**How To Find Structures in the PDB Related to a Topic of Interest**

**Learning Objectives**:

1. Identify PDB entries that are related to a topic of interest
2. Identify one or more examples of PDB entries that may be used to visualize/illustrate a molecular view of the topic of interest



**Steps**:

1. Getting Started

* Get overview of topic, key participants and roles from textbooks, review articles, and online resources
* Gather name(s) of protein(s) and function(s) in relation to the topic of interest. Where possible, gather sequence, domain(s), organization, and interaction information for proteins to be explored at a molecular level

1. PDB Query

* Search for structure(s) in the PDB using one or more of the following, separately or in combination.
  + Contents of PDB entry:
    - Text search: Name of protein and/or ligand, name of authors of the structure(s)
    - Sequence of protein of interest or its close relative
    - Ligand information: e.g. formula, chemical description (SMILEs/InChI)
  + Annotations of PDB entry:
    - Molecule of the Month features related to the topic of interest may provide background and PBD identifiers
    - Various classifications based on bio process, cellular location, function; Enzyme Classifications (E.C.), source etc.
  + Related entries:
    - By a specific criteria to launch a query by example: e.g. by protein name/UniProt ID, author name, E.C. number
  + Links to other resources:
    - Identifiers linking PDB entries to UniProt, PubMed, DrugBank or other databases and resources
  + Literature:
    - Based on keywords used in published literature/abstracts, PubMed IDs

1. Sorting/Refining Query Results

* Make reports and sort by various criteria (e.g. deposition/release date, resolution, etc.)
* By visual inspection, select from a collage of PDB structures in the query results
* Refine results using additional information/conditions

1. Selecting PDB entries that meet specific criteria
   * For a multi-domain protein, make sure selected structure(s) include coordinates of relevant domains/regions. Review Protein View page to see which domains are included in the PDB entry. Read the associated literature (primary research articles accompany most PDB entries) and related articles from PubMed and the PDB to learn more
   * Make sure desired mutations and/or ligands/drugs etc. are included in the structure 🡪 check sequence, mutations list, ligands.
   * Make sure structure quality meets your needs (review resolution, experimental details and validation report)
2. What if …
   1. There are too many results?
      * Use sorting and refining strategies to narrow down results list to only include structures relevant to the topic/story (see above)
   2. There are no results or the results are not relevant?
      * Try using other query options – e.g. synonym of protein name, sequence(s), other known details about structure
      * Look for structures of the same protein/domain from another organism (note selection criteria listed above)
3. Exploring Sequence and Structural Neighbors
   * Explore all PDB structures related to a UniProt ID (Access in Protein View, or by searching for UniProt ID). Note: For a multi-domain protein, check the domains/regions included in the PDB entry
   * Explore sequences and 3-D structure clusters of a relevant PDB entry from the structure summary page
   * Based on your interest, choose relevant PDB entries using selection criteria listed above for further exploration

At the end of the above steps you should have identified PDB entries that are relevant to your topic of interest and help tell a story about it.