Molecular Modeling Homework 3 (2020)

Due Friday Apr 10

Open SARS-2 spike protein cryoEM model **6vsb**. In SEQ window, note that some parts of the sequence are bold (known structure) and some parts are not bold (structure unknown). In this homework you will model the unknown parts using MOE's Homology Model tool. Note also, that one of the domains is sticking up in chain A only, and down in chains B and C. We will make an "all up" model.

PART 1 -- modeling missing loops using homology model

1) Save the sequence. Select the three protein chains (delete glycan chains 4,5,6). **Save | Sequence | Only Selected | chains**
2) Open the file you just saved.
3) **SEQ: Alignment | Align | Sequence only**. (The sequences are already aligned, so nothing happens)
4) Run **Protein | Homology Model.** Model each of the three monomers using its corresponding sequence. (The sequences are identical in this case.) *Allow up to one hour for MOE to complete this task!* Cancel | Homology Model if it runs too long.
5) Open the mdb file created by MOE. Read in the first model. Close the template.

Now, find structurally conserved regions and variable loops. Variable loops indicate regions of low confidence.

6) SEQ: Color residues by chain.
7) MOE: Display ribbons only. Hide atoms. Color ribbons by residue.
8) Superpose the three chains using the whole alignment.
9) Select each domain in turn. For each domain, superpose just that domain by setting **Selected residues** in **Superpose**.
10) Identify the sequence limits of variable regions **by eye**, loops that are modeled very differently in the three monomers (low confidence regions) are the variable regions. For the variable regions, color residues white in the SEQ window. (Loops turn white in ribbon because of (7)).
11) Proceed to label low confidence loop regions white throughout the structure.
12) Render as ribbons only. Hide all atoms.

Send the MOE file for PART 1 as HW3.1

---------- PART 2 -- "all up" model. --------

1) Delete the model from PART 1. Re-import the model from the mdb file.
2) One of the monomers (chain A) has its receptor binding domain (RBD) sticking up. We're going to make the RBD stick up in all three monomers.
3) **Copy** chain A, then **paste** it. Label the chain of the new copy "**Chain B up**"
4) Fix chain B. Unfix "Chain B up" . Select the C-terminal 600 residues in all chains.
5) In the Superpose panel, set chains A and C to "i" (ignore). Leave chains B and "Chain B up" as "*". Use selected residues. Superpose. Check the quality of the superposition in the C-terminal 600 residues. If it is not good,

Good quality superposition

do it again. It should superpose well (see figure).
6) Delete original chain B.
7) **Copy** chain A, then **paste** it. Label the chain of the new copy "**Chain C up**"
8) Repeat steps 4,5,6 substituting chain C for chain B.
9) Now you should have a model with all three RBD's sticking up. This is the conformation of the spike protein when it is bound to the receptor protein ACE2 on the surface of the pulmonary cell, just before it invades the cell by fusing the virus membrane with the cell membrane.

Set the rendering so that ribbons are colored by chain. **Send the MOE file for PART 2 as HW3.2**

http://www.bioinfo.rpi.edu/bystrc/courses/biol4550/homework.html