**COVID-19: On the way to a vaccine**

**Part V - Molecular visualization of SARS-CoV-2 S protein attaching to human ACE-2**

Learning Objective: This worksheet engages students in exploring the specific interactions involved in COVID-19 attachment to human cells for infection.

*Viral Attachment*

Attachment of the SARS-CoV-2 virus to human epithelial cells requires interaction between the viral spike (S) glycoprotein and the human angiotensin converting enzyme 2 (ACE-2). The transmembrane ACE-2 protein is expressed in different cells throughout the body, but the

COVID-19 infection is largely targeted to the lower part of the lungs and the epithelial cells. Understanding the specificity of interaction between the viral S and human ACE-2 proteins may help develop vaccines and drugs to block the infection.

To obtain an atomic level picture of the critical interaction between ACE-2 and the SARS-CoV-2 S protein, Yan *et al.* (2020) determined the structure of the receptor binding domain (RBD) of the SARS-CoV-2 S glycoprotein complexed with the ACE-2 protein stabilized by another protein, B0AT1 (a Na+/amino acid transporter) using cryo-electron microscopy.

In the first part of this worksheet, you will use a web based interactive viewer for

three-dimensional macromolecular structures (iCn3D), to visualize the 3D structure of the ternary complex of SARS-CoV-2 S protein RBD-ACE2-B0AT1 determined by Yan *et al.* (2020). In the second part of this worksheet, you will look at figures from the original paper, Yan *et al.* (2020) to explore critical interactions responsible for the attachment of the virus to the cell, at the interface between the SARS-CoV-2 RBD and the ACE-2 peptidase domains.

# Part 1 - Visualizing 3D images of RBD-ACE2/B0AT1

Goal: In this part you will be guided through visualizing the 3D structure of the ternary complex of S protein RBD-ACE2-B0AT1 as determined by Yan *et al.* (2020).

1. Yan *et al* (2020) solved the structures of full-length human ACE2, in the presence of a neutral amino acid transporter B 0 AT1, with or without the receptor binding domain (RBD) of the surface spike glycoprotein (S protein) of SARS-CoV-2. These three structures were deposited to the Protein Data Bank (PDB) and can be accessed with their unique identifier (PDB ID #). The PDB ID for the ternary complex you will be viewing is 6M17. Browse through the components making of the complex, open the structure summary page of the PDB entry 6m17 (<https://www.rcsb.org/structure/6m17>). Read the abstract of the article describing this structure, to get a brief overview of the research.
2. To visualize the structure, go to <https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>. This website will allow you to visualize and highlight different structures within the protein structure file.

The left hand portion of the window that opens should look like Figure 1.



Figure 1. iCn3D start-up window. (<https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>)

1. In the PDB ID input box (left of LOAD) in the iCn3D start-up window, replace the existing ID# with 6M17. Click LOAD and a new window with the protein structure will open.

*Alternatively, you can pull down File > Open File > PDB file and enter the PDB file number (6M17) in that box.*

1. Click on the structure and explore rotating it freely to view the different perspectives of this protein complex in 3 dimensions. Also try zooming in and out using your mouse or touchpad to see additional structural details.
2. After exploring the 3 dimensional structure, go to the pulldown menu under “Windows > View Sequences and Annotations”. *This will open a window with information about the proteins and chemicals present in the structure.*
3. Scroll down in the “View Sequences and Annotations” window to see the different proteins that make up the ternary protein complex you are viewing. These proteins are labeled 6M17\_A through 6M17\_F. Since this protein complex is a dimer, there are two copies of each protein in the displayed structure.

Fill out Table 1 with the identities of the proteins corresponding to each label. Table 1 Protein components of the PDB file structure.

|  |  |
| --- | --- |
| Subunit | Protein |
| 6M17A & C |  |
| 6M17B & D |  |
| 6M17E & F |  |

1. Scroll down to the section that lists any chemical, ion, or water molecules that are included in the structure file. What chemicals are included in the structure you are viewing? *By*

*clicking on the highlighted chemical, a new window will open with the identity of that chemical. You can close the window after you get the information.*

Answer:

1. Scroll back to the top of the “View Sequences and Annotations”. Click on the 6M17\_A sequence name (*using a mouse, draw a box around the name - it should turn yellow).* This selects that protein in the structure. You will now see the corresponding protein highlighted in the protein graphics window.
2. From the” Color” pulldown menu (near the top), select Unicolor > Red. You should now have a highlighted protein structure that is red. Repeat for the 6M17\_C protein. *You can choose any of the available colors as long as each pair of identical proteins are color matched.*
	1. Repeat for 6M17\_B and D and color them blue.
	2. Repeat for 6M17\_E and F and color them green.
3. Rotate the colored structure so that the 6M17\_A and 6M17\_C proteins are at the bottom and you can clearly visualize the two helical domains of ACE-2 that are near the B0AT1 proteins..
	1. This image of your structure by going to File > Save File > iCn3D PNG image. *This image can be uploaded later into a powerpoint file or other document.*
	2. Besides the three different proteins you colored in different colors, there are a number of other structures seen in the graphics window (shown in stick representation). Can you identify some of them, by moving your mouse over those structures? List the names of any 2 of these molecules.
	3. The B0AT1 protein (sodium-dependent neutral amino acid transporter) was used to stabilize ACE-2 and allow for the structure determination and visualization of its transmembrane domain. Hide those two B0AT1 proteins to focus on the ACE-2 and S glycoprotein RBD interaction.
	4. Click on the B0AT1 proteins. *If you hold down the shift key, you should be able to highlight both of the subunits. If not, repeat this step for each chain of the B0AT protein.*
	5. Go to Style > Protein > Hide. Those subunits will no longer be visible in the image.
4. To identify the N- and C-termini of the ACE-2 protein, click on 6M17\_B (one ACE-2 protein) in the “Sequences and Annotations” window. With the protein highlighted in yellow, go to Style > Protein > Spectrum. That ACE-2 protein should now be a rainbow of colors (Spectrum) . The N-terminus is colored purple and the C-terminus is colored red.
5. Repeat the process with 6M17\_E to find the N-terminus of the S protein.

What do the three different rendering styles (curly helices, thick arrows, and curved thick lines) you see displayed, tell you about those regions of the protein structure? This style of depiction is called “ribbon” representation.

1. To visualize the protein in a space filling model, select both the 6M17\_D and 6M17\_F. Go to Style > Protein > Sphere. *Now half of your image is in a ribbon representation and the other half in a space filling (Sphere) representation.*

What kind of structural information do you think is better communicated using a ribbon representation and when would it be more advantageous to use a space filling representation?

# Part 2. Analysis of structures presented by Yan et al., (2020) showing the molecular details of the interaction between ACE-2 and the S glycoprotein RBD.

Note: You can generate figures similar to the ones included in the Yan *et al.* (2020) paper by using the iCn3D website. Since doing so requires additional time and practice, in this part of the worksheet you will be using images from the manuscript to explore and discuss SARS-CoV-2 RBD

-ACE-2 interaction.

The protein structure paper (Yan *et al*, 2020) demonstrates the manner in which the S glycoprotein receptor binding domain (RBD) interacts with the human ACE-2 protein. The authors presented the structure of the ternary complex of the S protein RBD plus ACE-2 and B0AT1 (Figure 2 - the structure that you viewed in the first part). Note that the alpha-helical structural elements you viewed in the ribbon representation of the protein complex in Part I are shown as tubular rods in Figure 2.



Figure 2. Overall structure of the RBD-ACE2-B0AT1 complex This is Figure 3B from Yan *et al.*

(2020). (<https://science.sciencemag.org/content/367/6485/1444/tab-figures-data>).

The S protein RBD and the ACE-2 protein interact in a specific manner as shown in Figure 3. The molecular details of this interaction are shown at the amino acid level in parts B, C and D of the figure.



Figure 3. Molecular details of the interaction between the S protein of SARS-CoV-2 RBD (colored in orange) and the peptidase domain (PD) of human ACE-2 (colored in light blue). This is Figure 4 from Yan *et al*., 2020 [(https://science.sciencemag.org/content/367/6485/1444/tab-figures-data](https://science.sciencemag.org/content/367/6485/1444/tab-figures-data)).

The authors, Yan *et al.* (2020), overlaid the structures of the SARS S protein (S protein from the SARS coronavirus)-ACE-2 complex with the COVID-19 (SARS-CoV-2) S protein ACE-2 complex as seen in Figure 4. Only part of this figure is shown.



Figure 4. Overlay of SARS-CoV-RBD-ACE-2 structure with SARS-CoV-2-RBD- ACE-2 structure. This is Figure 5 from Yan *et al*. (2020). [(https://science.sciencemag.org/content/367/6485/1444/tab-figures-data](https://science.sciencemag.org/content/367/6485/1444/tab-figures-data))

Questions

1. Why did the authors need to use B0AT1 protein to determine their protein structure even though the B0AT1 protein is not involved in the interaction between

SARS-CoV-2 RBD and ACE-2 that they were trying to learn about?

1. Describe the non-covalent interactions between the S protein and the ACE-2 protein (Figure 5 of Yan et al., 2020) that ensure the attachment of the virus to the cell.
2. Based on Figure 4, what are the main differences in the observed interactions between the ACE-2 PD and the SARS S glycoprotein RBD versus the SARS-CoV-2 S glycoprotein RBD. Discuss the possible functional implications of the structural differences you listed.