## Laboratory 1:
Visualization and Measurement of Protein Structures using VMD

Nathan Hammond
nhammond@mit.edu

Issued: Wednesday, February 11, 2009
Due: Friday, February 20, 2009

Whether or not you do molecular simulations, it is often useful to be able to make high-quality graphical displays of a proteins to illustrate something about their structure. This type of graphic is common in textbooks, research articles, and presentations.

For example, here is an image from a publication by the Michael Sheetz research group on the binding of two proteins, talin and vinculin. These proteins are important in a cell's attachment to an external surface. This study is experimental, not computational, but the graphic helps to illustrate what is happening in the experiments. Below we see a segment of the talin molecule that contains several vinculin binding sites that are buried in the protein structure. Armando Del Rio wishes to illustrate how they have applied a force to the N and C termini in order to partially unfold talin and expose the vinculin binding sites located in α-helices 4, 6, 9, 11, and 12.
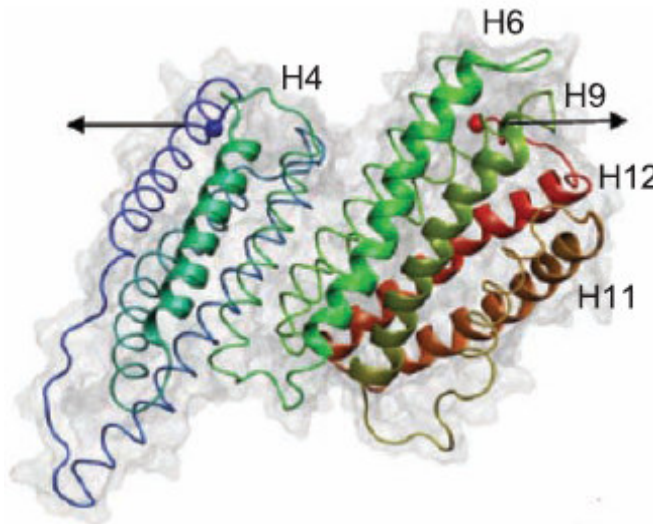
Armando del Rio, et al. 2009. Science 323:638-641.

In this lab you will:

1. Download the protein structure from the computer cluster that we will be using throughout Module 2 of the course
2. Recreate the image above
3. Download the structure file for a new protein of your choice from the Protein Data Bank
4. Use the downloaded PDB file to create your own graphical illustrations

Although I've set up the user accounts as teams, I encourage you to do this lab yourself. Of course talking with team members and classmates is encouraged—I just want you to get your own hands on the controls as much as possible.

Step 1: Obtain the structure of Talin (residues 482-889) from where it is stored on the cluster.

Access the cluster using SSH protocol

> To log into a Linux cluster and execute commands remotely, we will use a protocol called SSH. You will need to know your username, your password, and the host name or ip address of the cluster.
>
> - Your username is the lowercase word "team" followed by your team number, e.g. team1 or team5
> - Your password is the course name followed by your team number, all lowercase, e.g. engr3699a1 or engr3699a5
> - For the host name, use the ip address of the cluster, which is 67.202.16.75 (Note—this number is going to change later, but I'll let you know)
>
> *Linux users*
> If you are already on a Linux machine, you can log in as follows:
> ssh *username@hostname*
> Enter your password when prompted.
>
> *Windows users*
> You will need a program called PuTTY or another SSH client. Download putty.exe from
> http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html
> You don't need to install it. Just run the executable each time you need to log into the cluster.
> In the box labeled Host Name enter the following:
> *username@hostname*
> You may wish to click "Save" so that next time you run PuTTY you can load the saved settings.
> Connect to the cluster by clicking "Open" and entering your password when prompted.
>
> *Mac users*
> Open a command terminal. (Go to the "Applications" folder. Under "Utilities" open "Terminal".)
> Enter the following command:
> ssh *username@hostname*
> Enter your password when prompted.

> *All users*
> If this is your first time logging in, you should receive a prompt asking whether to trust the machine you are logging into. Type yes and press enter.
>
> After logging in, change your password by entering the command "passwd" and following the prompts. Make sure everyone on your team knows the new password.

Locate the files and copy them to your home directory

> You should now have a Linux shell prompt, something ending with a # or a $. From here, you can enter Linux shell commands to manipulate files. You can search for bash shell commands in Google and find whatever you need, but some of the basic commands are listed below.

ls	List files in the current directory
cd *directoryname*	Change directory
cd ..	Go up one directory
cd ~*team1*	Go to the home directory of *team1*
emacs *filename*	Open a file in the Emacs text editor
Ctrl-x Ctrl-s	Save a file in Emacs
Ctrl-x Ctrl-c	Exit Emacs
cp *file1 file2*	Copy *file1* and save it as *file2*
cp –r *directory1 directory2*	Copy *directory1* and its contents to *directory2*
pwd	Print working directory
mv *filename directory*	Move *filename* into *directory*
mv *file1 file2*	Rename *file1* as *file2*
rm *filename*	Delete a file
rmdir *directory*	Delete a directory
mkdir *directory*	Make a directory
more *file1.txt*	Display the contents of *file1.txt* page by page
grep "*keyword*" *file1*	List all the lines in *file1* that contain *keyword*
grep "*keyword*" *file1 > out.txt*	Write the results to the file *out.txt*
ps –u *team1*	Show processes being run by *team1*
kill -9 *pid*	Kill the process with process id number *pid*. (The 9 refers to a cat's 9 lives, meaning that this command will *really* kill it no matter what. Ah, Linux humor.)
exit	Close the SSH connection (logout)

> After you log in, locate the files you will need for this lab in a folder called "/share/lab1". The file containing the structure of talin (residues 482-889) is there, saved as "1XWX.pdb".
>
> Copy these files to your home directory. For example, if your home directory is /home/team1, you might use this command:

cp -r /share/lab1 ~team1

Within your team's home directory, you might wish to make your own directory to do individual work.

Download the file using an FTP client

You'll need an FTP client to transfer files between the cluster and your computer. If you don't have one already, you can download and install the FileZilla client from here: http://filezilla-project.org/

After you run the FTP software, enter the host (ip address) username and password that you used for SSH, and click on "QuickConnect". (It should default to Port 21.) You can navigate the directories and then transfer files to or from the cluster by dragging and dropping.

Download the structure of talin(482-889), in the file "1XWX.pdb". Get used to PDB files. These contain protein structures, usually from the Protein Data Bank.

Step 2: Create an image using the Visual Molecular Dynamics software

Download and install VMD

If you only remember one thing from this lab, remember this: *When you want to make cool graphics of a protein, use VMD.* It stands for Visual Molecular Dynamics, and it's free.

Download VMD from here:
http://www.ks.uiuc.edu/Research/vmd/
Look under "Download VMD" and under the newest version, pick the file that corresponds to your computer architecture (probably "Windows OpenGL"). You might have to register to get this download, but as near as I can tell they won't send you any junk mail. At least I haven't received any.

Load the protein structure

Start VMD and open the file "1xwx.pdb". ("File→New Molecule", "Browse". Make sure the file type is set to PDB and click "Load". After this, you can close the "Molecule File Browser" window.) You should see a graphical representation of talin amino acids 482-889, the segment studied in Armando Del Rio's paper.

Zooming, rotating, translating, and creating labels in VMD

Some useful commands for manipulating the image are in the table below. You
need to have clicked on the Display window for these commands to work.

s	scale. After pressing "s" click on the image and drag left or right to scale in and out.
r	rotate. Click on the image and move the mouse to rotate the molecule.
t	translate. Click on the image and move the mouse left, right, up, or down, to move the molecule.
1	Select an atom. Text will appear describing the atom. "ASP695:CA" means I have selected the alpha carbon in residue number 695, which is an aspartate
2	Measure the distance between two atoms (in Angstroms). After you click on two atoms you will see a line between them and the measured distance.
3	Measure an angle by selecting 3 atoms
4	Measure a dihedral angle by selecting 4 atoms

The labels created with 1, 2, 3, and 4 can be deleted by opening
"Graphics→Labels" in the "VMD Main" window.

Graphical Representations

Let's look at the graphical styles in which we can display a protein. Open
"Graphics→Representations". In the "Draw Style" tab and under the label
"Drawing Method", change from "Lines" to "Licorice" and notice the change.
Now try "New Ribbon". This more clearly delineates the protein's backbone
structure. Now let's choose "New Cartoon". This represents α-helices as spirals
and ß-sheets as flat arrows. In our image you should see the 12 α-helices and no
ß-sheets.

The 4th, 6th, 9th, 11th, and 12th helices each contain a vinculin binding site. We'll
highlight these by making them in the "New Cartoon" style while showing the
rest of the protein backbone in "New Ribbon".

Let's display just helix 4, which comprises residues 605 to 627. In the "Graphical
Representations" window, under "Selected Atoms" write "resid 605 to 627" and
press enter. Make sure "Display Mode" is on "New Cartoon".

Now we'll include helix 6 as well, residues 662 through 692. Change the
selection criteria to "resid 605 to 627 or resid 662 to 692".

Use more "or" terms to show helices 9 (761-787), 11 (819-846), and 12 (850-879).

Let's add some color. In the "Graphical Representations" window under
"Coloring Method" in the "Draw Style" tab we have a number of methods. Try a

few. Choose "Index". This creates some contrast between the helices, but not as much as we would like. Instead, we will color each helix individually.

Narrow the atom selection to only include helix 4. Under "Coloring Method" in the "Draw Style" tab, select "Color ID" and choose color 10, cyan.

In the "Graphical Representations" window, click "Create Rep". This creates a new graphic object whose properties we can control independently. Some of the properties are shown in a list in the "Graphical representations" window. For now there should be two, and they should both be labeled something like "NewCartoon ColorID 10 resid 605 to 627"

Make sure the second representation is highlighted. Change the atom selection to helix 6, and make the color green.

Add new representations for helices 9, 11, and 12, and make them yellow, orange, and red, respectively.

We can hide or redisplay any one of the objects by double clicking on it in the list. Try it out.

Create one more representation to select all the residues not included in the helices that we've already created. You can use this selection term:
"all not (resid 605 to 627 or resid 662 to 692 or resid 761 to 787 or resid 819 to 846 or resid 850 to 879)"

Change the drawing style to "New Ribbon". Set the coloring method to "Pos". Let's de-emphasize this structure a little bit by making the ribbon thinner. Reduce "Thickness" to about 0.2.

There is just one last image to add. Notice the light gray in Armando Del Rio's image showing the space filled by all the atoms in the protein. Create a new representation with atom selection "all". Set the drawing method to "Surf" (it may take a minute to calculate the surface), and set the color to ColorID silver. Under "Material" select "Glass1".

Finally, we'll make the background white and remove the axes. In the "VMD Main" window, open "Graphics→Colors". Highlight "Display" and "Background", and select "White". Now you can close the "Color Controls" panel. In "VMD Main" choose "Display→Axes" and set it to "off".

Rendering an image

Take a look at your image. Not bad, is it? Rotate, translate, and zoom to the orientation that you want, and render a .bmp file under "File→Render". Don't change the destination directory, because there is a bug in the program and it

won't be able to save. Just leave it set to "snap.bmp" or enter a new file name with no directory path. For your convenience, this will save the image in a hard-to-find location such as "c:/Program Files/University of Illinois/VMD/snap.bmp" and probably also open your image in Microsoft Paint.

Here's what I got:
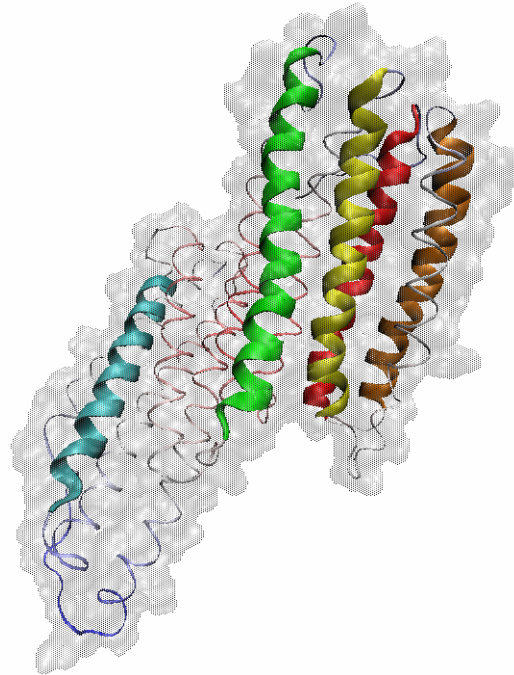
At 100% zoom this image looks fine, but as I zoom closer it starts to look pretty course. For that, you can play with some of the resolution options in VMD before you render the image. When needed, labels and arrows can be added in Paint or some other graphics program.

Save!!!

We may want to make minor changes to this image later, and we don't want to start from scratch. Let's save what we've done. Under "File" choose "Save State". Now we can retrieve our work by first loading the PDB file, then going to "File→Load State" and opening the VMD file we just created. You may have to redo minor things with the display, like turning off the axes.

Step 3: Find and download a protein structure from the Protein Data Bank (PDB)

Don't quit now just because you're tired! Fight the entropy in your brain! *Energize!* You'll need to know how to find protein structure files and create graphics for your final report. Besides, getting files from the PDB is easy. If you already have some idea of your project, you can even use the graphics you create now in your final report.

I'm going to step you through searching for the structure of talin. But you should search for your own protein.

Go to www.pdb.org

In the search bar, I enter "talin" and search as "PDB ID or keyword". I find 30 structures, each with a unique PDB ID such as "1u89". (You can also search by PDB ID if you happen to see it referenced in a publication. The image we used in this lab has ID 1xwx.) These may be full talin molecules, or fragments of talin, and they may come from a range of species. I'll have to sort through them to decide which one is best to use.

I click on the first result, 1u89, "Solution structure of VBS2 fragment of Talin." This structure contains a fragment of talin comprising amino acids (residues) 755-889. Under the "Structure Summary" tab I see information such as the authors and reference of the article that first presented this structure, the date it was filed with the protein data bank, the method used to obtain the structure (typically X-Ray crystallography or NMR), data on the polymer chains in the structure and the molecular weight, the source (in this case, a mouse gene expressed in E. Coli), and related structures in PDB.

On the right side of the window under "Display Options" I can click on one of the programs, such as "JMol," and it lets me view and rotate the molecule in 3D.

Other information such as sequence information and details of the experimental procedure is available under the other tabs.

To download the structure file (a PDB file), I click on "Download Files" in the menu at the left of the screen. Left click on "PDB text" and save the file to your computer.

Go through your own search results and decide which structure is most relevant. *For many proteins no structure exists, in which case there is no good way to do molecular dynamics.* However, there are many new structures being published every year, presenting the opportunity to do molecular dynamics studies on molecules that have never been studied before.

Step 4: Create an image with your protein structure

Open your new protein structure in VMD. Play around with VMD and explore its capabilities. Decide what you would like to illustrate in your protein and create a nice-looking graphic. Try to create an image that you can use in your final paper. Remember to "Save State" so that you can make adjustments later.