Lower-level Activity

Laboratory A: Molecular Visualization and Interpretation through iCN3D

Anthracycline Respinomycin D: A commonly used anti-cancer drug

Learning objective:

- Visualize a "mystery" macromolecule using a molecular visualization web-based software.
- Interpret the three dimensional structure and determine what it is.
- Examine how the macromolecule interacts with a small molecule.

Work in groups of 2 students. Answer the questions below and then read the introduction in page 2 - 3.

1. What is H-bonding?

2. What are covalent bonds?

3. What are ionic interactions?

4. What are London dispersion forces?

INTRODUCTION

According to the World Health Organization (WHO), cancer is the leading cause of death worldwide accounting for approximately 10 million deaths in 2020. Cancer is not a single disease but a group of diseases characterized by abnormal cell growth; many cancers have the potential to invade other parts of the otherwise healthy body. If diagnosed early, many cancers can be cured or effectively treated.

The number of cancer survivors continues to increase as the success of cancer treatment regimens improves. The number of people in the United States living after cancer was diagnosed was 16.9 million as of Jan, 2019 and is projected to increase to 21.7 million by 2029. In the early 1970s, 12% of the children who survived cancer of any type died within 15 years of diagnosis: by early 1990s, the proportion had decreased to 6%. Anthracyclines are still the chemotherapeutic drug class of choice for treating many cancers including, leukemias, lymphomas, and cancers of the breast, stomach, uterus, ovary, and lung, among others.

Respinomycin D, is a member of the anthracycline family of antitumor antibiotics that interact with **one** of the four types of biological macromolecules (i.e. proteins, lipids, nucleic acids, carbohydrates).

In this lab, you will:

- Determine the structure of the macromolecule that is a target of the respinomycin D drug.
- Explore the mode of action of respinomycin D. In other words, determine how this molecule interacts with its target macromolecule.
- Further investigate how the drug is used to treat cancer.

The four types of macromolecules, a brief description.

Carbohydrates: Carbohydrates are sugars, molecules containing carbon (C), hydrogen (H), and oxygen (O) atoms typically with a 2:1 (hydrogen-oxygen ratio). The general formula is $C_m(H_2O)_n$. In solution, sugars form cyclic structures. The simplest carbohydrates are called monosaccharides. This are the building blocks of disaccharides (2 monosaccharides) and polysaccharides (3 or more monosaccharides). Monosaccharides form covalent bonds called glycosidic bonds.

Proteins: Proteins are polymers of amino acids, thus, amino acids and the building blocks of proteins. There are four levels of protein structure: Primary (amino acid joined to the next by a covalent bond called peptide bond), secondary (alpha helix or beta sheets), tertiary (three dimensional fold forming a subunit) and quaternary (multiple subunits interacting with each other).

Nucleic acids: Nucleic acids are polymers of nucleotides. Thus, nucleotides are the building blocks of nucleic acids. Nucleotides join to the next via a covalent bond called phosphodiester bond. A string of nucleotides is referred as a strand. The nucleic acid consisting of a single strand is called RNA (ribonucleic acid) while the nucleic acid composed of two strands (joined to each other via hydrogen bonds) is called DNA (deoxyribonucleic acid). DNA forms a helical structure.

Lipids: Lipids do not form polymers, they form aggregates. The structure of lipids is quite diverse and includes linear and cyclic structures.

Now that you are familiar with the four types of macromolecules, you will use a three-dimensional molecular viewer called iCn3D to determine what is targeted macromolecule of the anticancer drug Respinomycin D.

Part I

PROCEDURE

1) Go to https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html or google icn3d and click on iCn3D: web-based 3D structure viewer (Fig. 1).

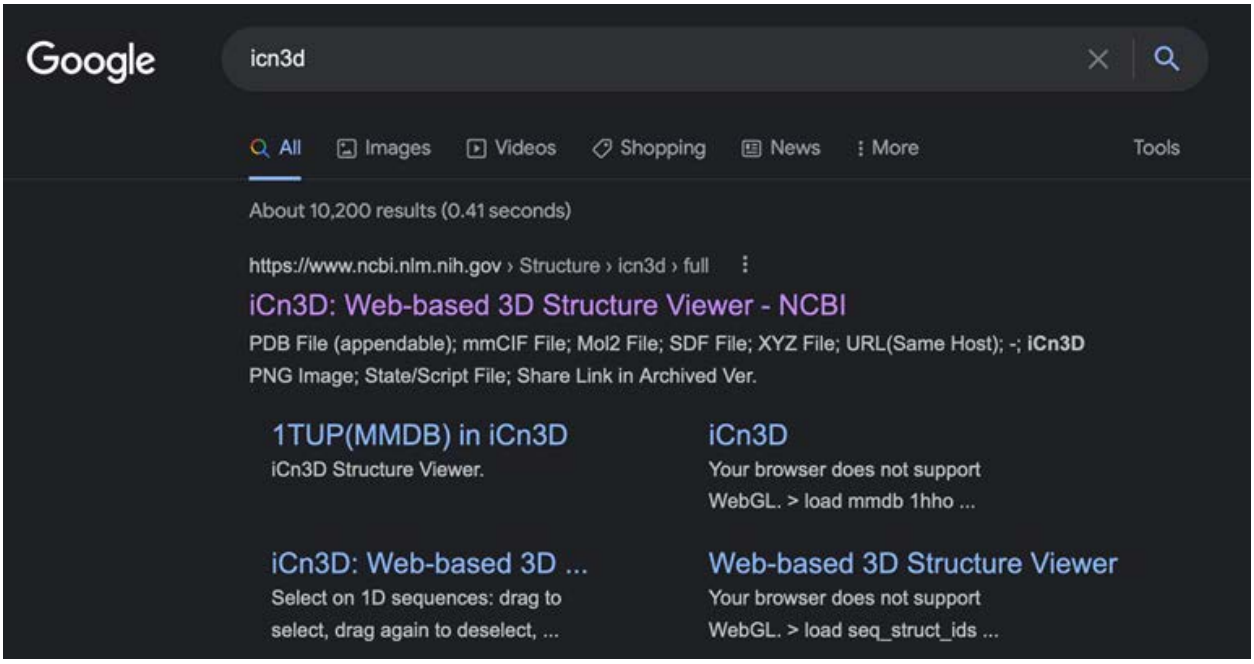


Fig. 1

2) Click on File – Retrieve by ID – PDB ID (Fig. 2).

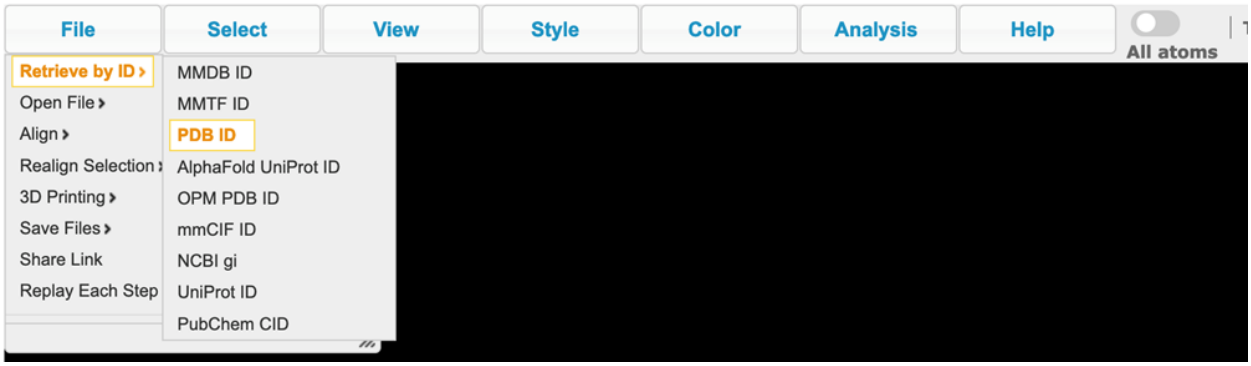


Fig. 2

3) Type in 1N37 and click Load button (Fig. 3).

Fig. 3

4) You should see the below visual (Fig. 4). See if you can rotate it with your mouse. You can also zoom in and out.

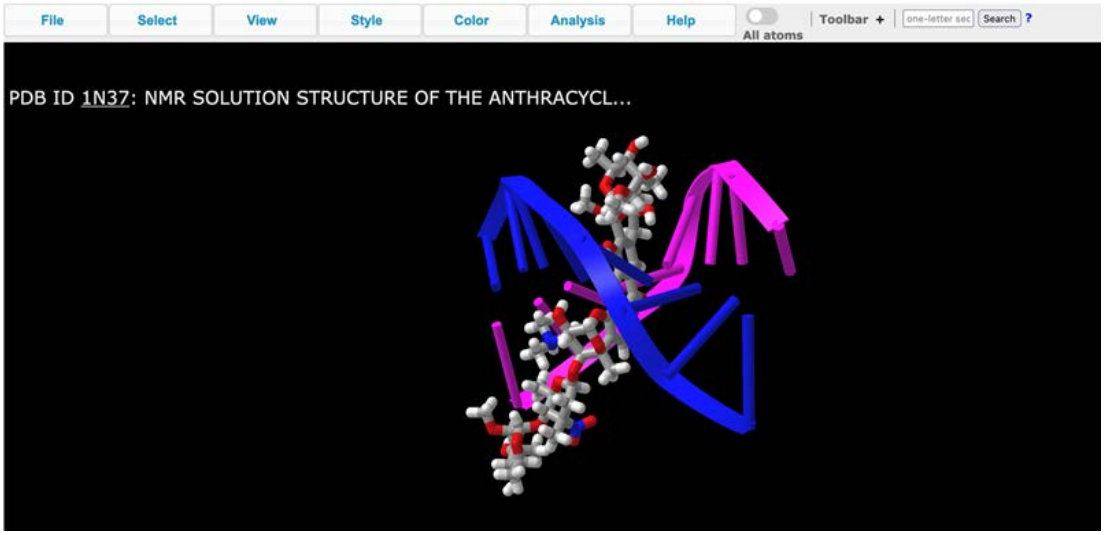


Fig. 4

In the image above, the targeted macromolecule is colored in blue and pink. The drug is colored in gray and red. Explore the image (rotate, and zoom in and out) and answer the questions below:

Question 1. Does it look like this macromolecule is composed of one or two strands? If two strands are present, are the strands straight or 'twisted" around each other? Would you say that the molecule is straight or helical?

5) The menu functions (yellow rectangle in the figure below, **File, Select, View, Style, Color, Analysis,** and **Help**) are what will allow you to explore the structure of any protein in detail. Click on Analysis → Interactions (A new box appears, see Fig. 5).

1. Unclick everything except Hydrogen bonds. 2. click selected. 3. Click selected. 4. Click 3D Display Interactions. Then close the box to see the interactions on your molecule.

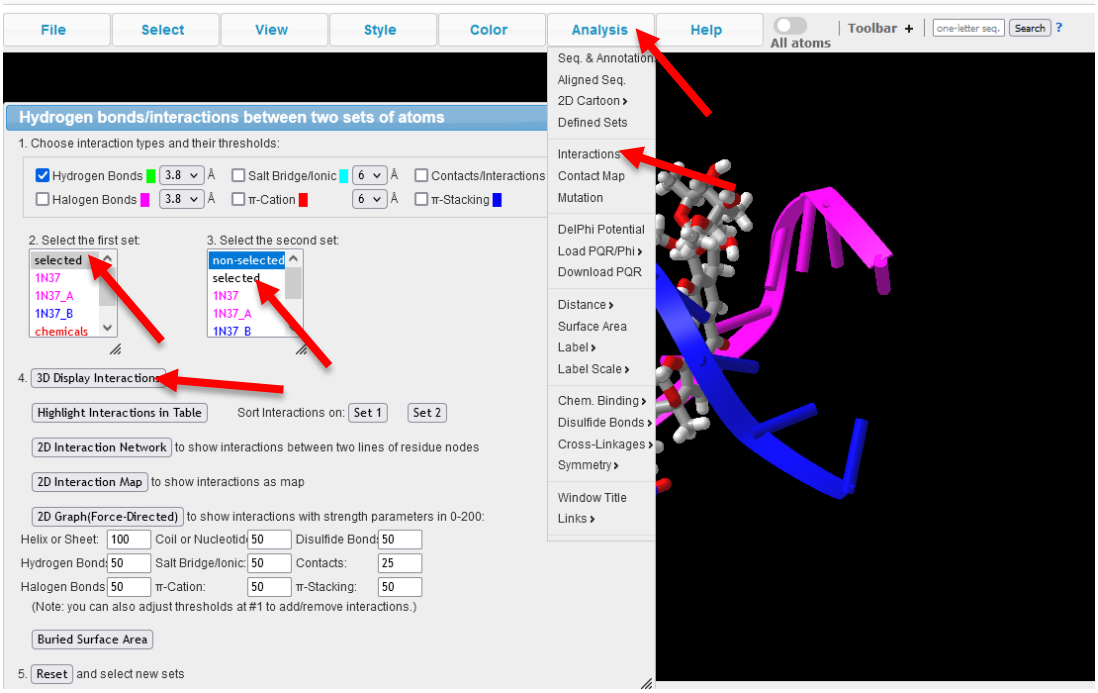


Fig. 5.

7. You should see the below image with all the H-bonds as green dotted lines. Take a screenshot and paste it on the space below.

Screenshot

Question 2: Observe and describe using your own words the location of the hydrogen bonds. What is the function of the hydrogen bonds?

8. Click Color - Atom

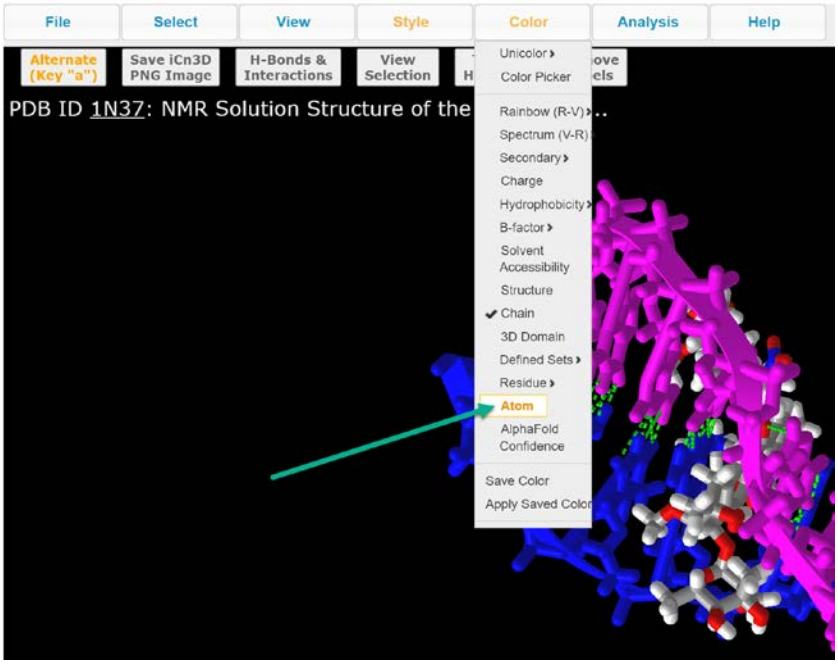


Fig. 6

9. You should see all the atoms in the strand. The coloring is as follows: Grey – carbon, Red – oxygen, Blue- nitrogen, Yellow – Phosphorus, and white – Hydrogen. Take a screenshot and paste it in the space below.

Screenshot

By coloring the atoms, you will notice an obvious pattern. The strands (ribbons and associated atoms) are in (which colors) _____ and _____ while the "inner" part of the macromolecule is colored in _____, _____ and _____. We will refer to the strand(s) as the backbone of the molecule.

Question 3. Based on the color code described above, we can say that _____ and _____ atoms are present in the backbone of this macromolecule. While nitrogen is only present is the "inner" part of the macromolecule.

Read the definitions of a sugar and a nitrogenous base below and determine which of these two are present in the "inner part" of the molecule.

Simple sugar: A simple sugar is a small molecule typically composed of C, O, and H

Nitrogenous base: A nitrogenous base is a cyclic, organic, nitrogen-containing compound.

Question 4: The inner part of the molecule is composed of _____.

In the space below summarize your findings:

a. The macromolecule is composed of _____ strand (s).

b. The strand (s) are _____ (straight / helical).

c. The strand (s) are held together by _____ (ionic interactions / covalent bonds / hydrogen bonds) that form between _____ (sugars / nitrogenous bases)

With all the evidence and observations that you have collected up to this point, you should be able to determine the identity of the "mystery" molecule that is a target of the anticancer drug Respinomycin D.

Question 5: What is the is (carbohydrate, protein, nucleic acid or lipid)? Be as specific as possible.

In the old days, three different models of this macromolecule were proposed. A description of each model is listed below. Based on your finding above, determine which of the proposed models is the most accurate (fill in the blank space with the name you identified in question 5):

Model 1: _____ model consisting of three chains for nucleic acid intertwined. In this model, the phosphates are located near the fiber axis with the bases pointing to the outside.

Model 2: _____ model consisting of three chains for nucleic acid. In this model, the phosphates are located on the outside and the bases on the inside linked together by hydrogen bonds.

Model 3: _____ model consisting of two chains for nucleic acid. In this model, the phosphates are located on the outside and the bases on the inside linked together by hydrogen bonds.

Summary: The macromolecule that is target of the drug Respinomycin D is _____ (carbohydrate / protein / lipid / nucleic acid. Specifically, _____.

Now that you know what macromolecule is a target of the anti-cancer drug Respinomycin, you will determine the mode of action of this drug. First, let's see how the drug interacts with _____.

Part II

How does Respinomycin D interact with its targeted molecule?

Close the iCn3D window and relaunch the browser. Import the image 1N37
https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html

10. Now look at where the anthracycline is interacting with the DNA. Is there any H-bonding occurring there? If so where? Note: It may be easier to see the anthracycline when colored by chain. Review previous steps in this tutorial and color the image by chain. Take a screenshot and paste it in the space below:

Screenshot

Question 6. Observe the diagrams below (Fig. 7) and determine which one would correspond to the structure of Respinomycin D.

Image A

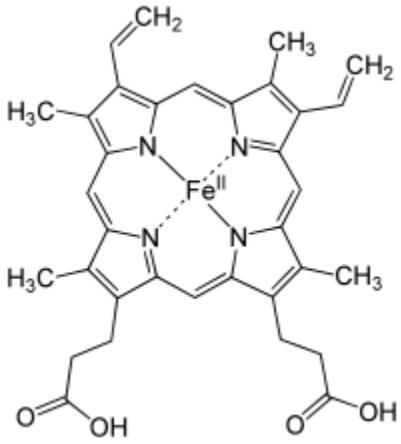


Image B

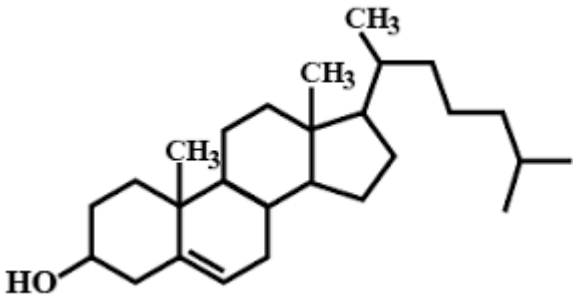


Image C

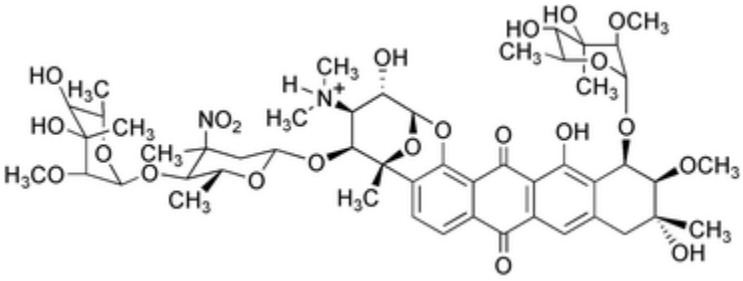


Fig. 7. Three images. One of them correspond to Respinomycin D.

Question 7. Based on your observation, determine which of statements below regarding the mode of action of Respinomycin D makes the most sense.

Statement 1. The drug Respinomycin D binds to the DNA backbone and prevents the binding of proteins (transcription factors) that are essential for transcription.

Statement 2. The drug Respinomycin D is an intercalating agent. This means that it inserts itself within the DNA structure. Specifically, in between base-pairs.

Statement 3. The drug Respinomycin D is a competitive inhibitor that binds to proteins (transcription factors) essential for transcription.

Question 7: Besides the mechanism of action by the drug Respinomycin D, in what other way would you expect other anticancer drugs work? Could other macromolecules be the target?

Question 8: The drug Gleevec® is used to treat cancer of the white blood cells. Investigate its mechanism of action and identify the target macromolecule.

Upper-level Activity

Structure of Lysozyme and Key Amino Acids Involved in the Interactions with NAG: An iCN3D Activity

Learning Objectives:

1. Understand the interactions between amino acids in various levels of protein folding.
2. Visualize a protein in 3D and 2D.
3. Connect the 3D structure of a protein to its function
4. Be able to use a molecular visualization program to create an image

Concepts that need to be explained before exercise

- Amino Acids (names, 3 and 1 letter abbreviations, and structures)
- Primary, secondary, tertiary structure of proteins
- Alpha helix and beta sheet
- How 3D structures are made (protein crystallography and 3D NMR)

Viewing Proteins in 3D

Proteins are a 3D structure. Without this 3D structure they would not be able to perform their function. In this lab you will explore how this 3D structure is held together through the use of a web-based molecular visualization program called iCN3D (pronounced I SEE IN 3D). This lab has two parts:

1. We are going to go through a tutorial that will help you learn how to use the program.
2. You will use the program to investigate the structure and function of a protein of your choosing.

Pre-lab Questions

We will be using the enzyme lysozyme for our tutorial. Answer the below questions on lysozyme from your textbook (pages XX-XX)

Pre-lab Question 1: What is lysozyme? What is the function of lysozyme?

Pre-lab Question 2: A small molecule referred as NAG is found in association to lysozyme. What does NAG stand for?

Tutorial:

The National Center for Biotechnology Information (NCBI) has free repositories of 3D structures of biomolecules. In the database each structure of a protein has a separate PDB file name by having a number followed by three letters. There have been many structures of lysozyme put into the database. The one we are going to be using has been given the code 1HEW. In this version of lysozyme, researchers crystalized a lysozyme that was bound to a NAG molecule. We will begin our tutorial by learning how to use the NCBI. On your computer go to

https://www.ncbi.nlm.nih.gov/ or type in NCBI in your search engine. On the NCBI home page you will see the many resources available there, feel free to come back an explore later. To find 1HEW in the database, simply type in 1HEW into the search box as seen in the image below (Fig. 1):

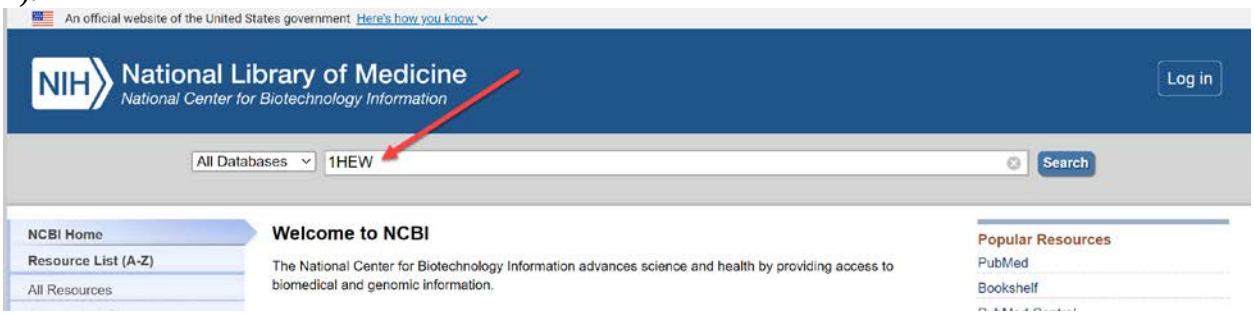


Fig. 1. NABI search box

Once you click the search button you can do two things: 1) view the information in the PDB file by clicking the top blue link 2) view the molecule in iCN3D by clicking the "view in iCN3D link" (see Fig. 2).

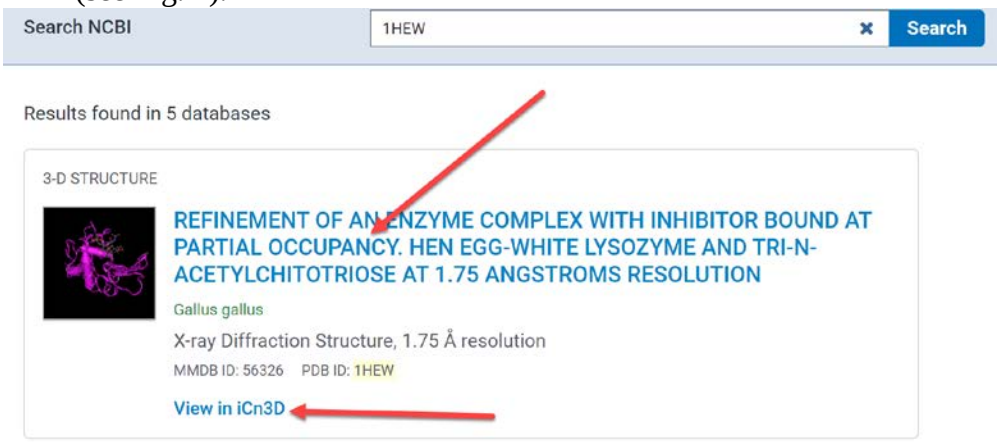


Fig. 2. *Gallus gallus* Lysozyme (1HEW)

PDB Structure Summary

We will start by exploring the PDB structure summary. Please click the top link.

Q1. Identify 3 pieces of information that you found in the PDB structure summary of 1HEW

1.

2.

3.

From the PDB structure summary you can also go straight into the iCN3D viewer by clicking on the **"full-feature 3D viewer"** link. Once you have explored the summary go ahead and click on the link (Fig. 3).

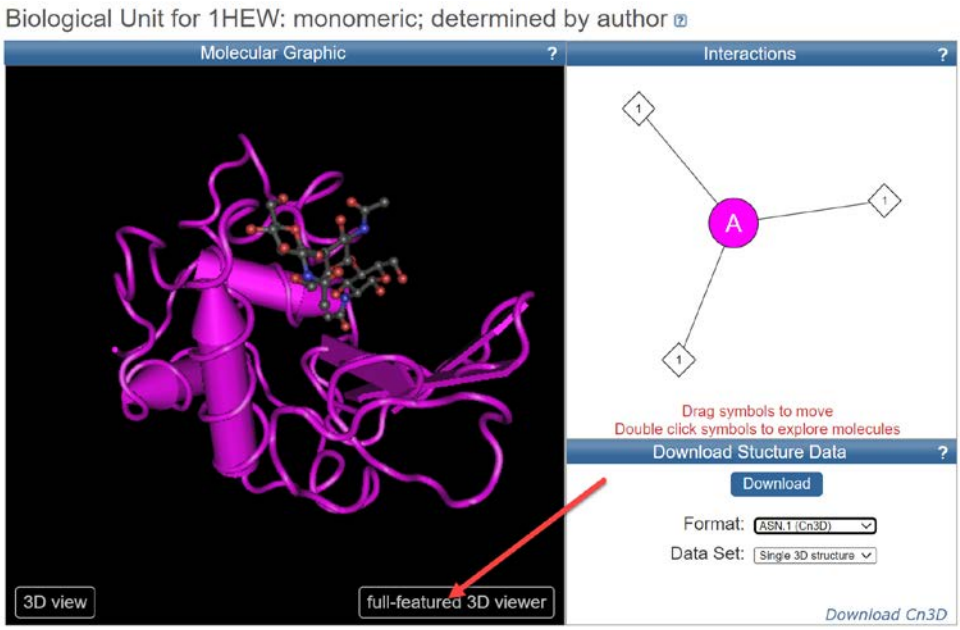


Fig. 3. 3D structure of lysozyme (1HEW) with "full-featured 3D viewer" button.

Once the protein has fully downloaded you will see the below screen. Notice how in the **"Sequences and Annotations"** box specific domains as well as NAG are identified (Fig. 4). Take a few minutes to use your touch screen or mouse to rotate the bound pair (lysozyme + NAG) and investigate it's structure.

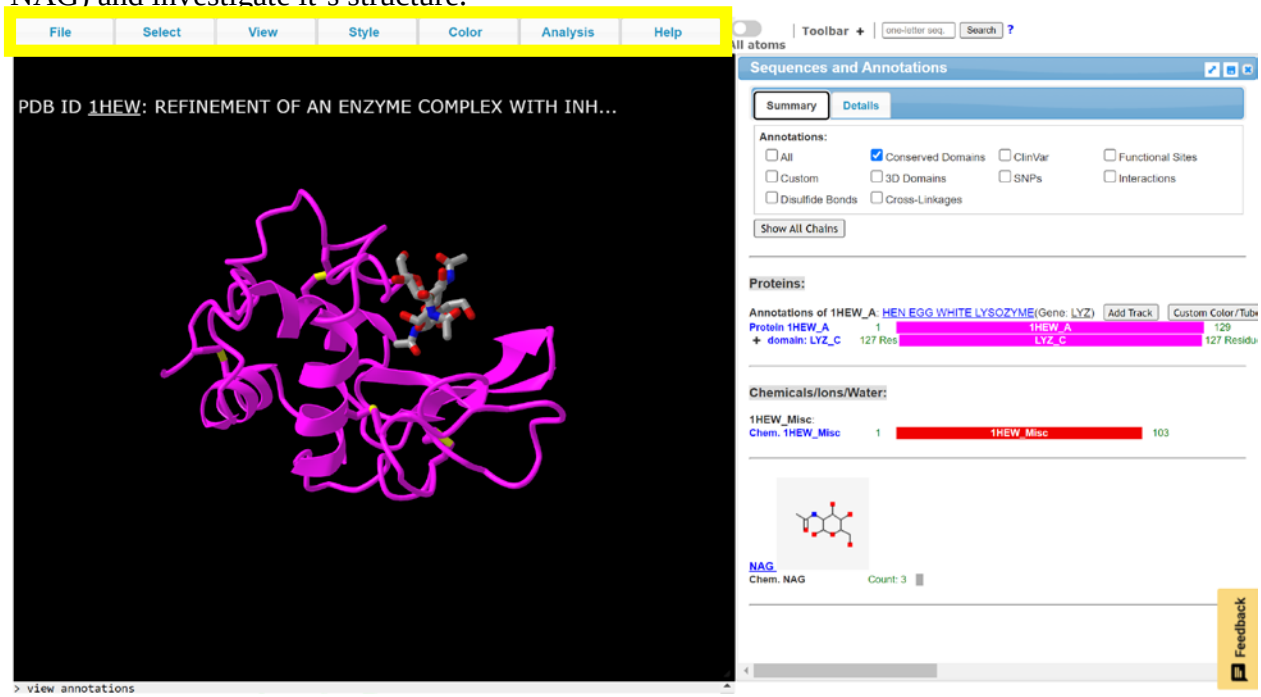


Fig. 4. Fully downloaded 1HEW structure.

Q2. How are alpha helices and beta sheets represented in the iCN3D structure?

Menus

The menu functions (yellow rectangle in figure 4, **File, Select, View, Style, Color, Analysis,** and **Help**) are what will allow you to explore the structure of any protein in detail. In this section of the tutorial we will explore some of the menu functions. There are many more functions beyond what is explained here. Once you learn these basic functions feel free to explore the others.

The first function we will explore is the Sequence and Annotations box that appeared on your screen when you first came to iCN3D. To open this box at any other time you need to go to the **Analysis** menu and then you will see the **Sequence and Annotation** option (Fig. 5).

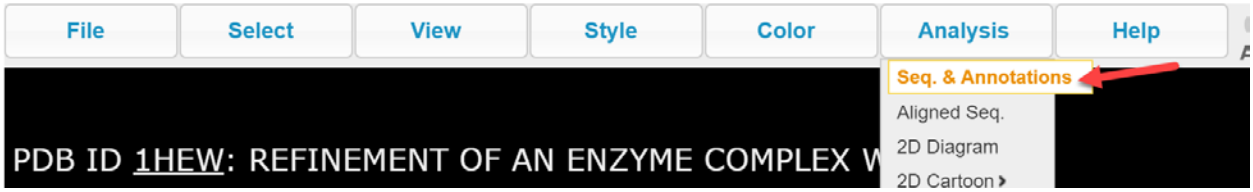


Fig. 5. The **Seq. & Annotation** option can be found within the **Analysis** tab.

Within the **Sequence and Annotations** box there are two tabs: **Summary** and **Details** (Fig. 6). Under the **Details** tab you will find the details about the amino acids in your protein. These amino acids are listed as their 1 letter abbreviation. Above the sequence of amino acids you will see a wave diagram indicating if the sequence forms a helix or an arrow diagram indicating a beta sheet (Fig. 6).

Q3. How many amino acids does lysozyme have? How many helices and beta sheets does lysozyme have?

Q4. Draw the structure of the NAG molecule below.

Fig. 6. Structure **Summary** and **Details** can be found within the **Sequence and Annotation** window.

Next we will explore the Color menu (Fig. 7). This menu will help you better visualize different aspects of lysozyme structure. Follow the directions below to change the color of the helices.

- Click the **Helix Sets** found in the **Sequence and Annotations** box. It is above the amino acid sequence of the protein is a gray box that reads: Helix Sets (see Fig. 7)
- All helices will now be highlighted
- Go to the Color menu and click Unicolor and select blue

.

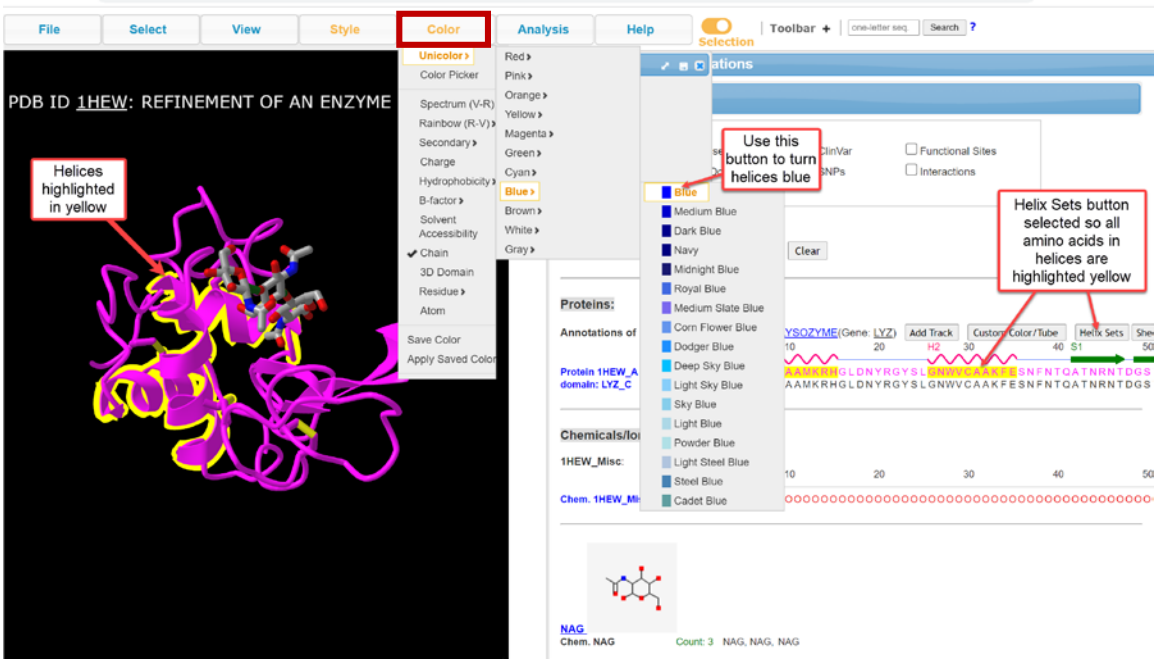


Fig. 7. Color menu

Q5. Insert a screenshot of lysozyme with blue helices. Then change the color to **Secondary→Spectrum** and describe what happens.

Now we will learn how to color a specific helix. Before we can do that we need to unhighlight everything (Fig. 8). To do that.

- Go to **Select** Menu
- Click **Clear Selection**

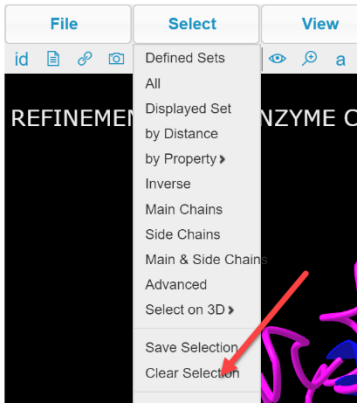


Fig. 8. Select menu.

To select a specific helix and color it a different color

- Click and drag from Gly4 [the fourth letter (G)] to Gly16 [the sixteenth letter (G)]. The numbers above the sequence will help you find the 4th and 16th letters (Fig. 9). Once you do this all the letters between the 4th and 16th will be highlighted yellow. If you accidently go too far, you can simply click and drag on the unwanted letter(s) and they will unhighlight.
- Color the helix by charge (acidic residues will be red; basic residues will be blue)

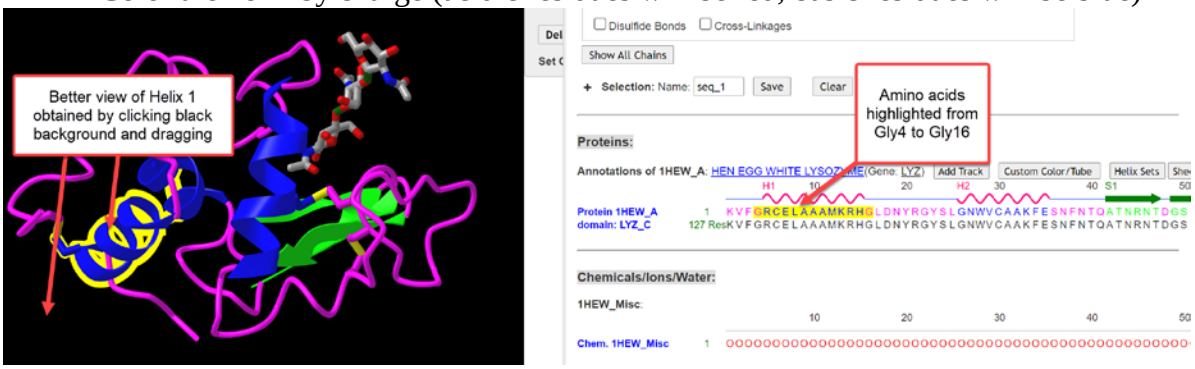


Fig. 9. Selecting and coloring specific amino acids

Q5. How many acidic and basic residues are in the first helix? Use the sequence and annotation box to determine which amino acids these are.

Now let's explore how the interactions within the protein and between the protein and the NAG molecule.

- With the first helix still selected, go to the **Analysis** menu and click **Interactions**

Fig. 10. Interactions can be found within the **Analysis** tab.

- In the Interactions box that appears (Fig. 11), unclick everything but Hydrogen Bonds. And in section 2 highlight 1HEW_A_H1 for first set and second set. This will show you the hydrogen bonds within the helix

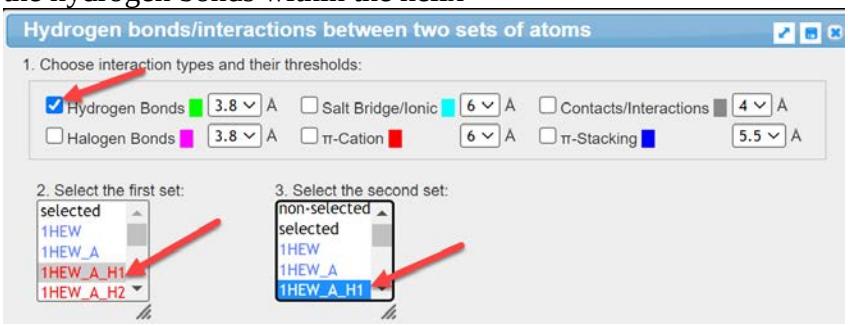
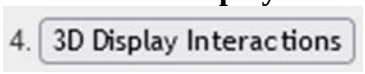


Fig. 11. Interactions menu and selection of hydrogen bonds.

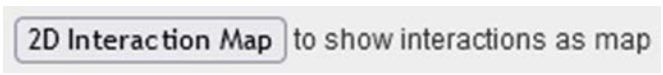
- Click the **3D Display Interactions** gray button (section 4) and close the box.



- To view only the helix, in the **View** menu click **View Selection**
- To better view what atoms are hydrogen bonding, in the Color menu click Atoms

Q6. Describe which atoms are hydrogen bonding within the helix.

- Go back to the Analysis menu and click Interactions. Put in the same settings as we did before (Fig. 11), but this time click the **2D interaction Map** button in a gray rectangle.



- You will get a map of which amino acids are hydrogen bonding to each other in the helix (Fig. 12).

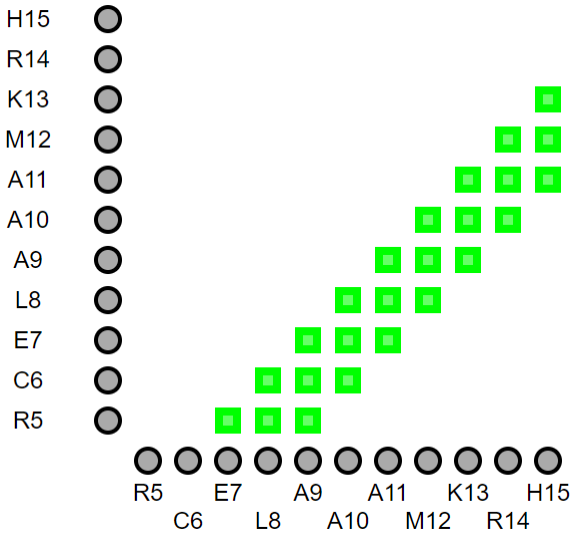


Fig. 12. 2D interaction map showing that hydrogen bonding occurs at positions other than the typical n+4.

Q7. Based on what you know about hydrogen bonding in a helix what amino acid should C6 (amino acid 6 which is a cysteine) be hydrogen bonded to? Explain your choice and does the chart confirm that?

- Now we will reset everything. Open the **Interactions** box again and click **Reset** (this will erase the hydrogen bonds). Reset is section 5 in a gray rectangle. Close the Interactions box and in the **View** menu click **View Full Structure**. To reset the color go to **Color → Secondary → Sheets in Yellow**
- In the **Select Sets** box, scroll down, find and click on **chemicals** (Fig. 13).

Fig. 13. The **Select Sets** box is marked with a red rectangle. The red arrow is points at **chemicals.**

- Once you have select chemicals, in the **Sequences and Annotation** box the three NAGs (found under the amino acid sequence) become highlighted (Fig. 14).

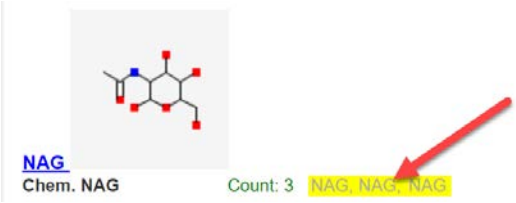


Fig. 14. There are 3 NAGs associated to lysozyme (count: 3). The three are highlighted in yellow.

- Then, go to the **Select** menu and click **by Distance** and type 5 into the box next to **Sphere with a radius**. This will highlight all the amino acids within 5 angstroms of the NAG and create a defined set with these amino acids.

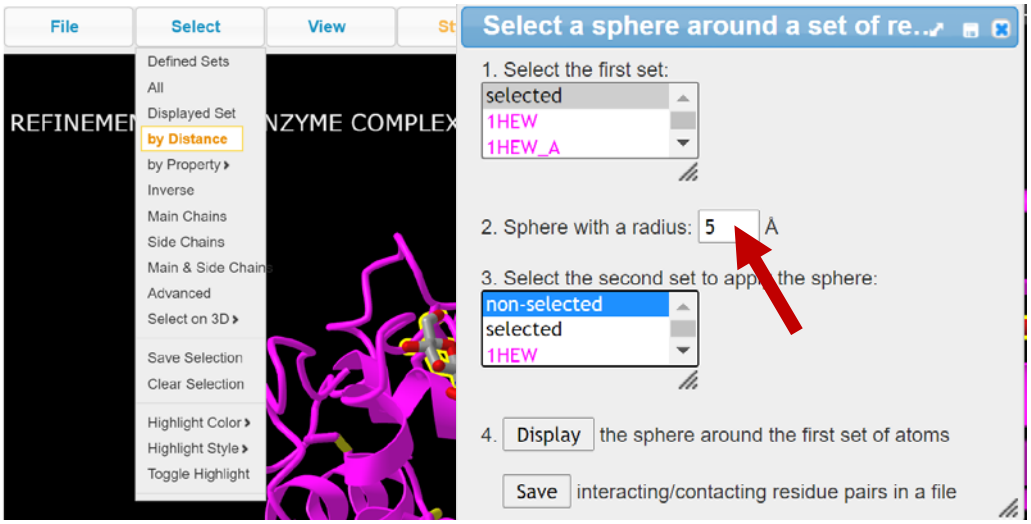


Fig. 15. Select by **Distance selection.** Specify 5 A in section 2, then click **display.**

- Click **Display (Fig. 15)**.

- To see the NAG and the amino acids within 5 angstroms. Go to the **Select** menu → **Defined Sets**

- A **Select Sets** box (Fig. 16) will open. Click while pressing the Ctrl key on **chemicals** and **sphere.NAG.NAG 1002-5A.** Both should be highlighted.

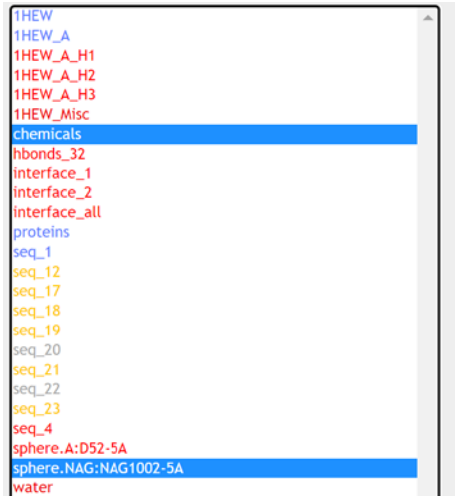


Fig. 16. Chemicals and sphere.NAG.NAG 1002-5A are both selected in the select sets box.

- To see only the highlighted atoms, click **View→View Selection**. You should see an image like below (Fig. 17).

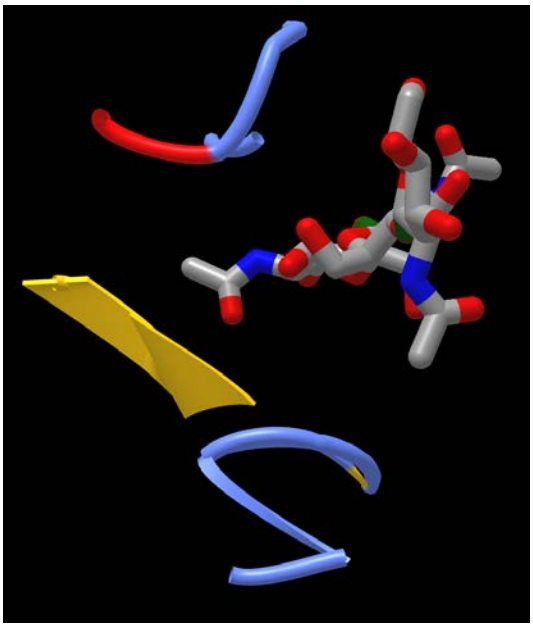


Fig. 17. NAG interacting with amino acids within 5 Å.

- To get a better view we need to get an atom view and not a ribbon view of the protein. In the **Style** menu click **Proteins→Stick (Fig. 18)**.

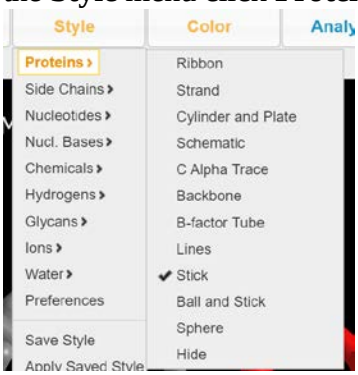


Fig. 18. Protein viewed as stick.

- Now we can change the color to indicate the polarity of the amino acids by going to the **Color** menu and selecting **Charge**
- To make the image even better we can color the NAG by atom (Fig. 19). To do this select the NAG only and in the Color menu select Atom (C-gray, O-red, N-blue, S-yellow). You now know two ways to do this: 1) Highlight NAG in Sequence and Annotation menu or 2) select chemicals in defined set box. The image should look like the one below.

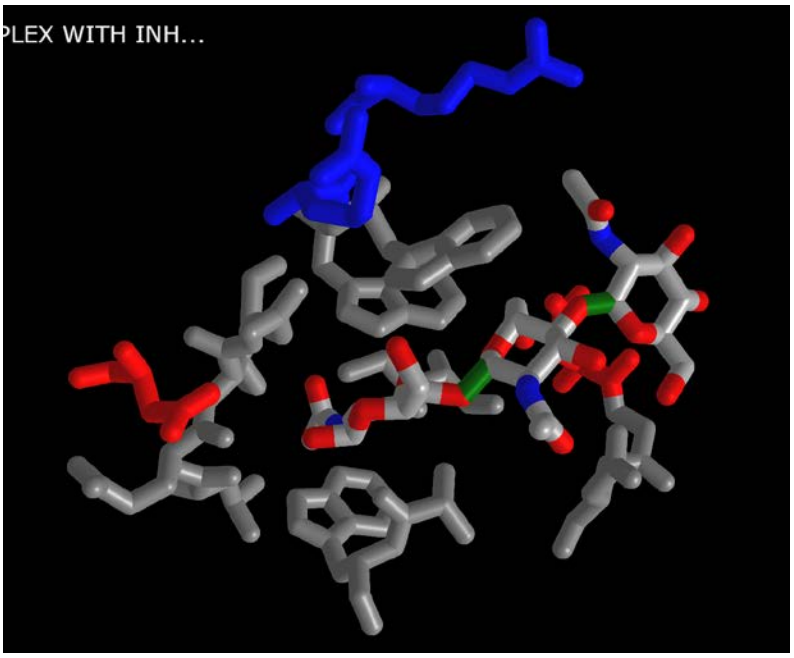


Fig. 19. NAG colored by atom interacting with the protein.

Q8. What type of amino acids are binding to the NAG. What type of interactions do you think they are using? Determine the identity of the amino acids using by placing your cursor over the amino acid. A gray box will appear with the three letter abbreviation of the amino acid.

- To verify your answer in question 8, let's see the H-bonds between the NAG and the amino acids use previous directions and the following setting for 1. and 2 (Fig. 20).

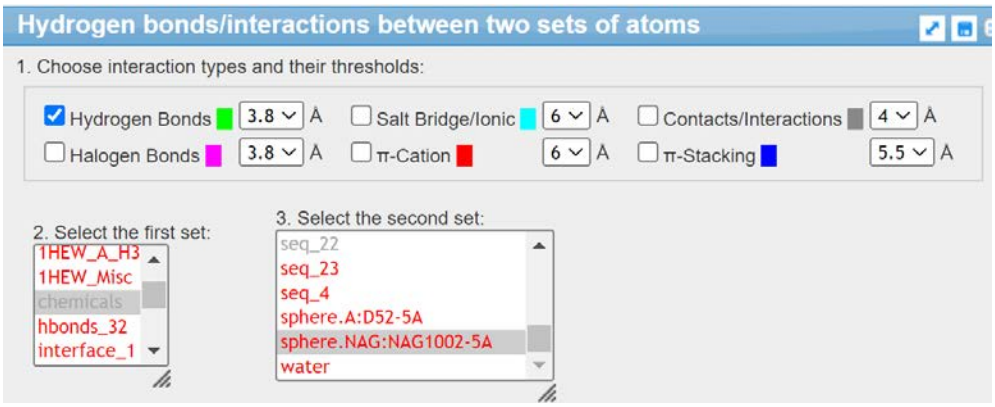


Fig. 20

- Click 3D Display Interactions

Q9. Was the result what you expected for where the H-bonding was located/not located. Explain where the hydrogen bonding was located (backbone atoms, nonpolar side chains, polar side chains, acidic/basic side chains).

Q10. What you are seeing is the active site of lysozyme. Based on your image what do you think are the key residues at the active site?

Q11. In human lysozyme the reaction peaks at pH 5.1 and lowers if the pH is lowered. Based on this additional information, how does that support/or not your hypothesis on what the key residues are.

Now that you know some basic functions of the program. You will choose your own protein from the protein databank and describe it similarly to how we did lysozyme.

Requirements of the lab report.

- Protein must be bound to something
- Must show at least 8 views of protein
    - Some examples include: secondary structure, location of charged amino acids, etc.
- Must have a written explanation of each view and why it is important
- Must connect structure of protein to function (one source must be cited)