Teaching enzymology with the Protein Data Bank: From Pandemic to Paxlovid

Exploring the SARS-CoV-2 Main Protease

Presenter: Dr. Paul Craig, RIT, NY.

Open the Structure Summary Page (SSP) for COVID-19 Main Protease PDB ID 6LU7

Learning Objective: To introduce or reinforce the use of a number of RCSB PDB resources for teaching protein structure and function. This handout has all of the links that you will need for the exploration that will be presented.

1. Review the elements of the COVID-19 Main Protease structure (6LU7) SSP
   a. SSP
   b. 3D view
   c. Annotations
   d. Experiment
   e. Sequence (including 3D view)
   f. Genome
   g. Versions

Poll Question 1: Which of the following elements on the COVID-19 Main Protease Structure Summary page do you plan to explore further?

2. Review the 6LU7 SSP
   a. Assemblies and 3D view options. Including 1D-3D view
   b. wwPDB validation for structure quality
   c. Literature/PubMed
   d. Macromolecules, including EC Classifications
      i. Mutations
      ii. Connection to UniProt
      iii. Entity Groups
      iv. Protein Feature View/Expand
   e. Biologically Interesting molecules (more on this later)

3. 3D view links with Mol*

a. Still on the SSP, looking at the left hand box
b. Expanded view provides access to tools
c. Structure takes us to Mol*. Click on the ligand. Show hydrogen bonds plus information
d. 1D-3D view - look for active site residue and interaction with the inhibitor
e. Electron density - click on a residue to see
f. Validation report provides information on clashes in Mol*
g. Global Symmetry 3D-View brings up the symmetry browser with symmetry axes
h. Find similar assemblies. Review refinements
i. Finish with Documentation for Mol*

Poll question 2: Which features of Mol* have you explored before?

4. Annotations - a rich source for developing lessons or for having your students prepare molecular stories
   a. ECOD
   b. Enzyme Classification browse and search
   c. MeSH - Medical Subject Headings
   d. Molecular Function
   e. Symmetry
   f. ATC - Anatomical Therapeutic Chemical Classification System

Poll question 3: Which annotations do you think would give your students the best opportunities to learn more about protein structure and function?

5. Ligand Interaction - Biological Interesting Molecule
   j. 2D diagram
   k. Ligand Interaction
      - Hide the protein chains
      - Hide the waters
      - Right-click and drag to zoom in on one ligand and its surrounding residues
      - Click on the … next to [Focus] Surroundings (5…)
      - Hide and show Ball & Stick Representation
      - Hide and show Non-covalent Interactions Representation
      - Zoom in and hover over a non-covalent interaction to learn more details

Poll question 4: As we explored the ligand interaction, which of the following tasks were you able to complete?

6. 1D-3D View
   a. Identify the active site residues
   b. Explore interactions

c. Find the covalent ligand

Chat Q: Name the ligand

Poll question 5: What features of the 1D-3D View screen are you most likely to use in teaching?

From Pandemic to Paxlovid

Presenter: Dr. Shuchismita Dutta, Rutgers University, NJ.

Part 1: Learn

Q1. Answer the following about the PF-00835231 molecule based on reading this passage.

a. What was this compound designed to do?
b. How was this compound administered? (orally, intravenously etc.?)
c. Why was this compound not used till 2020?

Q3. Search the PDB (www.rcsb.org) for the compound PF-00835231.
(Hint: Start typing the drug name PF-00835231 and run the search. Review the results returned to identify structures that include this molecule.)

a. List the PDB identifier(s) and structure titles of two different PDB entries that you would want to explore further.
b. Open the structures you have listed above and examine the structure summary pages to figure out which of the small molecule identifiers corresponds to the inhibitor compound.

Chat Q: What is the ligand ID of the compound PF-00835231?

Reflection: Are there other ways of identifying the Ligand ID of the drug molecule?

Q4. Search for all structures that include this compound.
(Hint: on the ligand summary page for the ligand that you have identified, click on the number of hyperlinked entries in the section for “Find entries here:” How many PDB structures include this compound? List the PDB IDs and structure titles here.

a. Examine the “Scientific Name of Source Organism” of the entries in the search results. List them below.

Q6. Explore the interactions and save an image of how the inhibitor binds to the SARS-CoV and SARS-CoV-2 Main protease molecules.
(Hint: Open the structure summary pages for a specific PDB entry >> scroll down to the small molecules section >> click on the Ligand Interaction box on the right hand column of the table.)

a. Save the images. Label one covalent and one non-covalent interaction between the inhibitor and the protease.

Q7. Compare the overall structures of the proteases from SARS-CoV and SARS-CoV-2 using the tool https://www.rcsb.org/alignment.
(Hint: use the PDB IDs of the structures used in the answer above for this comparison)

a. From this analysis, would you use the molecule PF-00835231 to develop an oral inhibitor for treating SARS-CoV-2? Justify your answer.

Part 2: Assess Learning

Begin by querying for the compound with the name PF-07321332 (or Paxlovid's key ingredient - Nirmatrelvir).

Q8. Use the top search box (on www.rcsb.org) to query for the compound name PF-07321332 in the PDB.
(Hint: After typing the name in the search box, scroll through and review the autocomplete options presented to see this name in Chemical Synonym. Select this and run the search using default options. Examine the titles of structures in the search results).

a. How many structures include reference to the compound name (PF-07321332)?
b. What is the ligand identifier for the compound PF-07321332?

(Hint: look through the list of small molecules in each of the entries in the results list and find the common ligand).

Chat Q: What is the ligand ID of the compound PF-07321332?

Reflection: Are there patterns and trends in the binding of this drug molecule? What can we learn from that about this inhibitor/drug?

Q9. For the PDB structures that include the compound PF-07321332, what are the organisms from which the main protease was derived?
(Hint: Examine the "Scientific Name of Source Organism" listed in the "Refinements" options in the left hand column.)

Q10. What can you conclude about the ability of compound PF-07321332 (Paxlovid) as a treatment for different coronavirus infections?

Part 3: Challenge Questions

Q11. Review the search results of Answer 8a to identify any two SARS-CoV-2 mutants with Paxlovid bound.

a. List the title of two mutant structures selected for review.

(Hint: Review the results - list the PDB IDs and titles of two SARS-CoV-2 mutants in this list)

b. How do the mutations impact Paxlovid binding to the protease?

[Hint: For each of the PDB entries identified, use the Explore Sequence Annotations in 3D view to visualize the location of the mutations with respect to the drug (Paxlovid) binding.]

Q12. Can you identify at least one protease mutant where the Paxlovid mutant may not be effective as a therapeutic agent?
Hint: For the set of structures that have PF-07321332 bound (Answer 8a). Select all the SARS-CoV-2 protease structures and group them

- Refine results by source organism "Severe acute respiratory syndrome coronavirus 2".
- Scroll up to the "Advanced Search Query Builder" section on the top of the page to change the
    - Return options from "Structure" to "Polymer entities",

    - “group by option” to “UniProt Accession” and
    - “display as” option to “Groups” before re-running the search
- Select the larger group of results to answer the following questions.

Polling Q 6: Why do you think that the polymer sequences shown line up with only a small part of the complete UniProt sequence?

Reflection: Are these the only SARS-CoV-2 Main Protease structures in the PDB?

a. Identify any PDB structures of SARS-CoV-2 Main Protease with mutations in/near the enzyme active site. List them.
b. Explore the impact of this mutation on Paxlovid binding.