1. What is the biological function of this protein? From what species was this protein taken? (Hint: check [pdb-101 page](https://pdb101.rcsb.org/motm/277))

The function of this protein is to break down nylon plastics. They were taken from bacteria, specifically flavobacterium sp. KI723T1.

1. What is the reference paper for this structure? Briefly describe the findings of the paper.

The reference paper is “Structural basis of the correct subunit assembly, aggregation, and intracellular degradation of nylon hydrolase” by Negoro et al. The study found that NyIC (an enzyme that breaks down nylon) is initially expressed as an inactive precursor that is then cleaved to generate an active enzyme compound of two subunits that assemble into a doughnut shaped quaternary structure. Then, the researchers altered the thermal stability of NyIC by substituting certain amino acids and analyzed the effects on the assembly of the subunits. The paper proposes a model that predicts whether the protein will assemble correctly or not.

1. How many protein chains form the structure? How many amino acid residues are in each chain?

There are 2 chains that form the structure.

Alpha subunit – 355 residues

Beta - 355 residues

1. How many ligands (small molecules) does this structure have? Name the ligand(s).

There are 3 ligands

1. Phosphate Ion
2. Glycerol
3. Chloride Ion
4. What was the experimental method used to obtain the 5Y0M structure? What is the “resolution” of this structure? What is the meaning of the “resolution”? What protein structure has higher quality: One with “high resolution” or “low resolution”? Hint: [check](http://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/resolution)

The experimental method was X-ray diffraction/crystallography. The resolution was 1.03 Angstroms. Resolution is a measure of the quality of the data that has been collected on the crystal containing the protein. If the crystal is perfect, all the proteins will scatter X-rays the same way. Lower resolutions allow us to see very fine details (i.e. every atom in the electron density map). Protein structures with low resolution are higher quality.

1. Can you think of a reason (there are multiple valid answers) as to why the resolution-value increases as protein size increases? (Do not forget that the smaller the resolution-value, the better the power of resolving two atoms from each other).

Larger proteins may form a complex structure and have many different conformations so it becomes more difficult to analyze with X-ray crystallography (resolution increases)

1. You can download the PDB file for proteins by clicking the “Download Files” button (next to the name of the protein on the **Structure Summary** tab page) and then selecting the appropriate format. After the long text (called PDB header), you should see several columns of text and numbers. Describe the information stored in columns 1 to 9 (for example column one contains the word atom, column two is the atom type, etc.; 9 points). Also, [check](https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/introduction).
2. The word “atom”
3. Unique number assigned to each atom in structure
4. Atom name (“N”, “CA”, etc.)
5. Residue Name
6. Chain Identifier
7. Residue sequence number
8. Atom x coordinate
9. Atom y coordinate
10. Atom z coordinate
11. What is a "protease" protein? What is its general function?

A protease protein is an enzyme that breaks down proteins by cleaving peptide bonds between amino acids. Its general function is to regulate protein activity and turnover, which is essential for many processes such as cell signaling and gene expression.

1. What is a “serine protease”? What common feature dose the 7vve structure (an engineered plastic eating enzyme) has with serine proteases?

A serine protease is a type of protease that uses a serine residue at its active sit to cleave peptide bonds. The 7vve structure has a similar active site structure to serine proteases, with a serine residue playing a crucial role in the enzyme’s ability to break down PET plastics.

**Q1.** Fill in the blanks in the following table by finding the residue names of the next three residues. Note that the R-group may be a hydrogen atom, as in glycine.

|  |  |
| --- | --- |
| **Residue Number** | **Amino Acid** |
| 1 | SER |
| 2 | GLY |
| 3 | PHE |
| 4 | ARG |

**Q2.** Use the Sequence Viewer to find the residue numbers and names of the last five amino acids in the COVID protease (NOT the inhibitor).

|  |  |
| --- | --- |
| **Residue Number** | **Amino Acid** |
| 302 | GLY |
| 303 | VAL |
| 304 | THR |
| 305 | PHE |
| 306 | GLN |

**Q3.** The first non-coil secondary structure is an alpha helix.

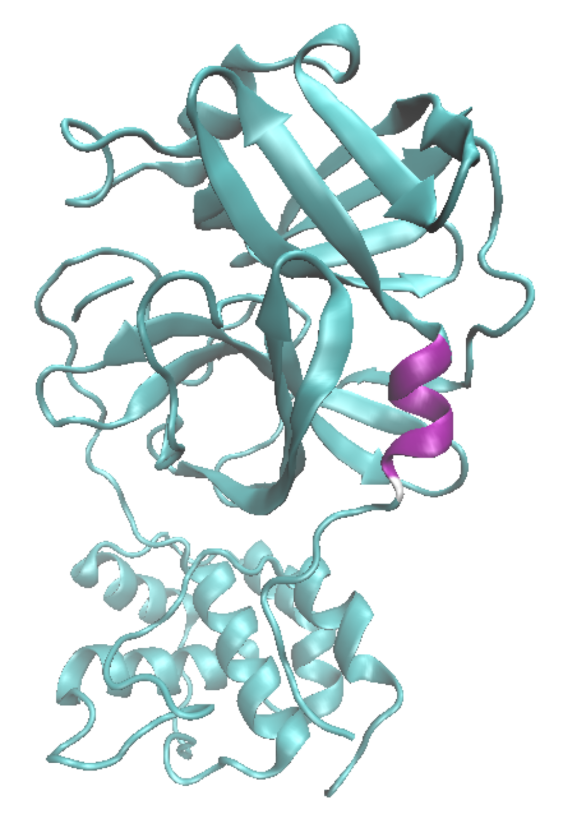
1. Which residue number does the helix start at, and what amino acid is it?

The first alpha helix starts at residue number 10 (SER amino acid)

1. What is the ending residue number, and what amino acid is that residue?

The first alpha helix ends at residue number 15 (GLY amino acid)

1. Select the alpha helix in the Sequence Viewer. Amend the representation of this alpha helix by using “**Secondary Structure**” as the coloring method and “**NewCartoon**” as the drawing method. Include a screenshot of this alpha helix along with the rest of the protease protein.



**Q4.** Try selecting protein residues 19, 58, and 155 at once using the Sequence Viewer. Find the representation created by the Sequence Viewer for these residues.

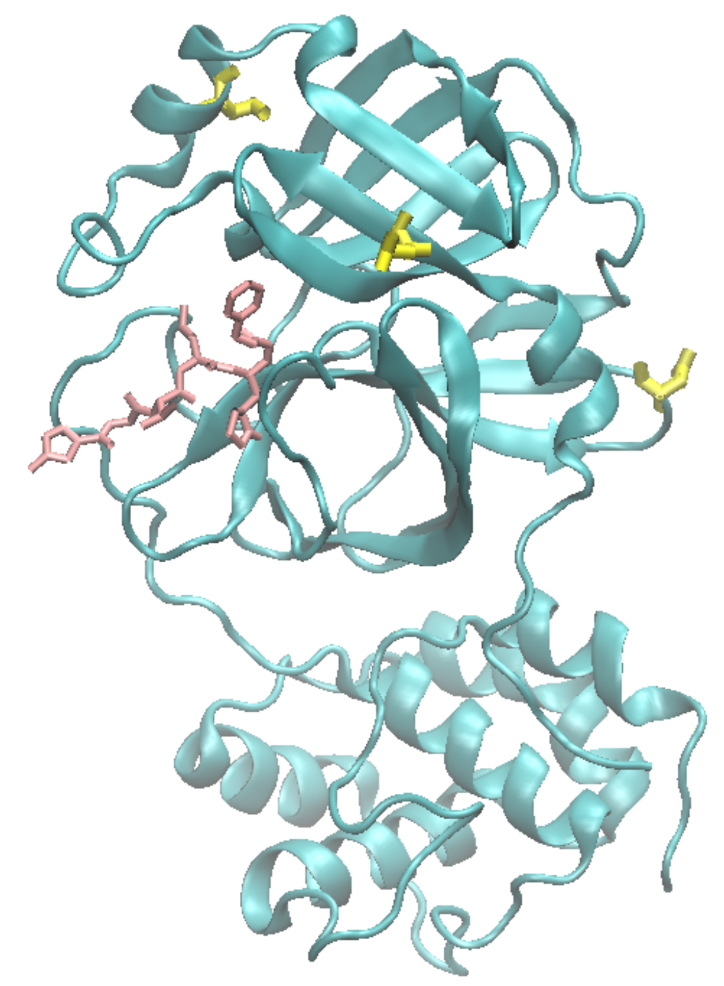
1. What is typed into the “**Selected Atoms**” box to give you these residues?

(chain A and resid 19 58 155)

1. What is the drawing method that the Sequence Viewer is using?

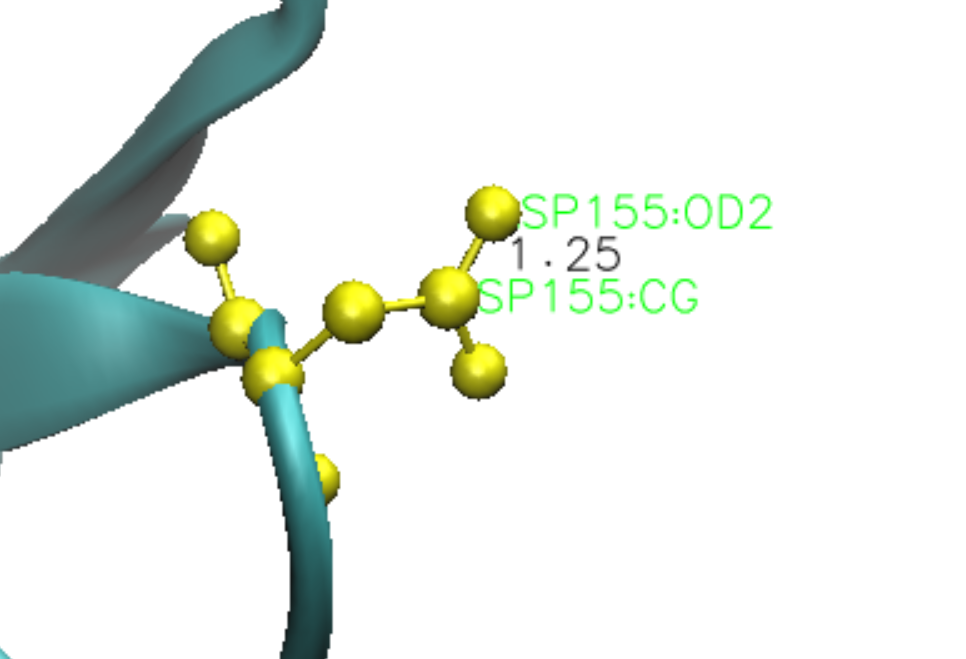
The drawing method it defaults to is Bonds

1. Include a screenshot of the *protein* shown as **NewCartoon** with these selected residues shown clearly. Note that, as mentioned before, there is also the *N3 inhibitor* present. Therefore, in this same screenshot, the N3 inhibitor should separately have the style DynamicBonds. Also, set the inhibitor’s coloring method to ColorID, and then pick a color that contrasts clearly against the protease.



**Q5.** Measure the distance between any two atoms

1. Insert a picture of the measurement that clearly shows the distance. You might need to modify the color for this label. Refer to prior instructions on how to do it. We recommend using Black as your new color.



1. Drawing on your previous knowledge, what unit do you think these measurements are in? If you are not sure, this information might be found by looking at the PDB page for 6LU7 (think about how the structure was obtained).

I think that the measurements are in Angstroms

**Q6.**  Again, find the first Serine residue on the molecule (**type protein and "resid 1 2" in the VMD selection window; then push the equal sign to center these residues in your screen; remember to hide any other representations to reduce clutter**). From this residue, follow the molecule with your mouse (from N-terminal to C-terminal) until you reach the atom to which an oxygen molecule is bonded. This is the carbonyl carbon, and the preceding carbon is the alpha carbon. Note that in the labels, the carbonyl carbon has “C” after the colon, while the alpha carbon has “CA”. You may find the “lines” or “CPK” drawing methods helpful. Measure the following bond distances. Please include the units in the answer. You don’t need to take a picture.

1. Amide N to alpha C:

1.46 A

1. Carbonyl C to alpha C:

1.53 A

1. Carbonyl C to N in the peptide bond:

1.33 A

**Q7.** Repeat this exercise for the same bonds in residue 33. Again, no picture required. (Hint: just show residue 32, 33, and 34 in the VMD screen.)

1. Amide N to alpha C:
   1. 1.46
2. Carbonyl C to alpha C:
   1. 1.52
3. Carbonyl C to N in the peptide bond
   1. 1.33

**Q8.** Of the three types of bonds we measured, which type of bond appears to be the shortest? Can you think of the reason why this might be the case?

The carbonyl C to N in the peptide bond seems to be the shortest. This may be because of the resonance in the peptide bond leading it to have a partial double bond character.

**Q9.** Describe the function of the COVID-19 main protease and briefly explain the purpose of the synthetic N3 inhibitor. (Hint: check the PDB101 page; you can google the answer!).

The COVID-19 main protease cleaves the viral polyproteins at specific sites resulting in the formation of smaller viral proteins that are required for viral replication and assembly. The purpose of the N3 inhibitor is to bind to the active site of the main protease inhibiting the protease’s activity and preventing the virus from replicating.

**Q10.** Go to the PDB page for the 6LU7 protein, at left corner of the window you see a protein structure; the first structure displayed is called "Biological Assembly 1". If you click on the arrow above this picture, you will see another picture called "Asymmetric Unit". What is the difference between the Biological Assembly and the Asymmetric Unit?

Hint, [check](http://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/biological-assemblies)

The asymmetric unit is the smallest portion of a crystal structure to which symmetry operations can be applied in order to generate the complete unit cell. An asymmetric unit may contain one biological assembly, a portion of a biological assembly, or multiple biological assemblies.

**Q11:** Describe one industrial application of Pepsin (Hint: you can use Google).

One industrial application of pepsin is in the food industry for the production of protein hydrolysates. These protein hydrolysates are used as flavor enhancers and nutritional supplements.

**Q12.**

1. How many hydrogen bonds are there in the largest alpha helix?

There are 6 H Bonds

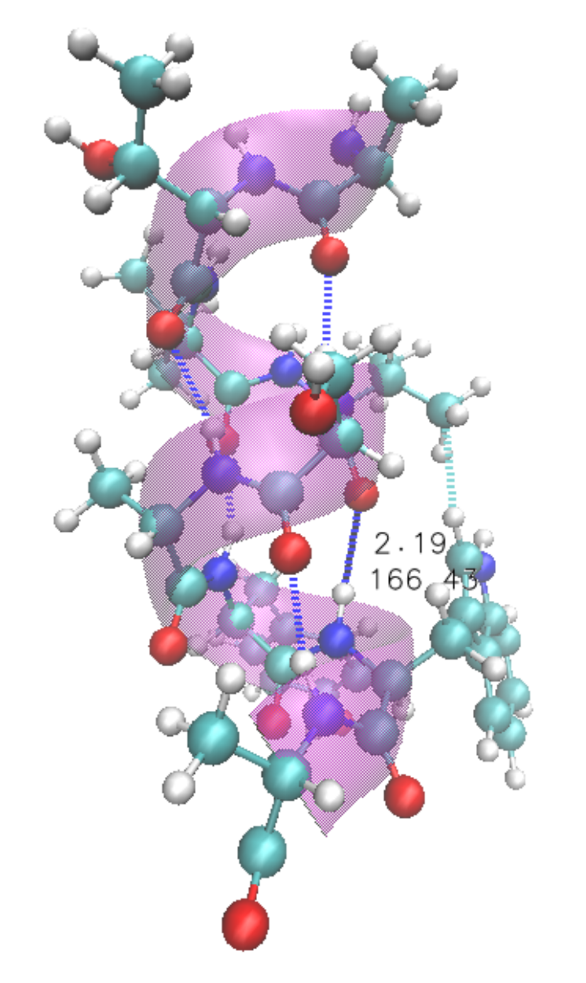
1. What sort of pattern exists between the residue numbers at the ends of each hydrogen bond (i.e., residue i has an H-bond with residue i+k)?

i -> i+4

1. Measure the distance and angle for one of these hydrogen bonds and attach a clear picture of your helix with its hydrogen bonding network.

Distance: 2.19 A

Angle: 166.43 degrees



**Q13.**

1. Do the side chains on an α-helix point into the center of the helix, or to the outside?

The side chains on an alpha-helix point to the outside of the helix

1. Why? (Try to show the sidechains in your helix representation)

This is because of steric hindrance. If the sidechains were on the outside of the helix, they would bump into each other

**Q14.**

1. How many cysteine residues are present in the molecule?

There are 2 cysteine residues in the molecule

1. What are their residue numbers?

Residue numbers 250 and 283

1. What unique structure do they form, and between which two atoms is the structure?

The unique structure they form are disulfide bonds, which are covalent bonds between two sulfur atoms in different cysteine residues.

1. What do you think is one function of these covalent cross-links in this molecule?

To add additional structure to hold secondary structures together. This disulfide bond specifically helps hold a beta sheet together.

1. Insert a picture highlighting these findings (a licorice drawing method might be helpful). For this picture, since we are studying cysteine residues in the context of beta sheets, you should take a snapshot that shows the cysteines belong to a beta sheet.

****

**Q15.** Identify the two catalytic aspartic acid residues (ASP32 and ASP215, corresponding to residues 32 and 215) in the molecule, and paste a **picture here that clearly shows the enzyme, the** **substrate, and the position of the active site residues on the enzyme; draw arrows in your picture pointing to these parts.**

A recommended way to visualize the enzyme-inhibitor interaction is as follows:

1. Make sure that your display is in “orthographic” mode
2. In general, always make a different representation for each part and color them differently
3. Create a representation for Selected Atoms “Chain E” with Drawing Method “New Cartoon” and Coloring Method “Color ID.” For the color, make sure that “0, Blue” is selected.
4. Create a representation for Selected Atoms “Chain I” with Drawing Method “Licorice” and Coloring Method “Name”.
5. Create a representation for Selected Atoms “resid 32 215” with Drawing Method “Licorice” and Coloring Method “Color ID”. Choose a color that stands out.
6. Finally rotate the molecule so that you get a clear picture for all of these selections.

