Complex Portal: Quick tour

This quick tour provides a brief introduction to EMBL-EBI's Complex Portal: a manually curated, encyclopedic resource of macromolecular complexes from a number of key model organisms. Learning objectives:

- Basic understanding of the Complex Portal and how it can help you to find macromolecular complexes of interest
- Know how to submit curation requests for complexes
- Know where to find out more about the Complex Portal

What is the Complex Portal?

The Complex Portal is a manually curated, encyclopedic resource of macromolecular complexes from a number of key model organisms, entered into the IntAct [2] molecular interaction database. Data include protein-only complexes, as well as protein—small molecule and protein—nucleic acid complexes. All complexes are either derived from:
1) physical molecular interaction evidence extracted from the literature

or

2) inferred by curators: either from information on orthologues in closely-related species or from background biological knowledge based on other evidence (for example, pharmacological studies).

Any complex which currently lacks direct experimental evidence is tagged as such, using controlled vocabulary [3] terms from the Evidence Code Ontology [4]. Content is added by expert curators and direct user requests are prioritised.

Why do we need the Complex Portal?

The Complex Portal provides annotations for the complex (as a biological entity), along with descriptions of their function and other properties. It differs from UniProt [5] because the latter only provides annotations for the protein subunits.

The portal contains an interactive viewer with details of stoichiometry [6] and binding regions. It provides cross-references to the complex from many other databases, such as The PDB [7], Reactome [8] and ChEMBL [9]. The portal also contains annotations to all three classes of The Gene Ontology [10], EC numbers [11], diseases and literature evidence.

What can I do with the Complex Portal?

With the Complex Portal you can:

- search for complexes by name, synonym [12], identifier [13], species, cross-reference [14] identifier;
- view complex details in tabulated and schematic forms;
- download complexes from our FTP [15] site;

Searching and visualising data from the Complex Portal

Searching for complexes

You can search complexes using:

- any name(s) or identifier [13](s) of participants in the complex (for example, using ChEBI [17], RNACentral or UniProt [5] IDs);
- any name(s) or identifier(s) of the complex itself, including identifiers from external databases such as ChEMBL [9], The Gene Ontology [10], The PDB [7] or Reactome [8];
- the species.

A search can be performed in either the big box on the homepage, the small box on top of every other page, or in the URL [1]. The search is ‘exact’ but allows the use of Boolean [18] operators [19] and the wildcard (*).
To see how the search results will be presented (and what filters are available), you can click on the Examples listed on the right-hand-side of the screen [2].

For more detailed searches using the "Complex Query Language" go to the Help [20] menu [3]. We also perform a search across all EMBL-EBI resources on your search terms [4].
Figure 1 Complex Search - showing the Complex Portal homepage and the search results page.

Search Results

Results are displayed as 10 hits per page.

Faceted filters are available in the left-hand margin, which allow you to search within the results [5].

Full EMBL-EBI search results are located in the drop-down table [4].
Details View

Each complex is extensively annotated:

- Each complex has a unique identifier [13], in the format: EBI-xxxxxxx (where x = {1-9}).
- The recommended name is the name most commonly used in the literature - all alternatives are captured as synonyms.
- We provide a systematic name, which is made up of concatenated gene names - these are based on the Reactome [8] naming rules.*
- Each complex is annotated with its function and, where applicable, further physical and biological properties.
- Diseases, ligands and assembly terms are given, if applicable.
- All participants of the complex are listed in a table, listing their unique identifiers (that link out to the relevant databases), names, descriptions, stoichiometry [6], molecule type and biological role (if applicable).
- Cross references to the experimental or inferred evidence, the occurrence of the complex in another database, gene ontology [10] and disease annotation [21] are captured extensively.
- The complex is visualised using an interactive viewer (see next page).
Visualising complexes

The Complex Viewer depicts the participants in a complex as nodes and known binding regions as edges. The legend in the viewer shows the type of complex participant. If no specific binding sites are known, either no edges are displayed or the edges link molecules by their *n-terminus* [23] (in the 'expanded view', see Figure 4).

The complex viewer is **interactive**:

- Click and drag a node to change the layout [1]. Click <reset> to bring it all back together.
- Click on a node to expand proteins to display binding regions in the sequence [2]. Hover over the *c-terminus* [24] to swivel the protein bars around.
- Press <shift> and click to expand the protein sequence to show the residue level [3].
- If there are many instances of the same molecule, the stoichiometry [6] for the whole complex will not be expanded, but it will be indicated in square brackets [4].

**Feature views:**

1. **MI Features**: provides detailed binding site [25] information. Hover your mouse over the binding edge to see binding site ranges [2].
2. **UniprotKB** [26]: displays all UniProt [5] features. Hover your mouse over the sequence for...
more detail [5].

3. SuperFamily [27]: Hover your mouse over to reveal domain [28] information [6].
4. Interactor: shows unique interactors in the same colour [7].

The graphic can be exported as a .svg file.
Figure 4 Complex Viewer.

**Getting data from the Complex Portal**

**Downloading complexes**
You can download complexes from the [Complex Portal FTP site](http://www.ebi.ac.uk/training/online) [29].

They are currently organised in folders for each species and are available in PSI-MI XML2.5 format.

**Programmatic access**
We have an [API for the Complex Portal](http://www.ebi.ac.uk/training/online) [30] with two methods:

/search/

/details/

**Contributing to the Complex Portal**
To make a request for us to curate a particular complex, please fill out our contact form and supply as much detail as possible, including:

- PubMed Identifier (PMID) that provides evidence for the existence of the complex and its function;
- list of the complex participants;
- suggested name and synonyms;
- suggested definition/functional annotation properties.

Your feedback

Please tell us what you thought about this Quick tour. Your feedback is invaluable and helps us to improve our courses and thus enhance your learning experience.

Get help and support on the Complex Portal

Support

- We have extensive documentation and help pages.
- For further queries and requests please contact our helpdesk.

References


Funding

- European Molecular Biology Laboratories Core Funding.
- Biotechnology and Biological Sciences Research Council [BB/L024179/1].
- National Institutes of Health [268201000035C].
- Wellcome Trust [WT101477MA]

Contributors
Birgit Meldal [1]

EMBL-EBI
Scientific Database Curator (Molecular Interaction & Complexes) - Orchard Team: Molecular Interactions

Birgit Meldal joined EMBL-EBI in 2012 as a Scientific Curator for the molecular interactions database IntAct. Her work focuses on curating macromolecular complexes into the Complex Portal, and she also contributes to the Gene Ontology. Birgit is a trainer for the IntAct and Reactome databases. She gained her PhD from the University of Southampton (in collaboration with the Natural History Museum, London) in 2004, having built the first comprehensive molecular phylogeny of the phylum Nematoda. She followed this with postdoctoral positions at the University of Cambridge working on the function of the breast cancer gene EMSY and on molecular epidemiology and host-virus interactions of transfusion-transmissible, blood-borne viruses (Hepatitis B and Hepatitis E virus).

Source URL: http://www.ebi.ac.uk/training/online/course/complex-portal-quick-tour

Links
[1] http://www.ebi.ac.uk/training/online/trainers/bmeldal
[7] http://www.ebi.ac.uk/training/online/glossary/pdb
[8] http://www.ebi.ac.uk/training/online/glossary/reactome
[9] http://www.ebi.ac.uk/training/online/glossary/chembl
[12] http://www.ebi.ac.uk/training/online/glossary/synonym
[16] http://www.ebi.ac.uk/training/online/glossary/curation
[17] http://www.ebi.ac.uk/training/online/glossary/chebi
[18] http://www.ebi.ac.uk/training/online/glossary/boolean
[19] http://www.ebi.ac.uk/training/online/glossary/operators
[21] http://www.ebi.ac.uk/training/online/glossary/annotation
[23] http://www.ebi.ac.uk/training/online/glossary/n-terminus
[25] http://www.ebi.ac.uk/training/online/glossary/binding-site
[26] http://www.ebi.ac.uk/training/online/glossary/uniprotkb
[27] http://www.ebi.ac.uk/training/online/glossary/superfamily
[28] http://www.ebi.ac.uk/training/online/glossary/domain
[31] http://www.ebi.ac.uk/support/intact
[32] http://www.ebi.ac.uk/training/online/glossary/pubmed
[33] http://www.ebi.ac.uk/training/online/glossary/pmid
[34] http://www.ebi.ac.uk/intact/complex/documentation/