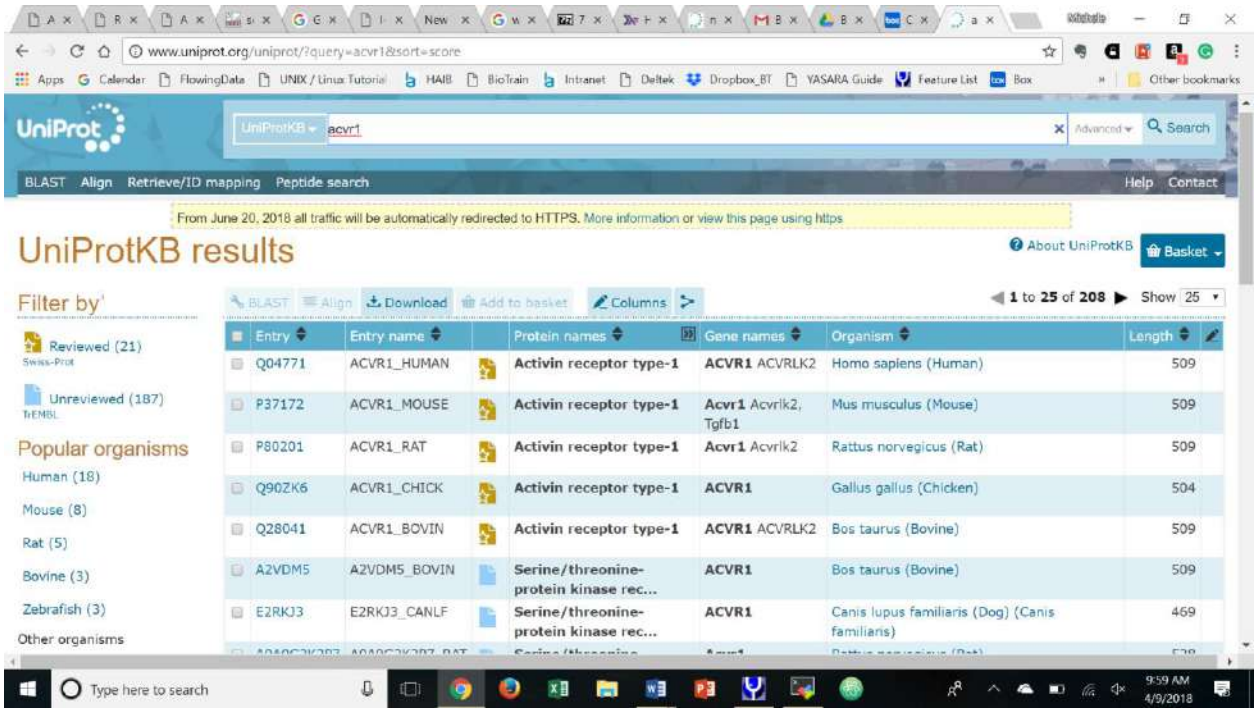
Homology Modeling with YASARA

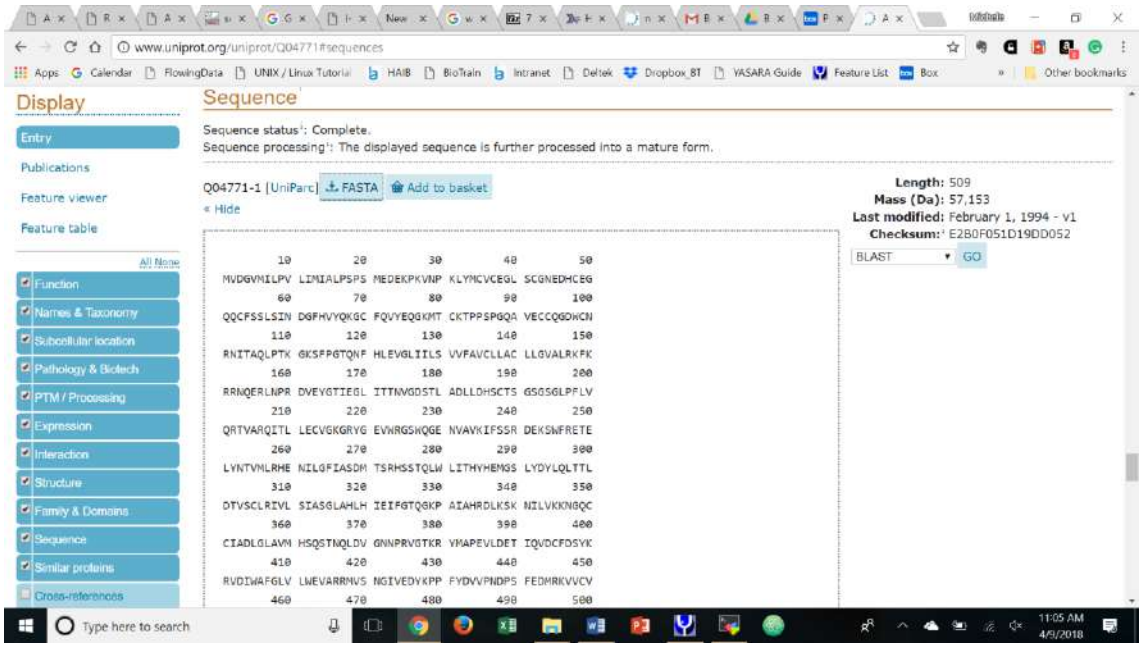
1) Search for desired protein on Uniprot.org



a) Select species of choice

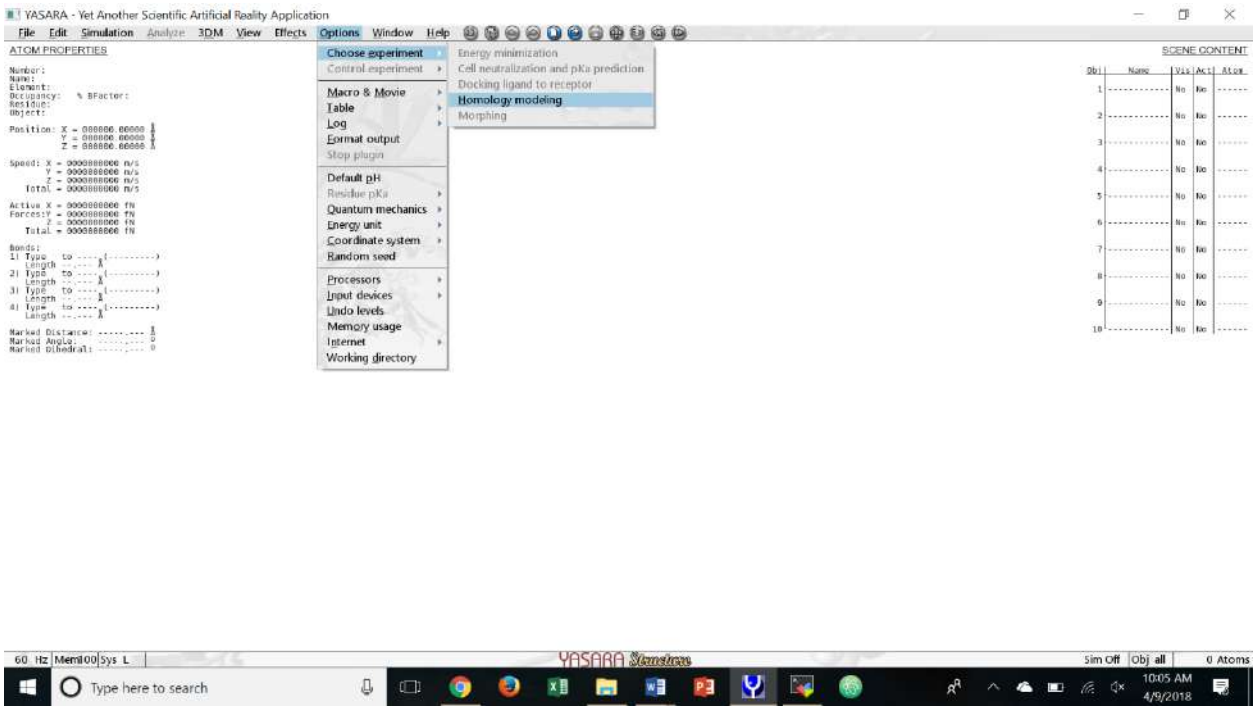
b) Scroll down to sequence

c) Right click on Fasta, save to a folder named after your protein

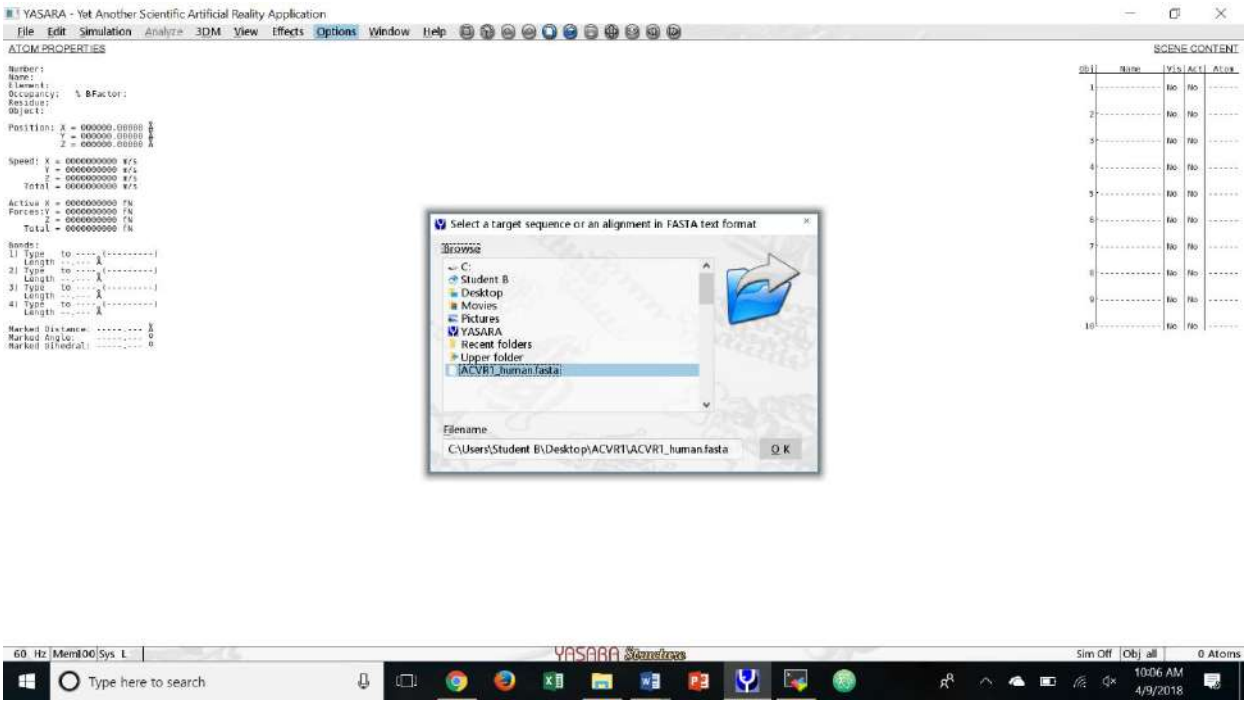


2) Open YASARA

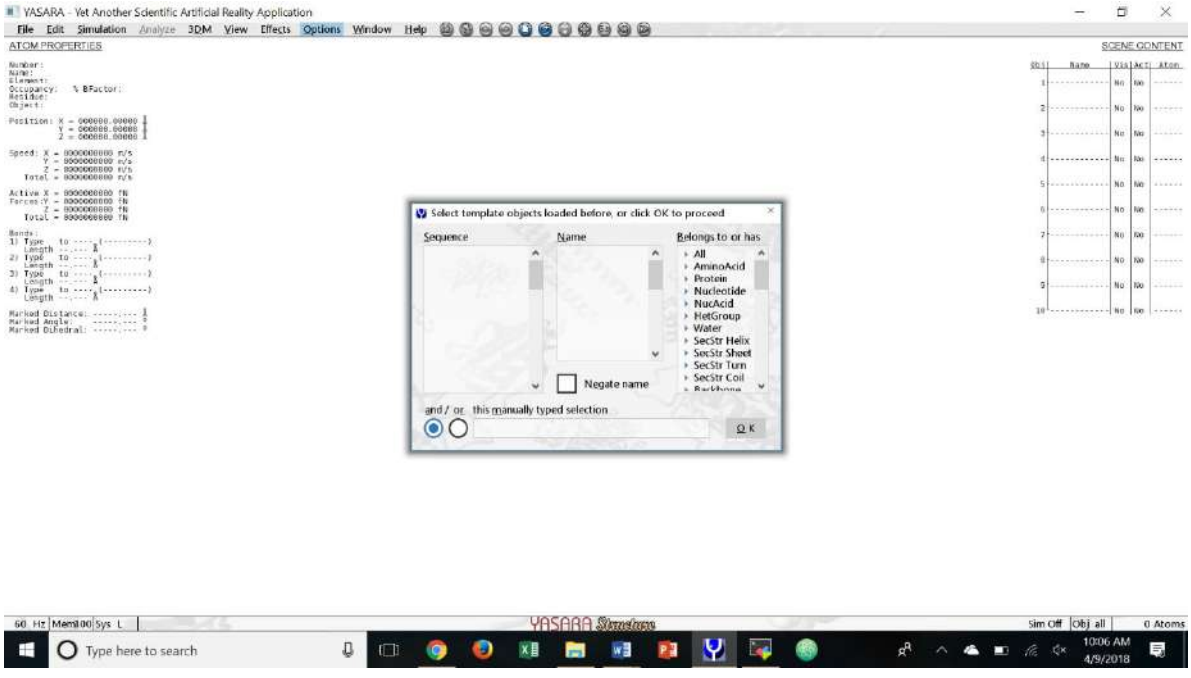
a) Click Options->Choose Experiment->Homology Modeling



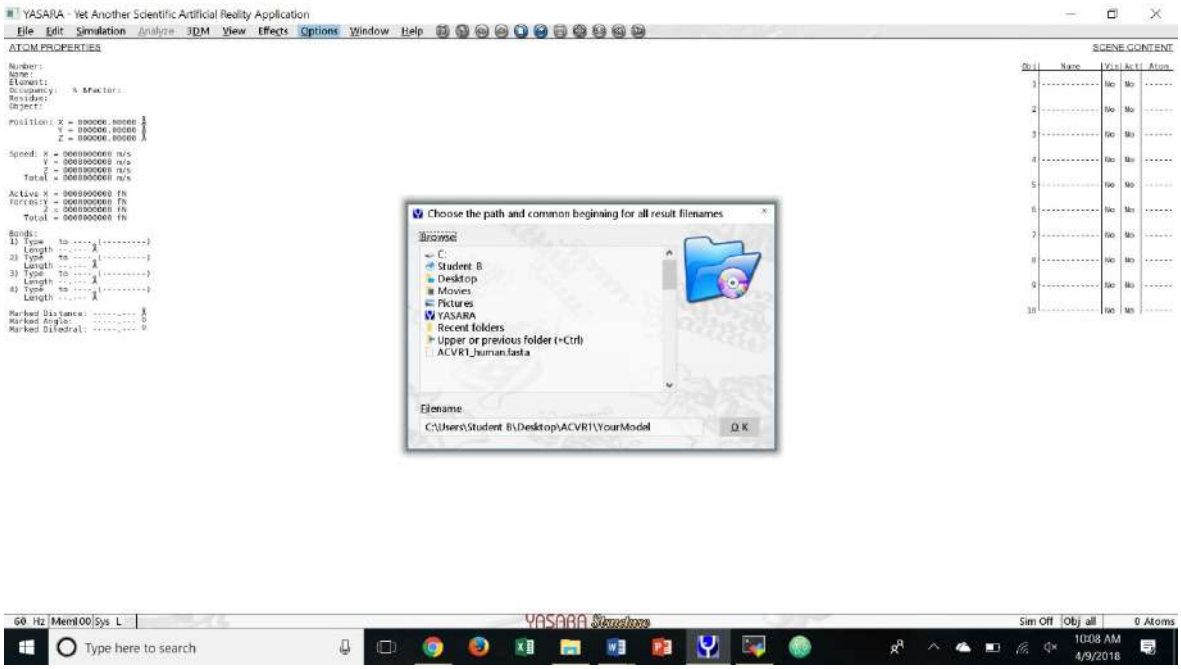
b) Find target sequence in File Browser, then click OK



c) If using a pre-modeled template select sequence in menu now, otherwise select OK



d) Designate the location that you want the output files to be saved to, this is typically the folder that you saved your.fasta sequence in.

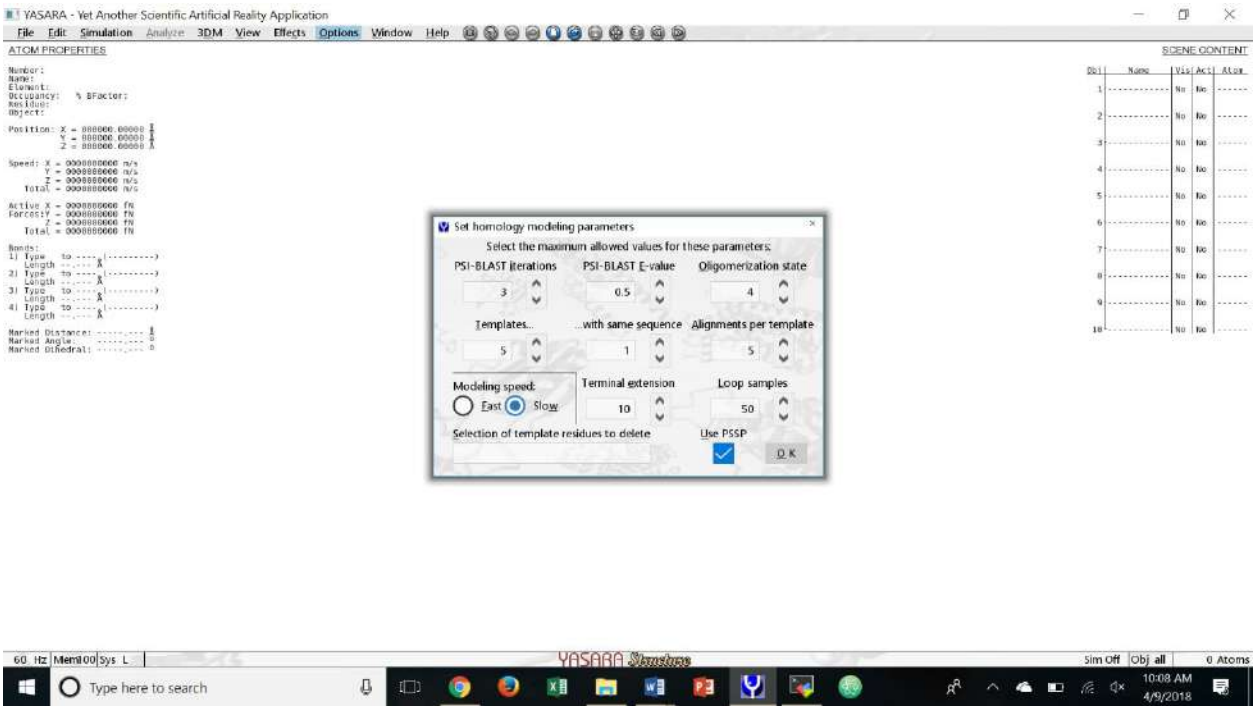


3) Modeling Parameters

   a) For Standard Running Parameters:

      i) Leave all options as is inside of the parameter box

ii) A typical Standard (slow) Running Model can take around a full day to run. Occasionally longer or shorter depending on protein size and complexity.



b) For Fast Running Parameters:

i) Change PSI-Blast iterations to 1

ii) Change Oligomerization State to 1

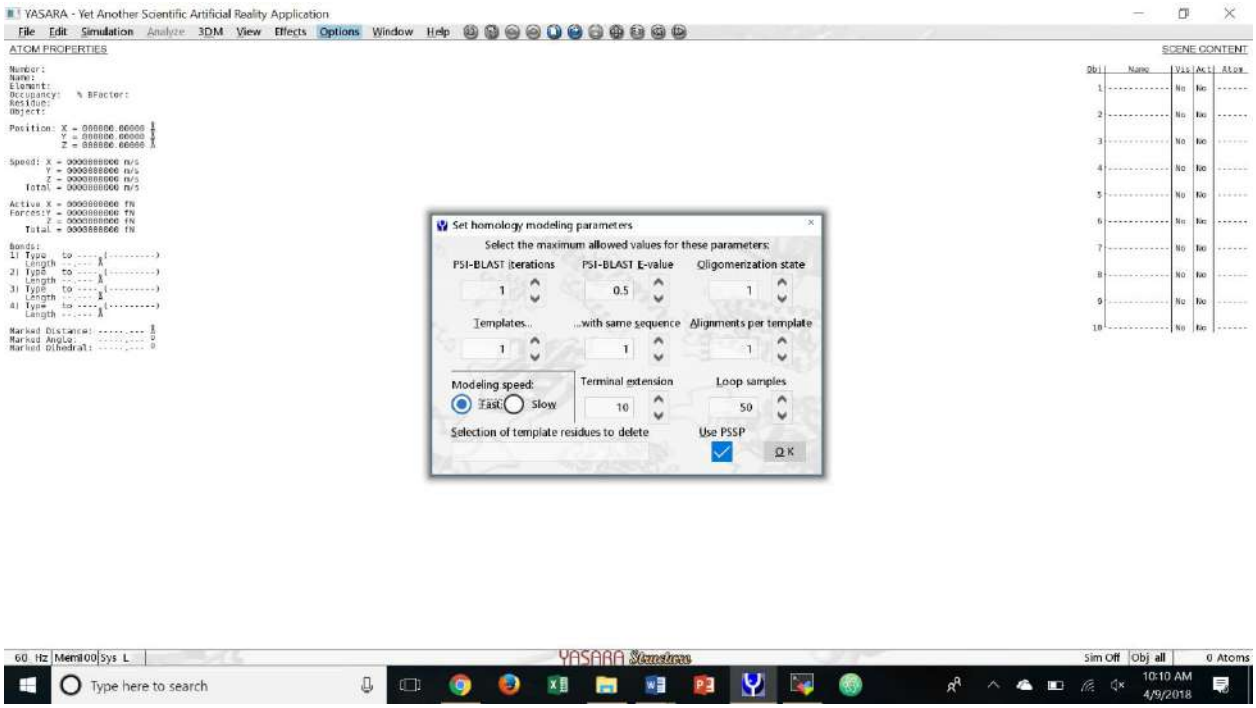
iii) Change Templates to 1

iv) Change Alignments Per Template to 1

v) Change Modeling Speed to Fast

vi) Select OK and modeling will begin

vii) A typical Fast Running Model will take anywhere from 10 minutes to a few hours depending on protein size and complexity

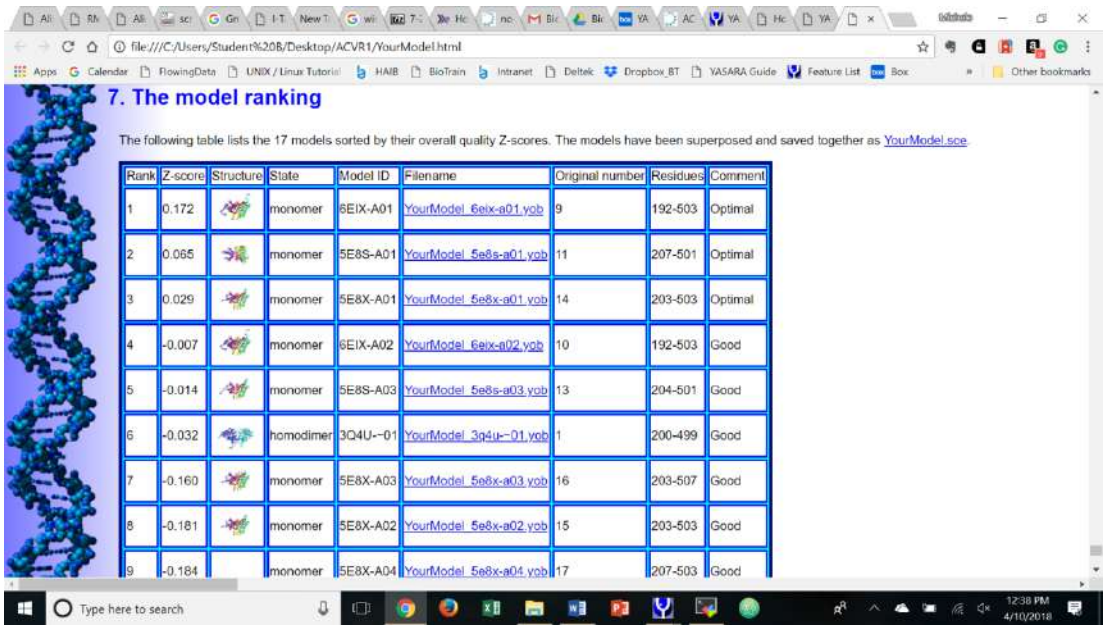


4) What to do when modeling is finished?

   a) If ran on <u>Standard Run Parameters</u>, a varying number of models will appear superposed upon each other. It may not let you toggle through each one.

      i) First, go to the Chrome HTML Document that will be placed in the chosen output folder

      ii) Upon opening this in your browser you will find an abundance of information regarding potential templates gathered, scoring parameters for models, as well as



visual representations of the top models. Scroll down and find the model with the highest overall Z-score, you can then rename to TopModel or your name of choice to indicate that it is the file that should be used or any future projects.

b) What is a Z-Score? A Z-score describes how many standard deviations the model quality is away from the average high-resolution X-ray structure. Higher values are better, negative values indicate that the homology model looks worse than a high-resolution X-ray structure. The overall Z-scores for all models have been calculated as the weighted averages of the individual Z-scores using the formula Overall = 0.145*Dihedrals + 0.390*Packing1D + 0.465*Packing3D. The overall score thus captures the correctness of backbone-(Ramachandran plot) and side-chain dihedrals, as well as packing interactions. It applies to globular proteins only and can be misled by artificial structures like long single alpha helices (which have perfect dihedrals and are free of packing errors, since there is no packing).

5) If ran on **Fast Run Parameters**, you will see the exact number of models as you set for templates.
   a) Typically, templates are set to 1 so you would only see the top model
   b) This saves you from having to analyze the Z-scores of multiple models.
      i) However, it is a good idea to document the Z-score of your model to assure accuracy.
   c) You will still need to save this model to a name of your choice to avoid confusion later on.