EMPIAR: Quick tour

Andrii Iudin [1], Ardan Patwardhan [2]

- Structures
- Beginner
- 0.5 hour

This quick tour provides a brief introduction to the Electron Microscopy Pilot Image Archive (EMPIAR), a public resource for raw, 2D electron microscopy images.

Learning objectives:

- Search and access raw, 2D electron microscopy images in EMPIAR
- Deposit electron microscopy data in EMPIAR
- Know how to access further support when using EMPIAR

What is EMPIAR?

EMPIAR [3], the Electron Microscopy Public Image Archive, is a free public resource for raw, 2D electron microscopy (EM) images [4] (EM) images (1 [5]).

Why do we need EMPIAR?

EMPIAR provides a way to easily access state-of-the-art 2D image data that underpins 3D cryo-EM structures of biomacromolecules [6] and molecular machines. It complements the Electron Microscopy Data Bank (EMDB [7]), where corresponding 3D structures are stored, and PDB [8] which stores atomic models of macromolecular structures (Figure 1). EMPIAR data has been used to facilitate development and validation of methods, which will lead to better 3D structures.
The development of EMPIAR was prompted by calls from the electron microscopy community on the urgent need for archiving of raw image data related to EMDB structures. Notably, two workshops run by the EMBL-EBI and the Open Microscopy Environment [9] (OME) kickstarted the process (2 [10], 3 [11]) and in 2014 the EMBL-EBI received funding to develop the resource.

EMPIAR has been designed to handle very large datasets with sizes in the terabyte range. EMPIAR uses Aspera [12] software plugin to enable the safe transfer of large datasets. As of February 2017 EMPIAR contains 67 entries averaging ~700 GB in size, with nine entries exceeding 1 TB.

**What can I do with EMPIAR?**

**Who is EMPIAR for?**

EMPIAR can be used by anyone working in the electron microscopy field. For those new to the field, EMPIAR is a useful source of training data. Experts in the electron microscopy field can use EMPIAR data for validation, reprocessing (possibly with a different scientific emphasis than the original depositor) and methods development.

**What can I do with EMPIAR?**

EMPIAR can be used to:
Discover electron microscopy datasets by searching for the EMPIAR accession code, title, authors, associated publication’s Digital Object Identifier (DOI) and PubMed ID, the format of the submitted files, or related EMDB entries;
- View the image data;
- Download the image data;
- Archive your own electron microscopy image data.

Where does the data come from?

EMPIAR stores and provides access to data from cryo-electron microscopy experiments performed in laboratories around the world. In addition to EMDB-related raw image data, Serial Block-Face Scanning Electron Microscopy (SBF-SEM) and Focussed Ion Beam Scanning Electron Microscopy (FIB-SEM) can also be deposited in EMPIAR.

The data is stored along with information (metadata) about the biological samples and the equipment used. EMPIAR curators manually annotate the data to a high standard using the web-based EMPIAR annotation system that displays the file information as it is extracted with such tools as IMOD and BSOFT.

Searching EMPIAR

You can view all of the datasets in EMPIAR by clicking on the 'EMPIAR home' tab on the EMPIAR homepage as shown in Figure 3. Alternatively, you can search for datasets using keywords or accession codes by using the search box shown below (Figure 3).
**Searching EMPIAR.**

**Viewing search results in EMPIAR**

Search results are displayed in real time, immediately after you have started typing in the search box (Figure 4). You can sort the search results using the clickable header. To view an experiment click on the accession code for that experiment. Datasets can be downloaded directly from the search results page once the Apera plugin has been installed.
**Figure 4** Viewing search results in EMPIAR.

**Viewing datasets in EMPIAR**

Clicking on an accession [13] code from the main page provides you with more information about that dataset (Figure 5). Each dataset contains information about related publications, EMDB [16] and PDB [22] entries and contains one or more image sets.
Figure 5 Viewing datasets in EMPIAR.

Before downloading the data, it is possible to preview it by clicking on a file in the directory browser (Figure 6). The mouse cursor will change into an eye icon when hovering over a file with a thumbnail available.
Figure 6 Previews the data before downloading.

Downloading data from EMPIAR

You can download full datasets, imagesets or individual images from EMPIAR.

Links to download the data are provided on each page.

There are two main options for downloading EMPIAR data:

1. Smaller files can be downloaded as ZIP files directly from the EMPIAR website.
2. For large datasets we recommend using Aspera connect. This is a small plugin that allows you to resume downloads if they are interrupted and checks that the datasets are complete once the download is finished.

Aspera can be blocked by firewalls at some institutes. Other options for downloading data from EMPIAR include Globus [23] and our FTP [24] site. Have a look at our FAQ [25] page for more information.

Depositing data in EMPIAR

You can deposit your data in EMPIAR using the EMPIAR deposition system. This tool can be accessed from the EMPIAR homepage (Figure 7).

Your data remains private until you tell us to release it.
When you deposit your data in EMPIAR you will need to fill in a form to capture metadata [26] about the image/dataset. An example submission form is shown in Figure 8.

The image data is uploaded using the Aspera plugin. This is a simple and secure way of making electron microscopy data and the associated metadata publicly available.
Other options for uploading your data include the Aspera command line and Globus [23]. If neither of these are suitable for you, please empdephp [at] ebi.ac.uk (contact us) and we can discuss the best way to submit your data to EMPIAR.

Help with depositing data in EMPIAR

Once signed in, you will find a link to our helpdesk on every page in the EMPIAR deposition system. You will also find a direct link for asking questions about a specific deposition (Figure 9).

Figure 9 Help with depositions in EMPIAR.

See the EMPIAR deposition manual [27] for more information on how to deposit data in EMPIAR.

Your feedback

Please tell us what you thought about this quick tour. Your feedback is invaluable and helps us to improve our courses and enhance your learning experience.
Get help and support on EMPIAR

Support

For help using EMPIAR to access, download and deposit 2D electron microscopy data please see:

- FAQ page [28]
- Deposition manual [27]
- empdep-help [at] ebi.ac.uk (Email us)

References


Contributors

Andrii Iudin [1]

EMBL-EBI
Scientific Programmer: Cellular Structure and 3D Bioimaging (CS3DB) team

In 2014 Andrii Iudin joined the Cellular Structure and 3D Bioimaging team of Molecular and Cellular Structure (MCS) cluster at EMBL-EBI to work on EMPIAR. Andrii is responsible for maintaining, developing and extending the EMPIAR entry pages, the API and the deposition and annotation system.
EMPIAR: Quick tour
Published on EMBL-EBI Train online (https://www.ebi.ac.uk/training/online)

Ardan Patwardhan [2]
EMBL-EBI
Project leader EMDB - Kleywegt team: Protein Data Bank in Europe

Molecular and Cellular Structure cluster

Molecular and Cellular Structure cluster aims to bring structure to biology and to biologists by making this complex field more accessible to non-specialists. The teams in the MCS cluster are involved in managing three of the major archives in structural biology: the Protein Data Bank (PDB), the Electron Microscopy Data Bank (EMDB) and the Electron Microscopy Public Image Archive (EMPIAR).

Funding

The work on EMPIAR is funded by a project grant awarded to PDBe by MRC and BBSRC (MOL2CELL; grant MR/L007835). It further benefits from funding from the Wellcome Trust and EMBL-EBI to PDBe.

Source URL: https://www.ebi.ac.uk/training/online/course/empiar-quick-tour

Links
[1] https://www.ebi.ac.uk/training/online/trainers/andrii_10648
[2] https://www.ebi.ac.uk/training/online/trainers/ardan
[3] https://www.ebi.ac.uk/pdbe/emdb/empiar/
[4] https://www.ebi.ac.uk/training/online/glossary/electron-microscopy
[5] https://www.ebi.ac.uk/training/online/course/empiar-quick-tour/get-help-and-support-empiar/#empiar_ref_1
[6] https://www.ebi.ac.uk/training/online/glossary/biomacromolecules
[8] https://www.ebi.ac.uk/pdbe/
[9] https://www.ebi.ac.uk/training/online/glossary/open-microscopy-environment
[10] https://www.ebi.ac.uk/training/online/course/empiar-quick-tour/get-help-and-support-empiar/#empiar_ref_2
[12] https://www.ebi.ac.uk/training/online/glossary/aspera
[13] https://www.ebi.ac.uk/training/online/glossary/accession
[14] https://www.ebi.ac.uk/training/online/glossary/digital-object-identifier
[15] https://www.ebi.ac.uk/training/online/glossary/pubmed
[16] https://www.ebi.ac.uk/training/online/glossary/emdb
[18] https://en.wikipedia.org/wiki/Focused_ion_beam
[19] https://www.ebi.ac.uk/training/online/glossary/annotation
[22] https://www.ebi.ac.uk/training/online/glossary/pdbe
[23] https://www.ebi.ac.uk/training/online/glossary/globus
[24] https://www.ebi.ac.uk/training/online/glossary/ftp
[25] https://www.ebi.ac.uk/pdbe/emdb/empiar/faq#question_Select
[26] https://www.ebi.ac.uk/training/online/glossary/metadata
[27] https://www.ebi.ac.uk/pdbe/emdb/empiar/deposition/manual/
[28] https://www.ebi.ac.uk/pdbe/emdb/empiar/faq