BIOS 11141 **Lab-6**

**Name:**

Please insert all answers and figures into this Word document, and then save the completed assignment as “Lastname\_CompAssin8.docx”. When you have completed the lab, submit it as an attachment to Canvas.

**Objectives:**

1. An introduction to the Protein data bank (PDB)
2. An introduction to a software package called **Visual Molecular Dynamics**1 **(VMD).** We will use VMD to analyze basic three-dimensional characteristics of a protein.

**\*If you have a Windows computer**, you need to change directories in VMD in order to access any of the files VMD creates or saves. To do so, you should add a cd “C:/Path” command to the end of VMD’s resource script (look in the folder where VMD is installed). You may find the following commands useful. If you need further help, ask your TA for assistance.

pwd see the path of your current directory.

ls list the files and directories within your current directory.

cd change directory.

cd ..back up a step in your path.

You can download VMD from the following link: <https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD>

Also, check the VMD tutorial:

<http://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>

**Part 1: Protein Data Bank (PDB)**

Proteins are vital biological molecules, which consist of smaller molecules called amino acids or residues, connected by peptide bonds. The resulting chain of amino acids is non-functional until it “folds” into a three-dimensional arrangement called the protein’s tertiary structure. The structure of the protein determines how it performs its function in the cell. For a brief description of proteins, please [see](https://cdn.rcsb.org/pdb101/learn/resources/what-is-a-protein/what-is-a-protein.pdf).

The Protein Data Bank ([PDB](https://www.rcsb.org/)) archive contains information about experimentally determined structures of proteins, nucleic acids, and complex biological assemblies. Each structure is identified by a four-character code like “1BCC”, and contains important information about the protein, such as spatial coordinates about individual atoms in the molecule, experimental procedure to obtain the structure, etc.

[PDB 101 page contains a lot of material about proteins that I strongly encourage you to explore.](https://pdb101.rcsb.org/)

Navigate to the PDB website and enter the code **5Y0M** into the search bar of PDB. This will take you to the specific page for this molecule that contains lots of information about this protein. First you see the authors who obtained the structure and the abstract for their paper. Below this, you will see the “Molecular Description”. This molecule consists of multiple protein subunits, which are separate chains of amino acids (“chains”), and each one is made of a specific number of amino acid residues (“length”). Answer the following questions:

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))
2. What is the reference paper for this structure? Briefly describe the findings of the paper.
3. How many protein chains form the structure? How many amino acid residues are in each chain?
4. How many ligands (small molecules) does this structure have? Name the ligand(s).

X-ray crystallography is the chief method for determining the 3-dimensional structures of proteins: about 88% of all structures deposited in the PDB were determined (“resolved”) using this technique. This first protein structure using the technique was the oxygen carrying protein myoglobin (isolated from a sperm whale!), and was performed in the 1950s, **winning the 1962 Nobel Prize in Chemistry.**

X-ray crystallography involves creating a highly regular and pure crystal of the protein in question, and then subjecting it to an x-ray beam and measuring the diffraction of the beam, from which the positions of atoms in the molecule can be calculated. An introduction to the process is available [here](https://www.youtube.com/watch?v=gLsC4wlrR2A).

It is generally more difficult to resolve proteins at a higher level of resolution. The figure below shows the mean number of residues in the protein and total atoms in the protein at a few different resolutions.



1. 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)
2. 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).
3. 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).
4. What is a "protease" protein? What is its general function?
5. What is a “serine protease”? What common feature dose the 7vve structure (an engineered plastic eating enzyme) has with serine proteases?

**Part 2: Introduction to VMD**

The following section is an introduction to basic operations in VMD. You need to begin by downloading the required proteins from the [Protein Data Bank](https://www.rcsb.org/), or PDB. The main unit of information deposited to the PDB is a text file (imaginatively called a .**pdb** file), which contains basic information about the protein: its origin, sequence, the experimental parameters (like its resolution), **the atomic coordinates,** etc.

The bulk of a PDB file is taken up by the *structure of the protein* — a list of the protein’s atoms and their respective coordinates; these coordinates describe the three-dimensional arrangement of all the atoms in the molecules. These structures are typically obtained via X-ray crystallography or NMR spectroscopy techniques (we will learn about these techniques later in class).

For this assignment, we will first look at the crystal structure of the COVID-19 main protease. The particular structure we will be looking at is the one with a PDB ID of “**6LU7**.” Note that there are a handful of structures of the coronavirus main protease, each slightly different from the other (with varying imaging techniques, resolutions, etc.). As such, it is important to be precise when communicating about a structure within PDB.

You have two options for importing this structure into VMD. Feel free to choose either one of them:

**1)** Using your web browser, navigate to the protein’s PDB page, click **Download Files** in the top right, and then click **PDB Format**. A drop-down window will open; save this .pdb file somewhere you will be able to locate later in the lab. Then, launch the program VMD**. You should see** three windows: A display window with a rotating “VMD” icon, a window with menu bars, and a terminal window. Minimize the terminal window. Select the “VMD Main” window and click **File** → **New Molecule**. In the new dialog box, select browse and navigate to the folder where you saved the 6LU7 structure. Click on the **6lu7.pdb** file, select **Load**, and then exit from the window.This will load the 306-residue coronavirus protein *and* a synthetic “N3 inhibitor” into VMD, where the latter is a potential therapeutic agent for the disease. For more information, please see: <https://pdb101.rcsb.org/motm/242>

**2) In the easier method,** you may simply **open VMD** as described above**, select File** → **New Molecule**, and then type in the PDB accession code **6LU7** under **Filename** andclick **Load**. This will automatically retrieve the structure from the PDB for use in VMD. Note: this method currently only works for VMD version 1.9.4 or higher. You should be able to view your version at the top of the display window.

Select the Display window. You should see a wire-frame model of the COVID main protease. In the structure, **teal** corresponds to **carbon** atoms, **red** to **oxygen** atoms, **blue** to **nitrogen** atoms, and **yellow** to **sulfur** atoms. You can change the background color to see the protein representation more clearly. Go to **Graphics** → **Colors** → **Display** → **Background** → **8 White**. **All of the images you take should be done with a white background.**

You can use mouse movement to change the perspective with which the molecule is presented in the window. There are three mouth modes in VMD, you can press keys to change between these mouse modes.

**R (Rotate)**

This is the default mode you have on VMD. You are in the Rotate mode if you haven’t switched to other modes. Otherwise, pressing R would get you to this mode.

Shape

Description automatically generatedThe following table and figure summarize the three ways you can move your mouse under Rotate mode. Hold down the indicated mouse button and move the mouse in the prescribed fashion; you should see the molecule rotate following your movement.

|  |  |  |
| --- | --- | --- |
| **Button** | **Movement** | **Result** |
| Left | Left/Right | Rotate around the vertical axis |
| Left | Up/Down | Rotate around the horizontal axis |
| Right | Left/Right | Spin CW/CCW |

**T (Translate)**

After you press T, you can drag and move the origin of the molecule (thus you are translating the graphic on your screen).

**S (Scale)**

After you press S, you can zoom in and out on the molecule by clicking on the screen.

Moreover, you can zoom in and out by scrolling as well.

**= (Restore to the default view)**

When you are adjusting the view, you can press = to bring the molecule back to the default view. This would be extremely helpful if you accidentally move the molecule out of the window, and you cannot bring it back.

Please familiarize yourself with these kinds of mouse movements and try to get a better view of the molecule.

There are several different ways to view a protein in VMD. In the VMD main window, click **Graphics** → **Representations**. Ensure that your single representation (Lines Name all) is selected; it will be highlighted green if it is. Next, navigate to the “**Drawing Method**” dropdown box towards the bottom left of the Graphical Representations window, select “**VDW,**” and observe its output on the screen. This gives you the Van Der Waals model, which is the space-filling representation of the protein atoms. Also, try the methods “**Licorice**” (which is a useful representation of the atoms and their bonds) and “**New Cartoon**” (which will help you identify the presence of α-helices and β-sheets).

To get rid of any extra atoms that you do not want to include in your visualization, you can make use of the **Selected Atoms** box. For example, after making sure the representation is selected (highlighted green), try typing “**protein**” in the box; you will notice that you can now only see the protein. **Some other commonly used “Selected Atoms” entries are:**

* “resname XXX” – Replacing XXX with the three-letter abbreviation for the residue will display all residues of this type (all letters in upper case). *Example: resname MET will select all Methionine residues in the protein.*
* “resid <number>” or “resid <number1> to <number2>” – Displays the listed residue number or all residues from <number1> to <number2>. The residues are indexed according to the numbering provided by the PDB file; note that multiple residues may be returned if the protein has multiple chains. *Example: resid 20 – will select the 20th residue in the protein; resid 20 to 40 – will select all residues between 20th and 40th residues in the protein, inclusively.*
* “helix” – Displays all atoms that are part of an alpha helix.
* “betasheet” - Displays all atoms that are part of a beta sheet.
* “backbone” - Displays only atoms in the backbone of the protein, i.e., only α-carbon, carbonyl carbon, oxygens in the carboxyl group, and nitrogen in the amino group will be displayed (see the picture below).

You can also combine selections by using a combination of Boolean operators: **and**, **or**, or **not**.

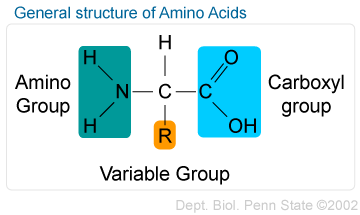
Additionally, using parentheses will enable you to change the order of operations in which your selections are made. *Example: not helix and resid 15 to 30 – will select all residues between residue 15 and 30, inclusive, that are not a part of a helix; not (helix and resid 15 to 30) – will select all non-helix residues that are not between residue 15 and 30*

You can add representations by clicking **Create Rep,** and you can change the drawing method, color, and selected atoms it displays without affecting any of your other representations. You can delete drawings you have created by clicking **Delete Rep**. If you want to temporarily hide your drawing without deleting it, double-click on the representation under “**Style Color Selection,”** and the selection will turn red. You can double-click it again to unhide it at a later time. For certain drawing styles, you can increase the resolution of a representation for higher-quality drawings. Note: for any changes to take effect, you MUST select the representation you intend to modify. The selected representation is indicated by the green highlighting. In some cases, you will need to click Apply in the bottom right corner of the Graphical Representations window for your changes to take effect.

**Try all of the above options before moving to the next section.**

**SARS-CoV-2 Analysis in VMD**

**Note:** for any questions for which you must take measurements, please take a screenshot of the measurement (shift+command+4 in Mac) and paste the photo into your report. Make sure the picture is large enough to be clear (page-wide recommended).



Make sure that the COVID protease protein (in complex with the inhibitor)is loaded and displayed in VMD. A protein consists of the generic amino acid backbone repeated many times. The generic form for an amino acid (a singular building block in a protein) is shown on the right.

Each amino acid in a protein consists of:

1. an **N-terminus** (N for amine, or the –NH2 group),
2. a **Cα**carbon bonded to a characteristic **R-group** (the R stands for the amino-acid **side-chain** and is the only difference between different amino acids), and
3. a C-terminal (C for carboxyl, or **–COOH** group).

**Proteins are read from N-terminals to C-terminals.**

We will begin examination of our 6LU7 structure by finding **its first residue**, which is a Serine (SER). Check the structure of a Serine residue using the included amino-acid side chain handout at the end of this document. If you are having trouble finding this, remember that proteins are folded from linear chains and thus Serine, as the first amino acid, must be on one of the ends of the chain.

CAUTION: This protein structure was imaged in complex with an inhibitor (N3). Check the Protein Data Bank [page](https://www.rcsb.org/structure/6lu7) to make sure you are looking at the proper chain. You will find that there are two chains (A and C), and that one of them corresponds to the protease and the other corresponds to the N3 inhibitor. Make sure you are considering the first residues from the proper chain.

After finding this Serine, work your way down the protein molecule and identify the next three amino acids in the protease. There are two ways to do so.

1. by looking at the amino acid side chains (and consult the table provided at the end of this document), or
2. by labeling atoms. Your mouse will be on 1-atom-labeling mode if you select **Mouse** → **Label** → **Atoms** in the VMD Main window **OR simply press “1”** on the keypad. Then, you can label an atom by clicking on it, which will display the selected atom, the residue name, and the residue number. Using the Drawing Method “CPK” will be helpful for making precise residue selections for this and future exercises.

To clear these labels, go to **Graphics** → **Labels**. After ensuring that “Atoms” is selected in the drop-down at the top left, you can select one or more labels and either “Hide” them, causing the text to turn red, or you can “Delete” them. Note that you can find additional information about a residue within the Labels menu, including its Cartesian coordinates, residue name, segment name, and more. You can also go to **Graphics**→**Colors** to change the color of the label if you can’t read it well. The color settings for labels will be in the **Labels** category. Note that you can modify the colors of atoms, bonds, angles, and other labels. This will be useful later in the lab.

**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 | ? |
| 3 | ? |
| 4 | ? |

Finding amino acids this way can be tedious and slow. A different, more efficient way of accomplishing this in VMD involves the Sequence Viewer, which can be found by clicking on the VMD Main window and going to Extensions → Analysis → Sequence Viewer. This tool gives a residue’s number and name, as well as any secondary structure the residue is involved in. The secondary structure letters stand for:

|  |  |  |
| --- | --- | --- |
| **Letter** | **Color** | **Secondary Structure** |
| T | Teal | Turn |
| **E** | **Yellow** | **Extended conformation (beta sheet)** |
| B | Gold | Isolated bridge |
| **H** | **Purple/Magenta** | **Alpha helix** |
| G | Light Pink | 3-10 helix |
| I | Red | Pi helix |
| C | White | Coil |

**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** |
| ? | ? |
| ? | ? |
| ? | ? |
| ? | ? |
| ? | ? |

**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?
2. What is the ending residue number, and what amino acid is that residue?
3. 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.

Sequence Viewer also allows you to highlight specific amino acids in the display window by clicking on their names in the Sequence Viewer window. A selected residue will turn yellow, and it will become highlighted in the display window. Multiple residues can be selected by clicking while holding down the **Shift** command. A continuous range of atoms can be selected by clicking and dragging, which will produce a red box that can be used to select atoms. If you want to see more or fewer amino acids, use the Zoom bar on the left side of the Sequence Viewer.

Selecting residues in the Sequence Viewer will produce a new representation of that selection in the Graphical Representations window. This representation will be changed the next time you make a selection, so to duplicate the current representation, you can click on it and click “Create Rep”.

You can go into the “Display” settings and choose one of the two display modes.

The “**Perspective**” setting gives the viewer greater depth perception at the expense of distorting scale relationships.

The “**Orthographic**” setting preserves these scale relationships and maintains the parallelism of lines, but with less depth perception. Try both options. **The “Orthographic” setting is better for creating pictures using VMD. Throughout this assignment, you are expected to submit every image in the assignment under the orthographic view with a white background.**

Before proceeding to the imaging questions, we need to clean up our visualization a bit.Firstly, recall that the 6LU7 structure was resolved via x-ray diffraction. As an artifact of this imaging technique, you will notice that there are extra red atoms (oxygen) indicative of water molecules. In order to exclude water, we can use the selection “not water.” We do not want the water in the image, so it is essential that it be removed.

**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?
2. What is the drawing method that the Sequence Viewer is using?
3. 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.

Recall the atom labeling mentioned above. With VMD, you can also label out and then measure bond lengths, angles, and dihedral angles of your protein. The following table summarizes the labeling procedure.

|  |  |  |  |
| --- | --- | --- | --- |
| Press (**Mouse → Label →**) | Hotkey | Number of Atoms Labeled | Value Measured |
| **Atoms** | 1 | 1 | N/A |
| **Bonds** | 2 | 2 | Distance between the two atoms |
| **Angles** | 3 | 3 | Bond Angle |

If you are zoomed out, it can be hard to select the center of an atom, so it may be helpful to zoom in. If it seems like it is not working, it is possible that you are not selecting the center of the desired atom. Remember that using the drawing method “CPK” might make it easier to click on a particular atom.

To un-label an atom that is labeled, simply right-click the atom. Alternately, you can go into the VMD Main window and click on Graphics → Labels to either hide or delete labels.

**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.
2. 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).

**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:
2. Carbonyl C to alpha C:
3. Carbonyl C to N in the peptide bond:

**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:
2. Carbonyl C to alpha C:
3. Carbonyl C to N in the peptide bond

**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?

**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!).

**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)

**Protein Secondary Structures**

Enzymes are molecules that catalyze chemical reactions in the cell. Molecules that enzymes catalyze their reactions are called **substrates**, and the finished molecule(s) are called the **product(s)**. Often, the structures of many enzymes are determined with compounds called **inhibitors** that, not so surprisingly, inhibit the action of the enzyme. Inhibitors can function in several fashions; the type of inhibition we will encounter today is **competitive inhibition**, in which the inhibitory molecule sits in the “active site” of the enzyme and blocks further access to this site by other substrate molecules. The active site refers to the part of the enzyme that the substrate fits into, while the **protein catalytic** residues are the specific residues that chemically interact with the substrate. To learn more about how enzymes work, watch this [video](http://pdb101.rcsb.org/learn/videos/how-enzymes-work).

The enzyme we will be using today, PDB ID **2ER0**, is an example of a class of enzymes called **aspartic proteases**. These enzymes are very highly conserved across many different phyla (meaning that nearly structurally-identical versions of this enzyme can be found in organisms as disparate as fungi and humans). This is because they serve a vital biological function: to cut down large proteins into smaller amino acids to use for energy and biosynthesis reactions. The protein does this by utilizing two of its vital amino acid residues (ASP32 and ASP215) via chemical reactions that are enabled during contact with unstructured (food) proteins in acidic conditions.

A name that you may be familiar with for an analogous human counterpart for the enzyme that we will examine today is **Pepsin**, which is released in an inactive form from chief cells in the lining of the stomach in response to the presence of proteins that you consumed, where it is activated by the low (acidic) pH of the stomach. Once activated, it begins to cleave proteins in the stomach to prepare the amino acids for further digestion and absorption in the small intestine.

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

In proteins, backbone hydrogen bonds stabilize the alpha helix and beta-sheet **secondary structures**, helping the protein maintain an energetically favorable conformation. This pattern of hydrogen bonding is essential in maintaining the secondary structure, and thus the hydrogen bond distances and angles are relatively consistent characteristics for each type of secondary structure.

A hydrogen bond is formed between an atom with a hydrogen bonded to it (the donor, D) and another atom (the acceptor, A) provided that the distance D-A is less than a cut-off distance and the angle D-H-A is less than a cut-off angle. Oxygen and Nitrogen atoms are the most common atoms involved in hydrogen bonding in proteins (usually Oxygen is the acceptor and Nitrogen is the donor).

For this section, **delete everything you may have currently loaded in VMD and open the PDB file 2ER0\_autopsf.pdb (available in the assugnmnet folder) that you can find on Canvas.** This is a molecule edited to contain actual hydrogen atoms so that we can examine hydrogen bonding in the formation of the secondary structures. After you have loaded the molecule, create a representation in the graphics tab and enter “alpha\_helix” in the selected atoms text box, and then choose “NewCartoon”.

Find the sequence range for **the largest** alpha helix in the protein (notice that once you do, you can use “resid x to y” to select it). Now delete all representations and create a new one with this sequence range. Use NewCartoon as the drawing method and choose a suitable color for it (or color it by secondary structure). Change the material to transparent (so that the other representations that we add hereafter are also visible).

Now, add a new representation. Select "resid x to y and name N O HN" and choose CPK as the drawing method. Since we are interested in knowing what atom the donor is and what atom the acceptor is in the hydrogen bonds, color it by name.

Add a third representation with all the **backbone atoms** of the selected alpha-helix and choose the drawing method as CPK. You can leave the coloring method as names for the backbone or choose some appropriate color. Now, if you wish, you can in each of these representations change the resolution parameters so as to yield a better rendering of the representations.

Finally, add the fourth representation with the selection "resid x to y and name N O HN" and choose “HBonds” as the Drawing Method and set the parameters Distance Cutoff to 3.5, Angle Cutoff to 35, and Line Thickness to 5. This will draw hydrogen bonds in the molecule.

Backbone hydrogen bonds are shown in dark blue, which may be difficult to see. To change the background color, you can go into Graphics → Colors, select the “Display” category, and change the value for “Background” color. If you follow all of these steps correctly, you should have a picture that shows the hydrogen bonds clearly.

**Q12.**

1. How many hydrogen bonds are there in the largest alpha helix?
2. 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)?
3. Measure the distance and angle for one of these hydrogen bonds and attach a clear picture of your helix with its hydrogen bonding network.

**Q13.**

1. Do the side chains on an α-helix point into the center of the helix, or to the outside?
2. Why? (Try to show the sidechains in your helix representation)

**Q14.**

1. How many cysteine residues are present in the molecule?
2. What are their residue numbers?
3. What unique structure do they form, and between which two atoms is the structure?
4. What do you think is one function of these covalent cross-links in this molecule?
5. 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.

Enzymes are often crystallized while bound to a molecule that they interact with in order to ascertain which features of the enzyme are important for the catalysis of the substrate. This enzyme was crystallized while bound to a “competitive inhibitor” named ***pepstatin***.

To view the enzyme only, in the graphical representations window enter the text “**chain E**” in the “Selected Atoms” text box. To view the inhibitor molecule only, create a new representation in VMD and enter the text “chain I” in the selected atoms.

**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.