: Contour level: Short description:	
☑	img-emdb-dummy.gif	img-emdb-dummy.gif	2024-02-19 15:30	14.31 KB	Entry image for public display Image must include the entire primary map	

Data model

mmCIF
_em_imaging.accelerating_voltage
_em_imaging.nominal_cs
_em_specimen.concentration
_em_sample_support.grid_type
_em_software
_em_image_recording.film_or_detector_model
_em_particle_selection.num_particles_selected
_em_3d_reconstruction.actual_pixel_size
_em_3d_reconstruction.num_particles
_em_3d_reconstruction.resolution_method



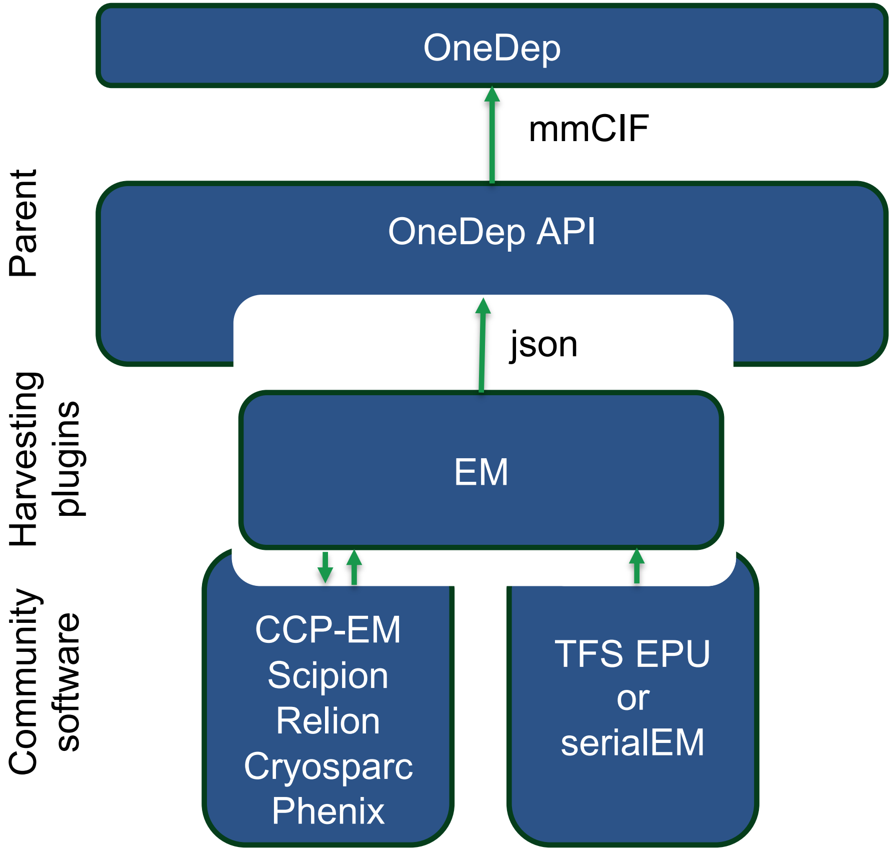
Archive

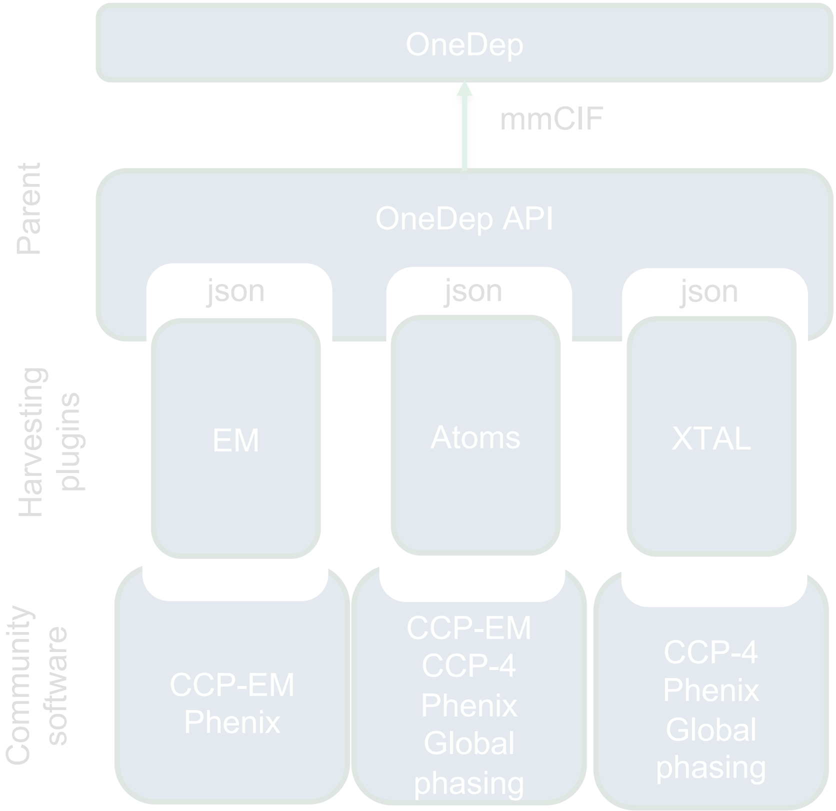
EMDB
Electron Microscopy Data Bank

OneDep

API

DepUI (www)


OneDep
mmCIF
Parent
OneDep API
json
Harvesting plugins
EM
Community software
CCP-EM
Scipion
Relion
Cryosparc
Phenix
TFS EPU
or
serialEM


OneDep
mmCIF
Parent
OneDep API
json
json
json
Harvesting plugins
EM
Atoms
XTAL
Community software
CCP-EM
Phenix
CCP-EM
CCP-4
Phenix
Global
phasing
CCP-4
Phenix
Global
phasing

Summary

- EMDB has data model in XML (historical) and in mmCIF (as part of wwPDB and OneDep)
- Both are documented online
- mmCIF can be used to pre-populate OneDep depositions

- API is operational on OneDep test server
- New end-points may need to be created but upload is working
- Upload using mmCIF to capture metadata

- We need your feedback!

Thank you for listening

https://www.ebi.ac.uk/emdb/

🐦 @EMDB_EMPIAR @kylelmorris

▶ @emdb_empiar

EMBL-EBI