Protein Visualization with Mol*

Glen Hocky, Department of Chemistry, New York University

Reminder

Proteins are **chains** of **amino acids**. They **fold** into three-dimensional structures, which depend on which amino acids make up the protein.

How do we know what proteins look like?

As will be discussed in class, scientists have different techniques for finding what proteins look like, which involve bouncing **X-rays** or **electron beams** off of the proteins, and then using computers to figure out what unique structure could lead to the patterns observed.

Scientists must publish their protein structures

When a scientist "**solves"** a protein structure, they release it to the whole world on a website called the **Protein Data Bank**. We all have access to all of this information for free!
[https://www.rcsb.org]

Each protein structure has a **four letter code**. For example, a structure of a chaperone protein discussed by Prof. Lupoli last week is:
[https://www.rcsb.org/structure/4IO8]

Visualization Software

Mol* (or molstar) is a new web software produced with the PDB so that we can look at and analyze protein structures in the web browser.

For example, this link will bring up the program and show the structure 4IO8 mentioned above:
https://molstar.org/viewer/?pdb=4IO8&collapse-left-panel=1

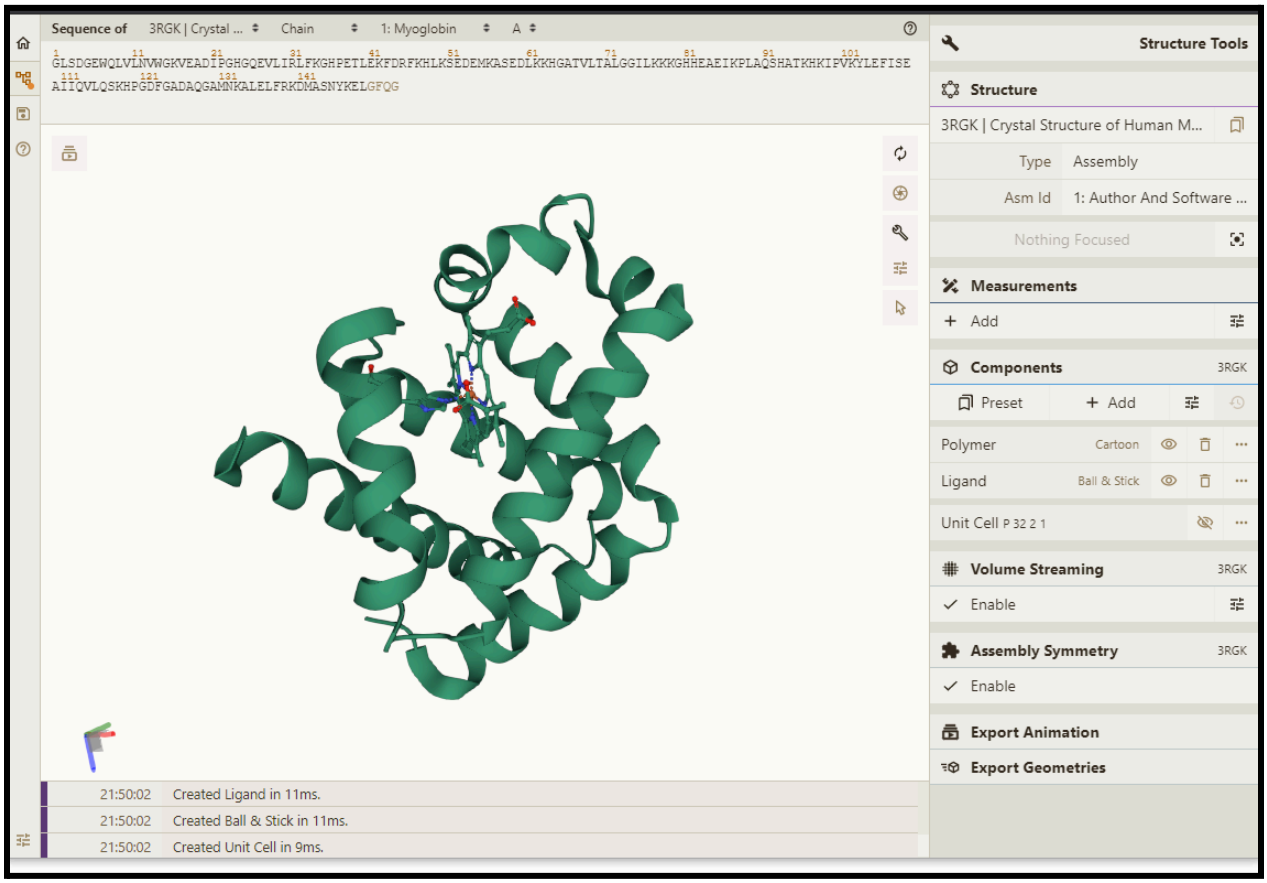
Assignment: Myoglobin

Myoglobin was the first protein solved by X-ray diffraction in 1958. It binds oxygen in our body using an *iron* (Fe) containing molecule.

Use the following link to show one of the many structures of myoglobin (3RGK) solved since 1958.

https://molstar.org/viewer/?pdb=3RGK&collapse-left-panel=1

It should look like this:



The full documentation is here in case you need it:
https://molstar.org/viewer-docs/

Task 1: try moving the protein around with your mouse/trackpad

- You can spin the protein around by clicking on the white area and dragging
- You can move the protein up/down/left/right by right-clicking on white area and dragging

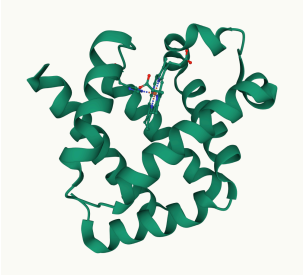
- You can zoom in and out by ‘pinching’ on a screen, two finger drag on a trackpad or with a mouse wheel if you have one

- Reset the view by clicking this button: 

Task 2: take a picture

- Hide the control panel with  [optional]
- Get the protein in a view that you like

- Click  to bring up a menu to take pictures.
- Decide what resolution you want, and whether you want the background to be transparent, and then download the picture. For example, here is one that I took

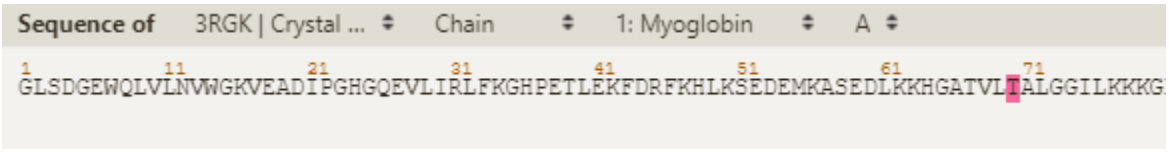


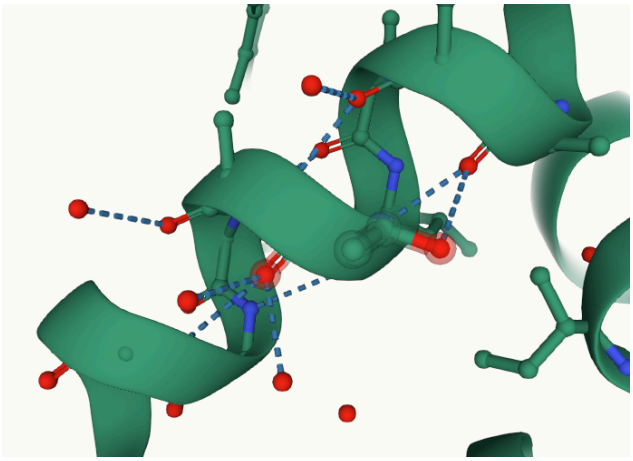
- Turn the control panel back on with 

Task 3: find different parts of the protein

- Put your mouse over the protein and note that as you move it along, different areas are lit up and info pops up in the bottom right corner. For example, this is amino acid #70 or **residue 70**

- You can also pick residues from the list of letters at the top of the screen



- Click on one of these letters or double click on the protein, and it will zoom in to that place. New things appear- the new lines are the actual chemical structure of the protein (e.g. oxygen in red, nitrogen blue) - *the rest is just a cartoon!*
- 
- After exploring, reset the view with 

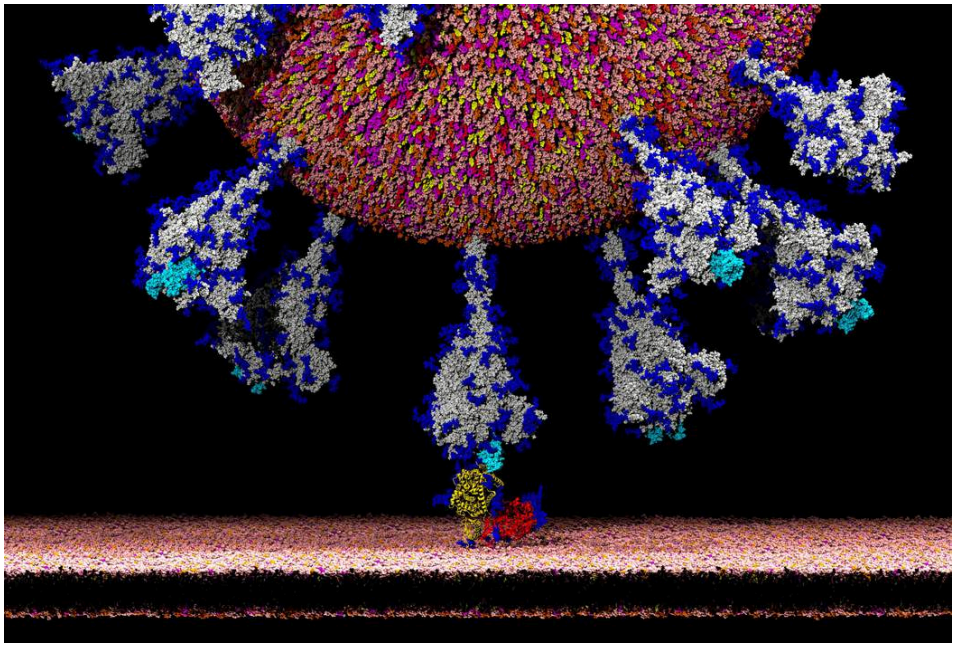
Task 4: find the iron atom (Fe)

- Does one part of the structure stand out from the rest? That could be the special molecule that lets myoglobin bind to oxygen in our body
- Can you find the iron atom by moving your mouse over it? Can you see how it is bound to the rest of the protein? What kind of atom is it bound to?

COVID Proteins

Spike protein

This picture illustrates the interaction between a virus particle and a cell membrane.



https://www.nytimes.com/interactive/2020/health/coronavirus-unveiled.html

It shows the **Spike** (S) protein (top) binding to a human protein **ACE2**.

First, let's look at the spike protein in more detail:
https://molstar.org/viewer/?pdb=6VXX&collapse-left-panel=1
This is in a **Closed** state. The protein **Opens**. To bind our cells.

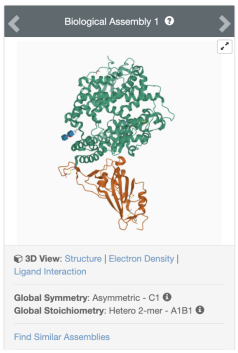
Here is a structure of the open state with the **D614G** mutation you may have read about in the news. It is believed to be more open.
https://molstar.org/viewer/?pdb=7BNN&collapse-left-panel=1

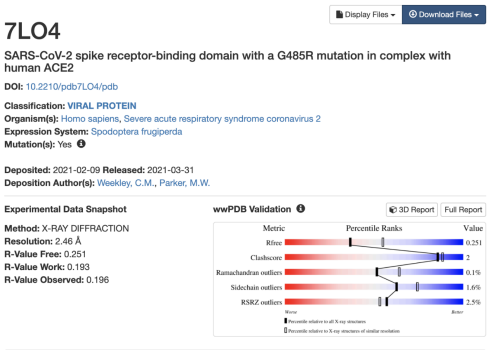
Can you see the difference?

This link may have the two structures on top of each-other to compare:
https://molstar.org/viewer/?snapshot-url=https%3A%2F%2Fwebchem.ncbr.muni.cz%2Fmolstar-state%2Fget%2Fe0277734-a7ec-4ee3-9790-fa90e514572d

The top section is the part that binds our cells.

Mutations in this area can make the virus bind better or worse, or help them avoid our immune system.





One example studied by my friend Claire is the G485R mutation. Glycine 485 on the spike protein is changed to arginine. *Let's look at the structure, what kind of effect could this have?*
https://molstar.org/viewer/?pdb=7LO4&collapse-left-panel=1

Variations in different animals' ACE2
Remember 2 week ago we learned that mammals have variations in their ACE2 proteins.

Here is a structure of the top of spike bound to ACE2
https://molstar.org/viewer/?pdb=6M0J&collapse-left-panel=1

Which amino acid locations do you think are most important for binding between the two proteins?

Prof. Arora told us that Q42E mutation (#42 changed from glutamine to glutamic acid) could affect the binding in American monkeys, let's see if we can see why.

SUPPLEMENT

What about amyloid diseases?

Prof. Lupoli discussed mis-folded proteins, and how they can **aggregate**. This may be connected to diseases such as Alzheimer's and Parkinson's.

Many of these proteins are **disordered** by themselves, and so we cannot get a crystal structure for them. Here is an example structure gotten by different methods:

https://molstar.org/viewer/?pdb=6SZF&collapse-left-panel=1

Here you can see different **models** of the same 42-amino acid peptide.

What happens when more than one protein comes together? Here are several possible model of an amyloid plaque:

https://molstar.org/viewer/?pdb=2MXU&collapse-left-panel=1

If we switch to a molecular resolution, can you see which atoms hold this together?