

[00:00] Welcome to Bioinformatics for Biologists. I'm Stevie Bain, a researcher from the University of Edinburgh.

[00:10] In this video, we are going to use the NCBI (National Center for Biotechnology Information) to visualise 3-dimensional protein structures. Firstly, we need to navigate to the NCBI homepage by typing `ncbi.nlm.nih.gov` into the address bar of your web browser.

[00:30] The NCBI homepage has a search bar that allows us to search the databases using a keyword search. In this video, we are going to focus on the 3-dimensional protein structure database. So, we click on the dropdown menu to the left of the search bar and select 'Structure'. Now we type the name of the protein whose structure we want to search for in the search bar. In this video, we'll type `deoxyhaemoglobin AND Homo sapiens [ORGN]`.

[01:00] The addition of [ORGN] after *Homo sapiens* means that we only want to search for deoxyhaemoglobin found in the species, *Homo sapiens*. It basically allows us to perform a more specific search of the database.

[01:30] We then click search, and in a few moments, we should see our results page. We can see that we have 106 results for deoxyhaemoglobin in *Homo sapiens* in this database. For each of these results, the line in blue is the description, it tells us the full name of the structure. And importantly, if we look underneath this we can see "taxonomy" which tells us what species this structure comes from.

[02:00] When using this database, it is important to take time look through the descriptions of results to ensure that you find the most suitable structure. You can always refine your search terms and search again if necessary. If we click on the description line of our chosen result it will take us to a page with more details. We can see the name of the structure and also some information about the scientific literature associated with this database entry.

[02:30] When we scroll down, we can see an image of the structure and also the molecular components. In this example, we have 4 protein chains: 2 that are Hemoglobin S alpha chains and 2 that are Hemoglobin S beta chains. Underneath the chain name, we have the gene symbol for each. If we click on this, we are taken to a page that gives us some info about the gene.

[03:00] For example, we can see the HBA1 gene is located on chromosome 16.

[03:12] On this page, we also have some more information about the chemicals and molecules that bind to these proteins. Here we can see that we have 4 molecules that contain iron - these are the haem groups. Once again, we can click on this description to get more information about this structure.

[03:40] If we want to look at this deoxyhaemoglobin protein structure in more detail, we can scroll back up and click full-featured 3D viewer. This may take a while to load but it's worth having patience as this feature will allow us to interact with a 3D view of our structure. Here we can hover the mouse over the structure to see which amino acids are in the protein.

[04:10] If we click down on the structure, we can also move it to focus on different parts. In this viewer, we can also very easily visualise secondary structures, for example, alpha helices are shown as curled ribbons. On the right-hand side, we can once again see each of the protein chains in this structure and the molecules.

[04:40] On the right-hand side, each of the protein chains and their associated conserved domains are represented by coloured boxes. If we wish to highlight any of these regions in the 3D viewer, we simply click on the coloured box and a yellow highlight will appear around that area. If we wish to remove this highlight, we simply go to the toggle at the top that says 'selection' and click.

[05:10] As you can see this selection has now be removed so we can click on a different protein chain.

[05:20] At the top of the viewer, there are a number of menus. One particularly useful menu is the 'color' menu. This allows us to colour the 3D structure based on particular properties or features. Right now, 'chain' is selected. That means that each of the protein chains in the structure is uniquely coloured. Let's change this to charge.

[05:50] Now we can very clearly see areas of the structure that have different charges: negatively charged areas are red, positively charged areas are blue and those that are grey are neutral. Once again, we can use the boxes at the side to make specific selections.

[06:14] We can also change the style of the structure. If we go to proteins, you will see that there are many options that allow us to change how the structure is styled.

[06:25] We can change this to lines for example or show the proteins as spheres. Let's go back to proteins and change their style to ribbon. Then we can then move along to the side chains option - which you will see right now is hidden - and change this to show the side chains.

[06:49] There are many options in this viewer, and we recommend that you have a go at exploring these yourself.

[06:56] We hope you found this video useful. If you would like some more information about our project and our free resources: [4273pi.org](http://4273pi.org) or [@4273pi](https://twitter.com/4273pi)