**BIOS 10603 Lab 7**

**Timeline Trajectory Analysis and Additional Protein Analysis**

**Name:**

Please insert all answers and figures into this Word document, and then save the completed lab as “Lastname\_Lab7.docx”. When you have completed the lab, submit it as an attachment to the Lab-7 assignment link on Canvas. The report is due by midnight the day after your lab period.

**Objectives: Learn how to use the VMD "Timeline" tool for simulation analysis.**

**Lab 7 Sections:**

1. Introduction to Timeline
2. Analysis of your own protein
3. Further analysis using the VMD scripting library

**Part 1: Timeline introduction**

Last week, you submitted your protein for the production simulation, which we will analyze in this lab. Remember, this is a critical stage in the process: please ensure that your simulation is free of errors at this point. As Dr. Haddadian told you individually, for your final project, you need to have two **40-nanosecond** simulations: one for your control and one for your alternate system.

In this lab, we are going to use VMD’s “Timeline” tool to analyze your protein’s trajectory. Timeline is a very useful tool that displays and analyzes how different aspects of the change of a protein’s secondary structure throughout a Molecular Dynamics trajectory.

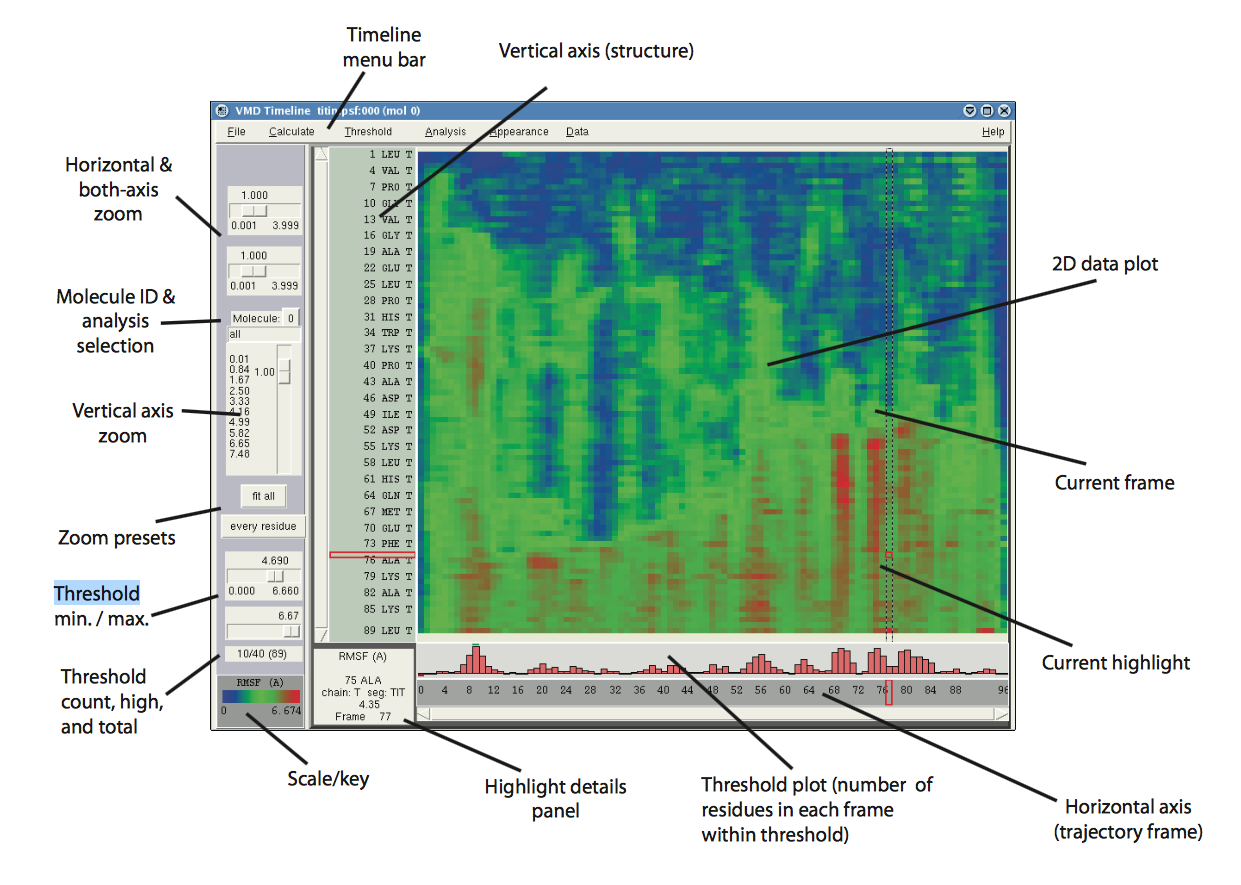
**We will first analyze the provided Titin trajectory, and then you will apply these ideas to your own protein simulation.** For more information about Timeline, please visit the following links:

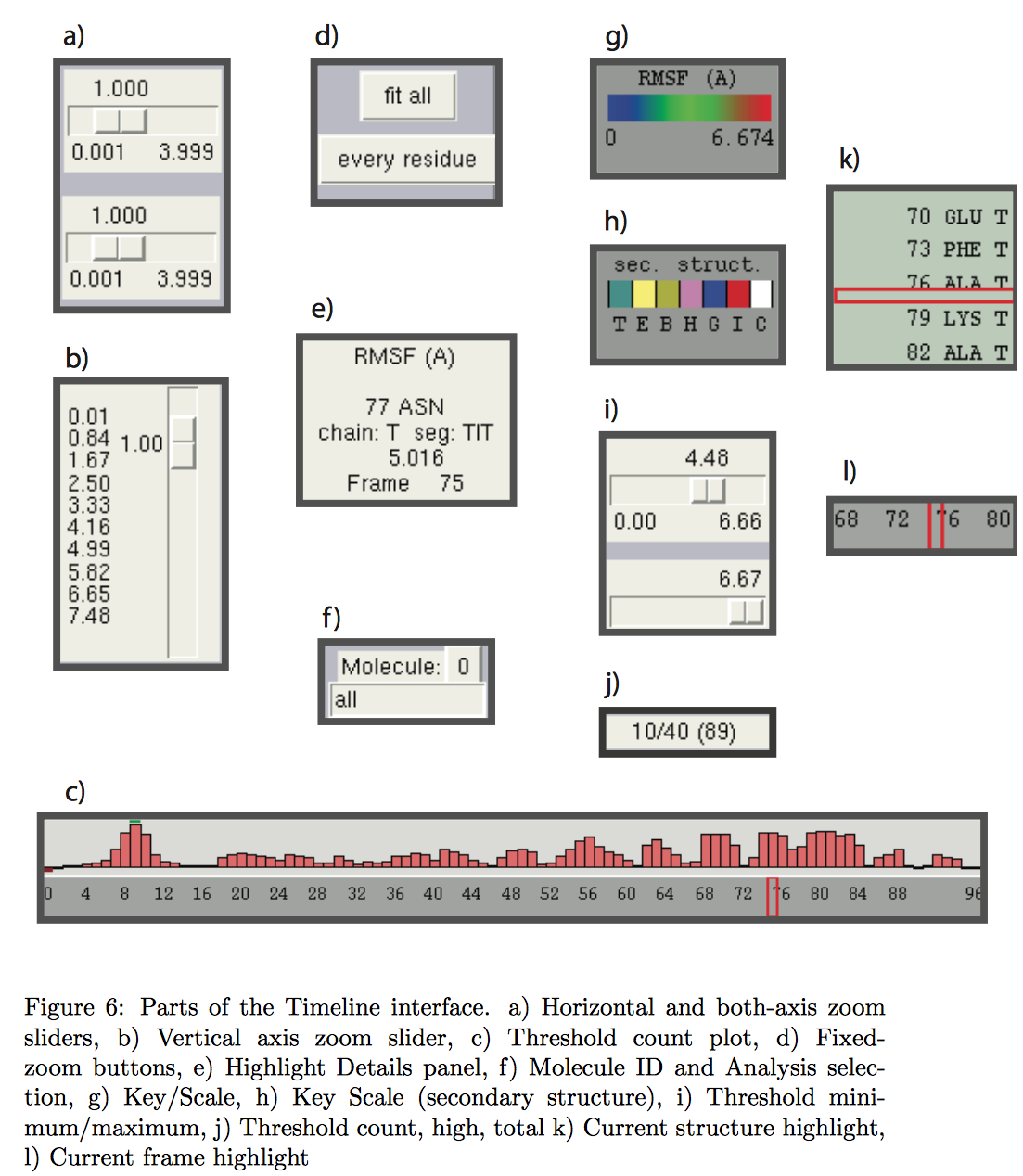
<http://www.ks.uiuc.edu/Training/Tutorials/science/timeline/tutorial_timeline.pdf> or

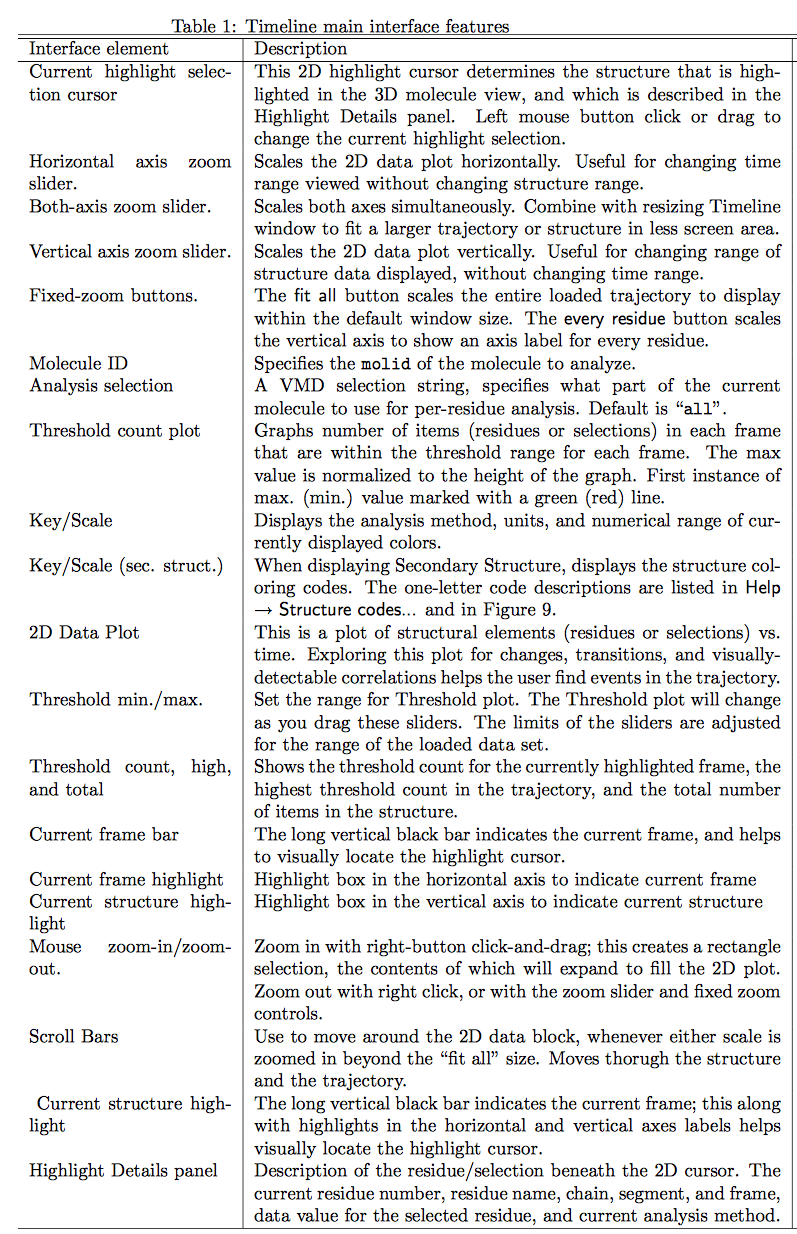
<http://www.ks.uiuc.edu/Research/vmd/plugins/timeline/>

The general description of the Timeline interface is shown below (taken from the Timeline tutorial).

**Please review the following TimeLine window and table very carefully before you continue.**







**Titin’s trajectory analysis using the Timeline tool**

On Canvas, you can find the PSF and DCD files for a titin trajectory. **Titin is a very large protein found in sarcomeres, which are responsible for** [**stretching and contracting muscle**](https://www.youtube.com/watch?v=8TZgf1zaVgU)**. Domains in titin unfold when the protein is stretched and refold when the tension is removed.** This particular trajectory is a single domain of titin. Load the PSF file into VMD, and then load the DCD file into the PSF.

**Q1.** Watch the trajectory a few times.

1. What is happening in the simulation of this titin domain?
2. Why might biologists be interested in this simulation?

**A) Secondary structure analysis**

Open Timeline by going to Extensions 🡪 Analysis 🡪 Timeline in the VMD Main window. A new window will open. Click on the “Calculate” drop-down box and select “Calc. Sec. Struct”. After a moment, you should see a plot with a lot of teal and yellow. There is a key at the bottom left corner of the timeline window, where each letter stands for:

|  |  |
| --- | --- |
| **Letter** | **Secondary Structure** |
| T | Turn |
| E | Extended conformation (beta sheet) |
| B | Isolated bridge |
| H | Alpha helix |
| G | 3-10 helix |
| I | Pi helix |
| C | Coil |

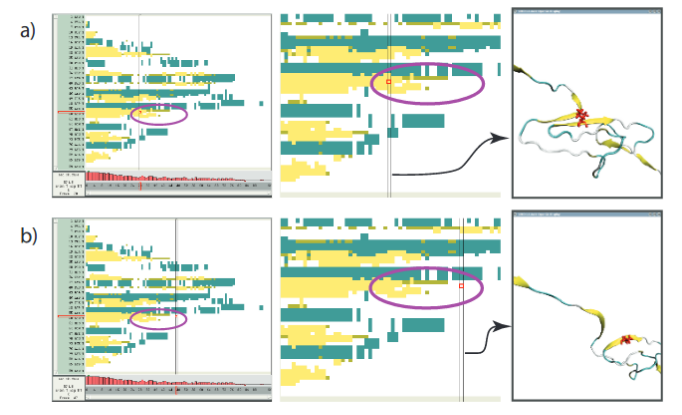
The two bars on the top left side of the Timeline window control horizontal and both-axis zoom, and the vertical bar controls the vertical zoom. You can always reset the zoom by selecting the “fit all” button. **Clicking anywhere on the timeline 2D plot with your left mouse button selects a particular residue from a particular frame. Information about the selection is displayed in a box at the lower left of VMD. The selection is also displayed in the Graphical VMD window, highlighted in red.**

**Q2.**

1. What are the axis labels in the Timeline window for secondary structure analysis?
2. Using Timeline, locate residue 28 at frame 66, take a screenshot of this residue in the VMD display window (not timeline), and attach it below.
3. What amino acid is this residue?

You can also click and drag with your left mouse button in Timeline to select a different residue or watch the action of a residue as the trajectory goes on. You can click and drag with your right mouse button to select an area to zoom in and click with your right mouse button to zoom out. Using these controls, watch the protein trajectory again.

In the image below, you can see that in (a), the purple region still shows a -strand secondary structure, while in (b), this secondary structure is lost, and the -strand turns into a coil. This can clearly be correlated to visual analysis of the same residues over the time frame, as seen in the image on the right (note the loss of secondary structure).



VMD assigns the secondary structure (-sheets and -helices) in the first frame and keeps the same assignment through the trajectory (to save on memory) even though the secondary structure may change. **This is misleading and needs to be corrected.** Download the sscache.tcl script from Canvas and *source* it in the VMD TK console (this script updates the secondary structure assignment for each frame). Then run the script with the following command:

start\_sscache molid (molid is the number assigned to the protein by VMD, found in VMD main)

**It would be best to show the protein with “new cartoon” and color it by “secondary structure” before you check the trajectory.**

**Q3.**

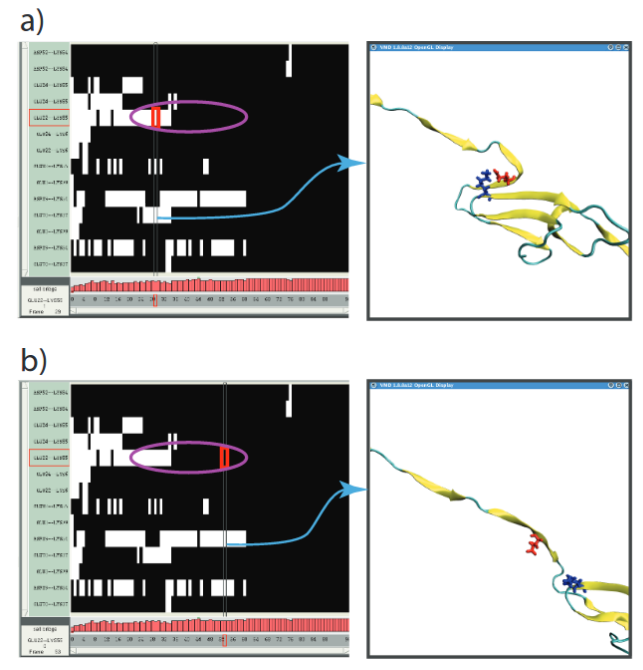
1. Where in the protein do the β-sheets unfold earliest (provide the residue numbers), and where in the protein do the β-sheets unfold latest (provide the residue numbers)?
2. Why do you think this is the case?

**Q4.** Observe, using the timeline plot, the loss of secondary structure in residue 22.

1. Attach pictures (just as above) showing the residue right before and right after the moment, the secondary structure was lost.
2. List the frame number when this unfolding event happens.

**B) Salt bridge Analysis**

Salt bridges are interactions between a pair of positively and negatively charged residues that help to stabilize the 3D structure of the protein. To calculate the salt bridges, reset your selection to “all,” and go into the “Calculate” menu, and select “Calc. Salt Bridges”. Use the default distance cutoff value of 3.2 Å between sidechain atom O of one residue and atom N of the other residue. This plot is slightly different from the ones you have seen previously. **Instead of showing every residue on the y-axis, it shows lots of individual salt bridges on the y-axis.** However, the controls remain the same.

In the image in the right, you can see an example of a salt-bridge breaking. In the purple region, we can identify directly on the timeline plot where this break (the distance between the O N atoms becomes larger than 3.2 Å) happens, and then, by visual inspection of the trajectory we can see the salt-bridge breaking.

**Q5.** Select the ASP29-LYS54 salt bridge, and watch it break by clicking the corresponding row in Timeline and dragging left and right.

1. Between which two frames does the salt bridge break for the final time?
2. Show these two residues with licorice in VMD and attach a screenshot of the salt bridge before and after breaking.

**Q6.**

1. Observe the salt bridge between GLU22-LYS55 (the top one) in the Timeline calculation window. Can you correlate the salt-bridge breakage with change in secondary structure (similar to the figure in the right)? **Explain your answer**.
2. From what you know about different types of amino acids, does it make sense that glutamic acid and lysine would form a salt bridge? Why?

Slide the upper threshold slider to a value of 0 and slide the lower threshold slider to a value of 1. The sliders can be found below the “Every residue” button. These are the minimum and maximum values for the count plot that appears below the main graph. Below the two sliders controlling threshold are three numbers of the form “X/Y (Z)”. X denotes the number of salt-bridges in the frame that you are currently selecting. Y denotes the maximum number of salt-bridges in any one frame, while **Z denotes the total number of different salt bridges formed throughout the trajectory.**

**Q7.**

1. Why are the salt bridges formed in Frame 75 and 76 nonexistent at the beginning of the trajectory?
2. What residues are involved in these salt bridges?

**Q8.**

1. What is the maximum number of salt bridges formed at one time, and in what frame?
2. What is the total number of different salt bridges formed throughout the trajectory?

A picture containing vector graphics, light

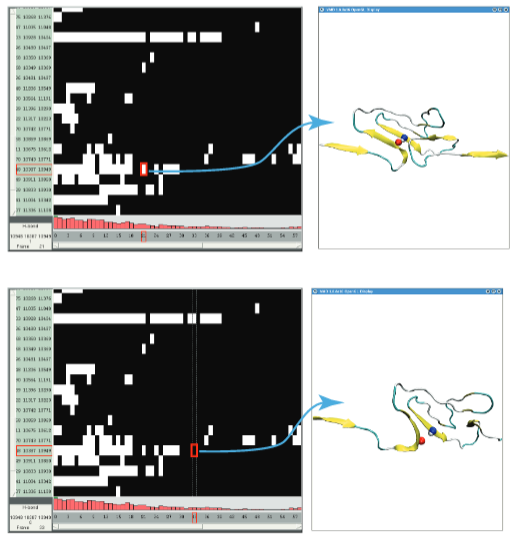
Description automatically generated**C) Hydrogen bond analysis**

Now we will look at the hydrogen bond breaking pattern during the titin extension trajectory. Since we are interested in the protein’s form, we will examine the bonds between the partner β-strands that make up the two β-sheets comprising titin's **β-sandwich domain**. Protein domains are structurally independent units that each have the characteristic of a small globular protein and have specific functional roles

<https://en.wikipedia.org/wiki/Beta-sandwich>

Change the Analysis Selection from all to “name *N HN O*” (the box appears below the molecule selection drop-down) to select only the **backbone** atoms involved in inter-β-stand hydrogen bonds. Calculate the hydrogen bonds that can be formed in this selection in Calculate → Calc. H-bonds. Enter a Bond distance cutoff of 3.5 and an Angle cutoff (deg) of 35, and then click the Calculate button.

The value of data in the Timeline 2D plot is either 0 or 1. Set the Threshold Plot range from 0 to 1 to see the number of formed hydrogen bonds in each frame. (Although the auto-ranging is already set from 0 to 1, you must still adjust the threshold minimum control once to display the Threshold Plot).

****

**Q9.** Zoom in on the Hydrogen bond plot to examine individual hydrogen bonds. Click on the atoms to display the names and numbers of the residues involved.

1. Name 3 titin **residue pairs** involved in Hydrogen bond formation at the beginning of the trajectory that break by the end of the trajectory.
2. Attach a screenshot of any hydrogen bond and correlate its breakage with some change in secondary structure (similar to the example provided below on the right). Hint: You may find it helpful to create a visual representation of the hydrogen bonds in VMD. First, make a “new cartoon” representation of your protein. Next, add another representation with “name CA C N O HN” and show it with licorice. Finally, add another representation by choosing “HBonds” as the drawing method, choosing the distance and angle cutoff as above, changing the color, and setting the line thickness to 5 for better visibility.

**Part 2: Your Protein Analysis**

Navigate to your protein’s directory on Midway and check that your job actually ran to completion. To do this, use the command: tail XXXX-mini-equil-npt1(or 2).log

If the output of the message contains “End of program” at the bottom, then your simulation has been completed successfully! **Otherwise, you will need to contact Dr. Haddadian or your TA about how to restart your simulation. You can also refer to question 12 in lab 5.**

Just as you observed last week, if your simulation is completed successfully, there will be files ending with .old and .BAK. You may remove these files, as they are unnecessary. **Carefully** enter the following line (making a mistake could delete all of your work!!)*:* rm \*.BAK \*.old

Download your latest NPT DCD file and load it into VMD **with a stride of 5**. If your VMD crashes while attempting to load the DCD file, try again, but change the “Stride” setting to 10. If your VMD timeline analysis for your protein crashes or take too long on your computer, use [**Midway ThinLinc to run VMD**](https://midway2-login1.rcc.uchicago.edu/main/)

**(See instructions at the end of the lab)**

**Q10.** **Hydrogen Bonds (“H-Bonds”):** Students working on antifreeze protein, please consult with Dr. Haddadian.

Using the H-Bonds tool built into VMD (Extensions 🡪 Analysis 🡪 Hydrogen Bonds), find the number of hydrogen bonds in your **protein backbone** over the course of the simulation. To achieve this, choose atoms “name N O” in the selection box and use a Donor-Acceptor distance of 3.5Å and an Angle Cutoff of 35 degrees.

1. Explain any trends you observe and try to correlate them to a visual analysis of the trajectory.
2. Attach a snapshot of your H-bond graph.
3. Why are the backbone H-bonds more important than the side-chain H-bonds?

**Q11.** Repeat the hydrogen bond analysis for your alternate simulation as above (either NPT1 or equilibrium simulations).

1. Explain any trends you observe and try to correlate them to a visual analysis of the trajectory.
2. Attach a snapshot of your H-bond graph.
3. Do you notice any differences in the number of H-bonds and patterns between your two trajectories?

This hydrogen-bond plot you obtained is interesting, **but we can get a lot more information out of Timeline.** Replace “all” with “backbone” in the selection box on the left side of the Timeline window. Now, in the “Calculate” menu, select “Calc. H-bonds”. Below the “every residue” button on the left side, there are two sliders controlling the minimum and maximum threshold for a bar plot. H-bonds can either exist or not exist, so the numerical values are either 0 or 1. Slide the lower slider all the way to the right, to a value of 1.00. You should see bars appear in the plot below the Timeline window.

**Q12.**

1. Attach the output plot for your initial and alternate simulation.
2. Does this plot agree with the plot that you created in questions 10 and 11? Explain.
3. How is the Timeline plot more useful?

**Q13.**

1. Using Timeline, calculate the secondary structure evolution for **both of your simulations**. Attach a figure and explain any interesting trends you observe (**We recommend using Midway for this as it is a rather computationally expensive calculation**).
2. Do you see large secondary structure changes (for example, alpha-helix or beta-sheet unfolding events)? Why or why not? Explain your answer.

**All of the above analysis tools can be focused on the important parts of your protein (such as an active site) by defining the residue range. This could be helpful for your final report.**

**Part 3: Further analysis using the VMD scripting library**

Looking at the distance between residues in the protein throughout the trajectory is another good way to tell how our protein’s structure changes over time. We will be using the ca-dist.tcl script from the VMD script library to measure the distance between the CA’s of all residues.

This script works by first calculating the distance between each of the residues in the selection. Then it maps those distances onto colors (basically assigns each distance value to a color). Finally, it plots the colors on a grid where the row indices and column indices represent residues.

For instance, the element on the grid with index (1, 2) will represent the distance between CAs of the first and second residues in the **Titin trajectory**. For more information about the script check

<https://www.ks.uiuc.edu/Research/vmd/script_library/scripts/ca-dist-v2.0/>

**Delete your protein trajectory and load the Titin trajectory**

**Q14.** Source the script in the Tk Console and create a selection with just the protein called sel. Run ca\_dist $sel in the console. This should create a grid in the display window, as well as create a new molecule for the grid (be sure to press “=” after running the script to ensure that you can actually see the grid).

1. Can you discern which color maps to a distance of 0?
2. What color maps to the maximum distance between two residues?

**Q15.** Run the ca\_dist script at 3 different frames in the simulation (once at the beginning, once in the middle, and once at the end)

1. Attach screenshots of the 3 plots below.
2. Describe any patterns you see in the three plots as well as how the patterns change between the plots.
3. How do these patterns correspond to what we actually see in the titin trajectory? (NOTE: when you want to rerun the ca\_dist function at a new frame, you will need to first delete the previous ca\_dist pattern in the VMD main window, select the new frame in the main window, and then rerun the script).

**Q16.** There is also a built-in tool in VMD to analyze residues-residue distances

( <http://www.ks.uiuc.edu/Research/vmd/plugins/contactmap/> ).

Go to Extensions > Analysis > Contact Map. Select the frames that you want to calculate the distances for in the main window (choose the same three frames as in Q15), then select Calc. res-res Dists under “Calculate.”

1. Attach screenshots of the 3 plots below.
2. Do these plots agree with the ones that you got in Q15? Explain.

**Timeline using Midway ThinLinc**

Timeline is very good at making 2-dimensional plots of a variety of different metrics for simulation analysis. Some of these metrics can be computationally expensive. You should try to run them on your own computer, but in case your VMD crashes while plotting the graph, you can use Midway to run the analysis. The RCC allows Midway users to log in to a virtual machine for Midway that runs with the Scientific Linux distribution. To use that, you can access Midway via

https://midway3.rcc.uchicago.edu

or

https://midway2.rcc.uchicago.edu

A picture containing text

Description automatically generated

There you can log in with your Midway account (just as you do to log in to my.uchicago). After login, you will see a window similar to the one on the right opens on your desktop:

In the bottom left corner, you can open the terminal (either by clicking the icon with a terminal screen or going to Applications > Terminal), which has the exact same directory hierarchy you use when you *ssh* to midway.

In the terminal, you can access the VMD GUI by using the following commands:

module load vmd #This loads the vmd module to your session

vmd #This starts vmd

On the flip side, if you ever want to use the Tk console directly on Midway (without the GUI), say, when you ssh to Midway, you can use the following commands:

module load vmd #Same as before

vmd -dispdev none #Dispdev stands for display device; you are telling VMD to open # without a display device

This way, you can load your protein without any graphics representation and run analysis scripts directly on Midway. This would greatly expedite the analysis!