**COVID-19: Molecular Basis of Infection**

Didem Vardar-Ulu1 and Shuchismita Dutta2\*

1Biological Chemistry, Boston University, Boston MA

2Institute of Quantitative Biomedicine, Rutgers University, Piscataway NJ 08854

\*contact: sdutta@rcsb.rutgers.edu)

**Part 2: Beginning the Infection**

The first step in the viral life cycle is infection and it begins with the SARS-CoV-2 Spike protein binding a host receptor protein (Angiotensin Converting Enzyme 2 or ACE2 protein on lung cells). In the following 3 sections of this part of the case we will explore these proteins separately and then look at a structure of them interacting with each other to understand the molecular basis of this infection.

* Watch the ACS Reactions video <https://www.youtube.com/watch?v=gDY8pH6OWBc> for an introduction to the two proteins that play a key role in the SARS-CoV-2 infection – the viral Spike protein and the human ACE2 (receptor) protein.

*A. The Viral Spike Protein*

The SARS-CoV-2 Spike glycoprotein is over 1200 amino acids long. Explore the structure of the protein as determined by electron microscopy and discussed in the video you watched (PDB ID 6vsb). The focus in this exploration will be to:

1. Learn about the overall assembly and domain organization of the spike protein (domains).
2. Identify the domain within the spike protein that binds to the ACE2 protein.
* Open the structure explorer page for the entry by entering the PDB ID in the top search box on [www.rcsb.org](http://www.rcsb.org).

Q1. Review contents of the page and complete the following table about the entry.

|  |  |
| --- | --- |
| PDB ID |   |
| Author(s) of entry |   |
| Year when the structure was published/released |   |
| Structure determination method |   |
| Number of protein chains in the entry |   |
| Names and number of copies of ligands (Small Molecules) present in the structure |   |

* Click on the 3D view tab at the top of the structure summary page to view the structure interactively. Take a screenshot of the page and answer the following question:

Q2. How many protein chains do you see?

* To visualize and explore this structure of the SARS-CoV-2 Spike protein in detail:
	+ Go to the iCn3D website at <https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html>
	+ Click on the button called File >> Retrieve by ID >> PDB ID so that a new window opens. Input the PDB ID of the structure you wish to visualize and click on Load.
	+ The structure opens in a new tab – rotate the molecule and examine the overall structure.

Q3. Orient the structure so that the C-termini of the protein chains are at the bottom of the page. Take a screenshot of the structure and paste it below.

* Visualize a single chain of the Spike protein and explore it as follows:
	+ Click on the Select button >> Defined sets
	+ In the new window that opens Select the chains B and C with the shift button pressed (select from 6vsb\_B to 6vsb\_C and chemicals)
	+ Hide the selected chains/molecules by clicking on the button Style >> Proteins >> hide; Style >> Sidechain >> Hide; View >> Disulfide bonds >> Hide. Clear selections by clicking on the button Select >> Clear selection.
	+ Select chain 6vsb\_A and color the chain using the Color >> Spectrum option.
	+ You can view and select specific amino acid residues by clicking on the Windows >> Sequences & Annotations >> Details (in the new window that opens up). By clicking and dragging on specific amino acids in the sequence window you can select the amino acids in the graphics window.

UniProt of the Spike protein (<https://www.uniprot.org/uniprot/P0DTC2>) lists the Receptor Binding Domain (RBD) of the Spike protein (part of the protein that binds to the receptor protein ACE2) as the amino acids between 319 and 541.

Q4. Select the receptor binding domain in the spike protein seen in the chain A for PDB ID 6vsb and color it magenta (click on Color >> Unicolor >> Magenta). Orient this domain to be positioned at the top, save an image and paste it below. Label the following in the image:

* + N- and C-termini
	+ RBD
	+ where you think the viral envelope (or membrane) is located in this image.

*B. The Host Receptor: ACE2*

The ACE2 protein is a membrane bound Carboxypeptidase, a protease that cleaves amino acids from the C-terminus of proteins, in the presence of a zinc ion. Explore the structure of the catalytic domain of this protein as determined by X-ray crystallography (PDB ID 1r42). The focus in this exploration will be to:

1. Learn about the overall structure of the ACE2 protease domain.
2. Identify the domain that binds to the SARS-CoV-2 (and SARS) Spike proteins.
* Open the structure explorer page for the entry by entering the PDB ID in the top search box on [www.rcsb.org](http://www.rcsb.org).

Q5. Review the contents of the page and complete the following table with information about the entry.

|  |  |
| --- | --- |
| PDB ID |  |
| Author(s) of entry |  |
| Year when the structure was published/released |  |
| Structure determination method |  |
| Number of copies, Name of protein chains, Chain IDs in the entry |  |
| Names and number of copies of ligands (Small Molecules) present in the structure |  |

* Visualize the structure of the ACE2 protein structure in PDB ID 1r42 using iCn3D. Hide chains B-E which represent the disordered segment of collectrin homology domain. Select the ACE2 protein and color it, using the Spectrum option. Clear selections.

UniProt lists the active site residues for the ACE2 enzyme as E375 and H505 (<https://www.uniprot.org/uniprot/Q9BYF1>). It also lists 2 amino acids that if mutated can abolish the SARS Spike protein from binding (K31 and K353) in the Pathology and Biotech section. Visualize this information on the structure of the ACE2 enzyme to understand the location of the enzyme active site with respect to the Spike binding site.

* Open the sequences and annotation window (using the same steps as when you explored the Spike protein), locate and select the enzyme active site residues and the residues to which the SARS spike protein binds by clicking and dragging on these residues in the sequence and annotation window. Display the side chains of the selected residues by clicking on the button Style >> Side chain >> Ball and Stick.

Q6. Save an image of this structure and paste a copy below. Label the enzyme’s active site on the image. Assuming that the SARS-CoV-2 Spike protein binds in the same location as the SARS-CoV Spike protein, draw a circle around that region and label that location in your figure as SARS-CoV-2 binding site.

*C. Viral Attachment to Host*

The first step in viral infection is attachment to the host cell receptor protein. In the case of SARS-CoV-2, the viral Spike protein binds the ACE2 extracellular domain. Examine a structure of the SARS-CoV-2: ACE2 complex (PDB ID 6m0j).

* Open the structure explorer page for the entry by entering the PDB ID in the top search box on [www.rcsb.org](http://www.rcsb.org).

Q7. Review the contents of the page and complete the following table about the entry.

|  |  |
| --- | --- |
| PDB ID |  |
| Author(s) of entry |   |
| Year when the structure was published/released |  |
| Structure determination method |  |
| Number of protein chains in the entry |  |
| Names and number of copies of ligands (Small Molecules) present in the structure |  |

* Visualize the structure of the SARS-CoV-2 Spike:ACE2 complex in PDB ID 6m0j using iCn3D.

The video that you watched mentioned that the binding between SARS-CoV-2 Spike protein and ACE2 is stronger. Of all the non-covalent interactions that facilitate protein-protein interactions salt bridges are one of the strongest.

* Examine the structure to see if there are any charge based interactions (salt bridges) between the Spike and ACE2 protein using the following steps in iCn3D:
	+ Click on the Select button >> Defined sets
	+ In the new window that opens Select the chains E (6moj\_E)
	+ Examine any salt bridges between this selected chain and the other chain in the structure by clicking on View >> H Bonds & Interactions >> check to turn off all the box next to different types of interactions and select the box next to Salt Bridges >> click on Display.
	+ Look for any Salt Bridges formed between the Spike and ACE2 proteins.

Q8. How many salt bridges do you see between the SARS-CoV-2 and ACE2 proteins? Make a figure using an image from your iCn3D visualization of the SARS-CoV-2:ACE2 co-structure where you circle the salt bridge(s) and Save an image and label the residues forming the salt bridge(s).

(Note: Mouse over any residue in the graphics window to see the residue number. Convert that NCBI reference number to the PDB/UniProt number by reading off the corresponding number from the Sequences and Annotations window).