

Whitehead Bioinformatics and Research Computing - Demo

<http://jura.wi.mit.edu/bio/>

Analyzing the structure and function of cholera and E. coli toxins

Goal: Explore the structure of toxic protein complexes produced by cholera and E. coli and link their structures to mechanisms of action.

1 - Open DeepView / Swiss-PdbViewer.

DeepView / Swiss-PdbViewer is a free program from <http://ca.expasy.org/spdbv/>. This program reads "PDB" files (Protein Data Bank; at www.rcsb.org) that lists the 3-dimensional positions of all atoms of all molecules in a protein, protein complex (more than one protein that are bound together) or other structure. These files are the result of difficult experiments where a pure solution of a protein is crystallized and then analyzed using a technique like X-ray crystallography.

Part I. Cholera toxin

2 - Open the cholera toxin protein complex (after downloading from PDB):

File >> Open PDB File >> 1XTC.pdb

This file came from a Protein Data Bank page for cholera toxin:

<http://www.rcsb.org/pdb/explore/explore.do?structureId=1XTC>

3 - This toxin is produced by *Vibrio cholerae*, a bacterium that causes one of the most rapidly fatal illnesses known, killing some people less than 3 hours after infection.

4 - **Make the structure easier to see by going to the "Display" menu and checking "Use OpenGL Rendering" (if you're on Windows) and "Render in Solid 3D"**. This will give you a ball-and-stick version of the protein complex. If manipulating the structure is too slow on your computer, you can turn off "Use OpenGL Rendering" until you're ready to get a final view.

5 - **With the mouse, try clicking and then dragging over the molecule.** It should rotate.

6 - Since it's hard to tell what's what, try making each protein chain a different color. **Go to the "Color" menu and select "by Chain"**.

7 - **How many protein chains are there? _____ Count the colors.**

8 – Confirm your count of protein chains using the Control Panel. Open the Control Panel by going to the Window toolbar (Windows) or the Win Toolbar (Mac). The Control Panel lists each amino acid (or other molecule) by 3-letter code and its position (a number) in that chain, like "ASN1" for asparagine at position 1. The first column lists each chain by letter (starting with "A"), and the "col" column show the color of each amino acid. **To get the number of chains, you can either count the number of chain letters or count the number of colors (since you already gave the command to indicate each chain with a different color). Is this consistent with the structure you see? _____**

9 - **Look at the sequences of all chains in the crystallized cholera toxin at once by opening (double-clicking on) the file "cholera_1XTC_sequence.txt".** This sequence also came from a Protein Data Bank page for cholera toxin. For each chain, there is a description line starting with ">" followed by the protein sequence (following lines) showing each amino acid with a one-letter abbreviation. **Can you see anything in common between chains D, E, F, G, and H?**

10 - **Find these chains D, E, F, G, and H on the structure.** They should be colored green (D), red (E), gray (F), pink (G), and light blue (H).

11 - To get a clearer view of these 5 chains, we want to change from a ball-and-stick view to a ribbon view that simplifies the proteins into helices and beta-sheets, two common secondary structures of proteins. **To do so, use the Control Panel. Find the first amino acid in chain D (THR1 for threonine). In the column labeled "show", click the mouse on the "v"/checkmark symbol. This will hide (not show) that amino acid. Still holding down the mouse, drag it down the rest of the column to hide the rest of the amino acids in chains D - H. Go back to first amino acid in chain D (THR1). Under the column "ribn" (for ribbon), click your mouse, and a "v"/checkmark should appear (selecting this amino acid for a ribbon view). Still holding down the mouse, drag it down the "ribn" column to show these 5 chains as ribbons. Rotate the cholera toxin to get a clearer view of the 5 chains (the pentamer). Note that they form a doughnut-shaped ring. After you swallow cholera toxin, this donut-like ring binds the cells that line your intestines.**

12 - How about the rest of the toxin? **To get a space-filling view of the rest of the toxin, locate the last amino acid of chain C (the blue chain).** The end of chain C is labeled as OXT240, which just means the oxygen molecule at the end of a chain (part of LEU240, or leucine). **Under the "::-v" column (for van der Waals surface), check OXT240 and still holding down the mouse, drag it up, selecting all of chains A and C.**

13 - What happens if the cholera toxin gets inside your stomach and then intestines? While on the surface of a human cell, chains A and C actually bound together as one long chain. After the pentamer ring (chains D - H) binds to an intestinal cell, the wedge-shaped chain AC is "injected" into the cell (perhaps thanks to a hole in the membrane made by the pentamer ring). Once inside the intestinal cell, chain AC is split into separate chains. Chain A (the yellow chain; the main part of the wedge) activates a molecule called cAMP. This makes a lot of changes in the cell, and as a result, the cell pumps out a lot of water into the intestine. This leads to severe diarrhea, further leading to dehydration (loss of lots of water and salts) and death (unless this water and salt are replaced).

14 - **To zoom in on any region you want, click on the 'two rectangles' icon near the top left of the DeepView toolbar. To move without rotating the protein complex, try selecting on the 'hand' icon next to the 'two rectangles' icon.**

15 - Does this have anything to do with E. coli infections? See the next section....

Part II. E. coli enterotoxin

16 – **At the top of the Control Panel, uncheck the “visible” box to hide the cholera toxin structure.**

17 – **Open the E. coli enterotoxin protein complex (after downloading from PDB):**

File >> Open PDB File >> 1TII.pdb

This file came from a Protein Data Bank page for E. coli enterotoxin:

<http://www.rcsb.org/pdb/explore/explore.do?structureId=1TII>

18 – **If the top of the Control Panel shows “cholera_toxin_1XTC”, click on the name and select “E_coli_entertoxin_1TII”.**

19 – This toxin is produced by dangerous types of E. coli – not the same as the “normal” E. coli that naturally lives in our intestines and helps us digest some foods. These pathogenic types of E. coli are sometimes in the news for causing intestinal problems, diarrhea (sometimes called “traveler’s diarrhea”) and can, in severe cases, be fatal.

20 - Since it's hard to tell what's what, try making each protein chain a different color. **Go to the "Color" menu and select "by Chain".**

21 - **How many protein chains are there? _____ Count the colors.**

22 - **Look at the sequences of all chains in the crystallized cholera toxin at once by opening the file "cholera_1XTC_sequence.txt".** This sequence also came from a Protein Data Bank page for cholera toxin. For each chain, there is a description line starting with “>” followed by the protein sequence (following lines) showing each amino acid with a one-letter abbreviation. **Can you see anything in common between chains D, E, F, G, and H? If so, what?**

Part III. Comparing the cholera toxin and the E. coli enterotoxin

23 – To generate a side-by-side structural comparison of these two toxins, start by clicking on “cholera_toxin_1XTC” at the top of the Control Panel and select “cholera_1XTC”. Now check the “visible” box. You should be able to see both structures.

24 – [Optional] To generate similar side-by-side views the challenging way, use a combination of the structure selection at the top of the Control Panel (that you just used to change the active structure), the “visible” and “can move” checkboxes at the top of the Control Panel, and the moving icons (the hand and the 2 icons to the right).

25 – To generate similar side-by-side views the easy way, go to the “Fit” menu and select “Magic Fit”. This should superimpose the two structures by trying to minimize the distance between atoms of each structure. Uncheck “can move” for one toxin, click on the “hand” icon, and slide one toxin next to the other. You may need the adjacent 2 icons to get both toxins on the screen.

26 – Why is cholera toxin so much more dangerous to humans than the E. coli enterotoxin? One reason is thought to be differences in the ends of chain C that face the cell when the pentamer donut binds the cell surface. **Selecting the cholera toxin structure, find amino acids 232 – 240 of chain C (blue) and mouse-over the blue-colored squares. When a “Color” pop-up window appears, choose another color (like red) to indicate these last 10 amino acids.** Unfortunately, the 3-dimensional shape of the corresponding region of the E. coli toxin was impossible to determine, so it’s been left out of the E. coli structure.

27 – Where is the active site of the cholera toxin? Once chain AC enters the intestinal cell and splits apart, chain A (yellow) is involved in the enzymatic reactions via its active site. These amino acids were identified by mutational analysis; if they were mutated to another amino acid, activity of the downstream chemical reactions decreased. **The active site is somewhat hidden, so to start, select the cholera toxin structure and remove the space-filling view of all of chain A by unchecking (sliding over with your mouse) the “show” column next to all amino acids of this chain. Then go back and check the following amino acids that make up the active site region:**

ARG7, ASP9, HIS44, SER61, SER63, VAL97, TYR104, PRO106, HIS107, GLU110, GLU112, SER114

Note that these amino acids are not next to each in the linear protein sequence but they are close to each other in the 3-dimensional protein structure.

28 – **Make the same changes with these amino acids to the E. coli structure which can be aligned to the above cholera amino acids: ARG5, ASP7, HIS42, SER59, THR61, VAL95, TYR102, PRO104, TYR105, GLU108, GLU110, ALA112**

29 – **Admire your final structural comparison of the cholera and E. coli toxins!**