HudsonAlpha Institute for Biotechnology

CODE
Characterizing Our DNA Exceptions

Analysis of Molecular Dynamics Simulation files

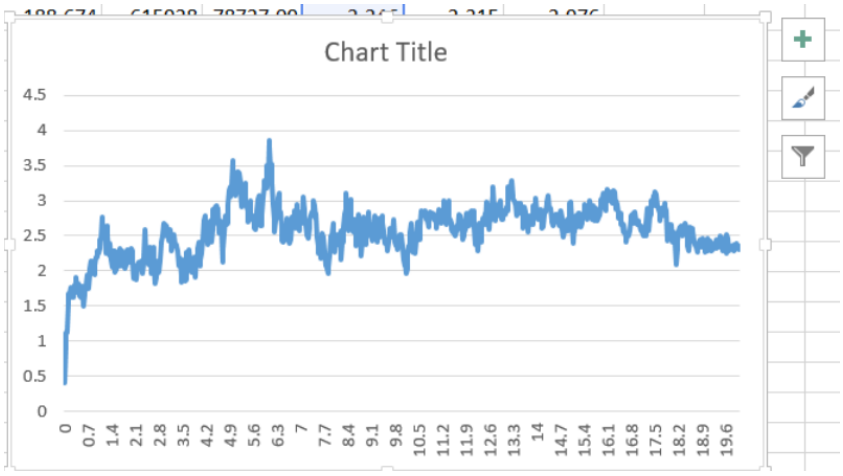
Upon completion of MDS you will find the analysis files generated under the 800 sim files. There will be two files of data titled analysis.tab & analysisRes.tab. Analysis.tab uses root-mean square deviation to track the global movement of the carbon-alpha backbone. AnalysisRes.tab uses root-mean square fluctuation to track individual residue movement of particular amino acids. These two files allow us to visualize the trajectory data by using simple line graphs in Excel.

1. Locate your folders with your output simulation (sim) files.
    a. If running multiple simulations on the same protein, it helps to stay organized by separating the folders into Wildtype and any mutants that you want to analyze.
        i. Ideally, these separate folders would be created BEFORE running the MDS.
    b. Check to see if you have all the sim files (there should be 800)
2. Fine the *analysis.tab* file -
    a. Located inside the list of files is a .tab file with a file name ending in _analysis.tab
3. Once you are confident that you have found the location of the _analysis.tab file open a blank spreadsheet in Microsoft Excel.
4. From within the Excel application, open the file ending with _analysis.tab for the protein that you wish to analyze.
    a. File > Open > Click file name
5. The Text Import Wizard should now be open.
6. Leave the “Fixed Width” option selected. Select “Next”
7. There is no need to change the column breaks. Select “Next”
8. There is also no need to change anything on Step 3. Select “Finish”
9. You should now have an Excel Worksheet with all of the raw data from the simulation.
    a. Name this document and save it. You will not edit this version of the document to prevent losing any of the original data.
    b. Now save the same document with a new name. This is the document that you will edit. Keep your original document to avoid losing any of your original data.
10. The first column is the time that the simulation ran. By default, this value is in picoseconds. These values should be converted to nanoseconds.
    a. Convert the picoseconds to nanoseconds by inserting a new column to the left of column. In cell A2 insert the equation =B2/1000. Auto populate the values for all of Column 1.

11. If the previous step was done properly you should now have a set of values that range from 0 to 20 nanoseconds in Column 1. This will be important when we graph the deviation over a course of time.

Graphing the _analysis.tab data

12. Hold the Ctrl-Key while selecting Column A (time in nanoseconds column) and Column J (Root Mean Squared Deviation [RMSD] Column).
13. VERY IMPORTANT: When you are selecting the columns, scroll down and be sure that you do not select the cells that say "#VALUE!" or the cells for "Mean,Min,Max" for the RMSD column. If you accidentally select these cells it will alter your graph.
14. Once you have these columns selected, look at the Excel Tool Bar and select the following tabs and drop-down items:
    a. Insert-> Charts-> Line
    b. Select the Line Graph that looks like the following:

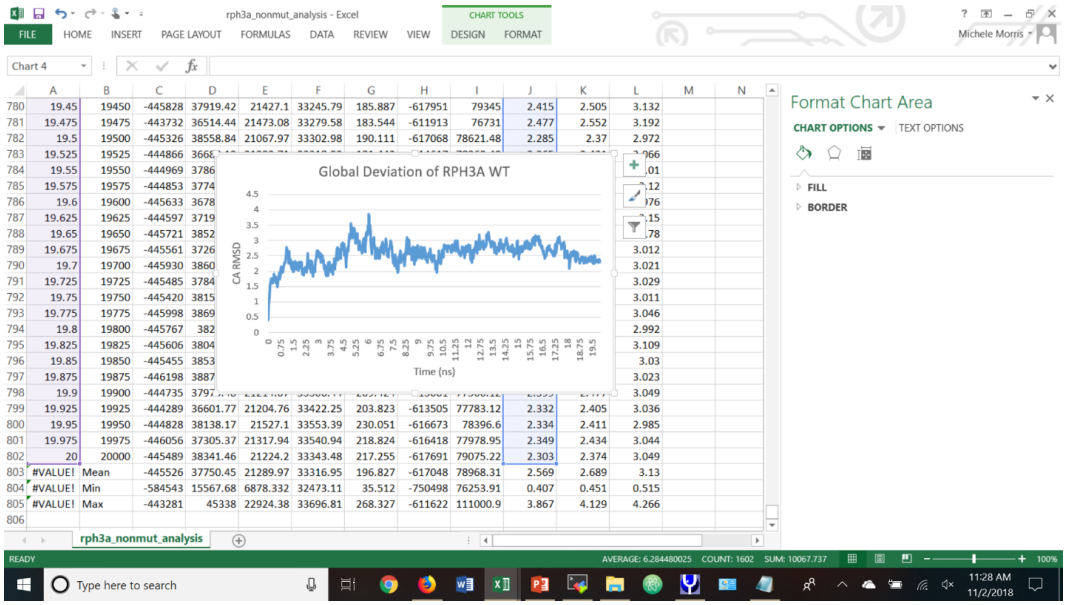


    c. Note: If you don't get the desired results, try Insert>Charts>Scatter>straight lines

Clean the graph to make it easier to decipher

15. In the Excel Tool Bar-Chart Design tab, fine "Add Chart Element". Select the green plus button in the top corner of the Add Chart Elements option.
    a. In the drop-down menus, select Axes
        i. Primary Horizontal and Primary Vertical should be checked.
    b. In the drop-down menu, select Axis Titles (Horizontal and Vertical). Type in your desired titles.
    c. In the drop-down menu, select Chart Title and type in your desired title.
    d. In the drop-down menu, select Gridlines; check to turn on gridlines for the chart as desired.

16. These next few steps are aesthetic changes that are not the only acceptable way to present your data. If you do not enjoy the look of these options feel free to play around with the graphs to make it how you want it.
    a. Change the Chart Title to: Global Deviation of “Insert Protein Name Here”
    b. Change the Axis Title on the Y-Axis to CA RMSD (Carbon-Alpha Root Mean Square Deviation)
    c. Change the Axis Title on the X-Axis to Time nanoseconds



17. Save your file and graph as an Excel workbook file.

Graphing the “_analysisres.tab” file data

1. Locate your folders with your output simulation (sim) files.
    a. If running multiple simulations on the same protein, it helps to stay organized by separating the folders into Wildtype and any mutants that you want to analyze.
        i. Ideally, these separate folders would be created BEFORE running the MDS.
    b. Check to see if you have all the sim files (there should be 800)
2. Fine the “*_analysisres.tab”* file -
    a. Located inside the list of files is a .tab file with a file name ending in “_analysisres.tab”
3. Once you are confident that you have found the location of the _analysisres.tab file open a blank spreadsheet in Microsoft Excel.
4. From within the Excel application, open the file ending with _analysisres.tab for the protein that you wish to analyze.
    a. File > Open > Click file name

5. The Text Import Wizard should now be open.
6. Leave the "Fixed Width" option selected. Select "Next"
7. There is no need to change the column breaks. Select "Next"
8. There is also no need to change anything on Step 3. Select "Finish"
9. You should now have an Excel Worksheet with all the raw data from the simulation.
    a. Name this document and save it. You will not edit this version of the document to prevent losing any of the original data.
    b. Now save the same document with a new name. This is the document that you will edit. Keep your original document to avoid losing any of your original data.

Create a graph of the _analysisres.tab data

10. Use the Residue and RMSDs columns.
11. Copy the columns and put them next to each other in a new Excel spreadsheet.
    a. The columns for Mean, Min and Max are not needed for your graph.
12. Name this document and save it
    b. Make this new document to avoid losing any of your original data.
13. Highlight the columns
14. Make the graph by selecting the following:
        i. Insert>Charts>Line
        ii. Make the X-axis the Residue (amino acid) names
15. Save your file and graph as an Excel Workbook file.

Compare the two Analysis graph results for the wild type and the mutant models

16. In a new Excel sheet, copy the Time column
17. Add the 1st RMSD (wild type) column of data
18. Copy and paste the 2nd RMSD (mutant) column of data
19. Relabel the headers to reflect wildtype or mutant
20. Highlight the columns
21. Create the graph by choosing:
    a. Insert>Charts>Scatter>straight lines
22. Save your file and graph as an Excel Workbook file
23. Go back to your *analysisres* data files, use the Residue and RMSD info to create a graph with both WT and Mut results.
    a. In a new Excel sheet, copy the Residue column
    b. Add the 1st RMSD (wild type) column of data
    c. Copy and paste the 2nd RMSD (mutant) column of data
    d. Relabel the headers to reflect wildtype or mutant
    e. Highlight the columns
    f. Create the graph by choosing:
        i. Insert>Charts>Scatter>straight lines
24. Save your file and graph as an Excel Workbook file
25. Try to replicate your results.
    a. Run the MDS again for both models (wild type and mutated)

        i. Set up new folders for the second MDS
b. Create the same graphs again
c. Compare to original graphs