

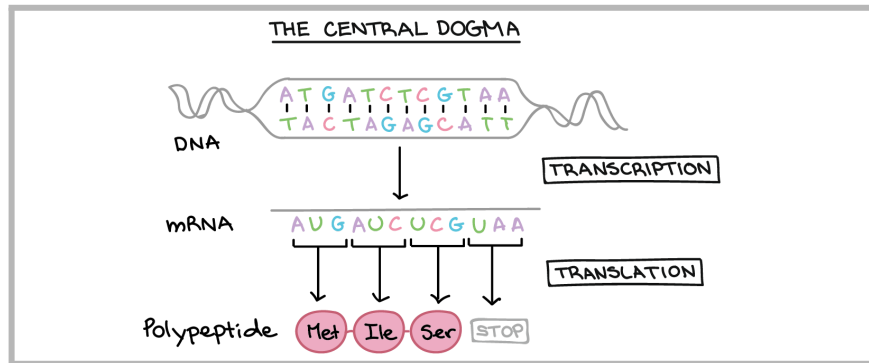
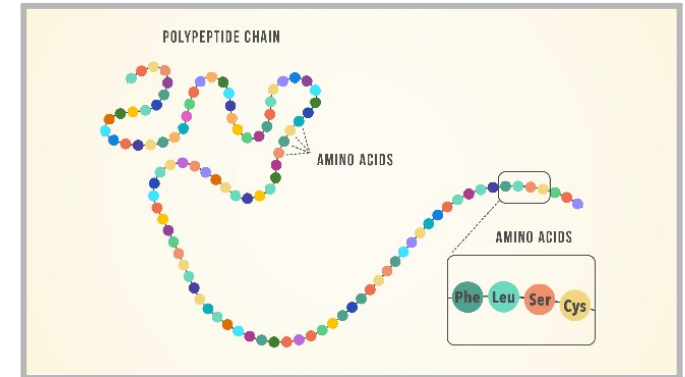
# AP Biology - DNA, Proteins & Protein Folding

## Goals

- Learn more about SARS-CoV-2 and its proteins
- Review the connection between DNA, RNA and proteins
- Learn how proteins fold and achieve their 3D design
- Consider the age-old questions of “Why are we studying this?” and “How does this affect the life today of an average person?”

## Introduction

Proteins are macromolecules made up of long chains of amino acids joined by peptide bonds that form complex folded structures. Proteins are responsible for numerous functions including structure, storage, hormones, enzymes, the immune system, transport of molecules (such as hemoglobin) and receptors & markers. A protein's 3-D structure is determined by the interactions of its side chains on the amino acids that make it up. This structure determines how a protein functions and interacts with other molecules.



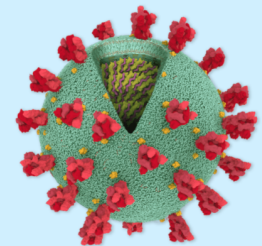
The primary structure of proteins is determined by the mRNA sequence, which is determined by the DNA sequence. This is the definition of the central dogma.

## Connection to SARS-CoV-2

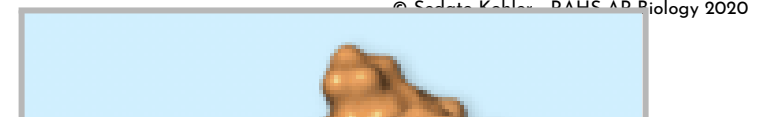
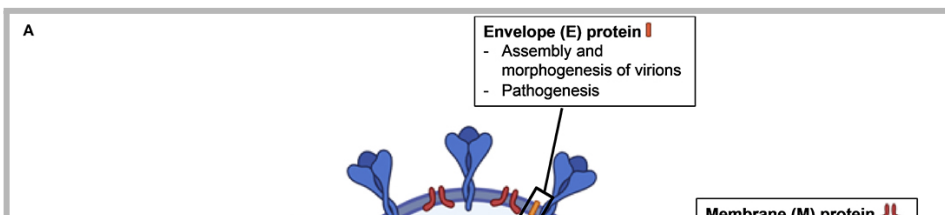
All things that have genetic material have proteins - that's the whole point of DNA/RNA. Note how it's all “things”, not all “living things”. Viruses have DNA/RNA. But... viruses are not living.

## Bad News Wrapped in Protein: Inside the Coronavirus Genome

By Jonathan Corum and Carl Zimmer April 3, 2020



The genome of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is less than 30,000 “letters” long. (The human genome is over 3 billion.) Scientists have identified genes for as many as 29 proteins, which carry out a range of jobs from making copies of the coronavirus to suppressing the body's immune responses. [Here's a great NYT article](#) about many of the proteins that SARS-CoV-2 uses.



© Sadava, et al. RAHS AP Biology 2020

## The ORF6 protein - the signal blocker

This accessory protein blocks signals that the infected cell would send out to the immune system. It also blocks some of the cell's own virus-fighting proteins, the same ones targeted by other viruses such as polio and influenza.

According [to a recent paper](#) in the American Society for Microbiology Journal of Virology, "SARS-COV ORF6 protein is localized to the endoplasmic reticulum (ER)/Golgi membrane in infected cells, where it binds to and disrupts nuclear import complex formation. Retention of import factors at the ER/Golgi membrane leads to a loss of STAT1 transport into the nucleus in response to interferon signaling, thus blocking the expression of STAT1-activated genes that establish an antiviral state."

## Part 1: DNA

The DNA below was sequenced from a sample of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) isolated from an individual from [Sauk County, WI on August 20, 2020](#). This is the sequence for the ORF6 protein.

ATGTTTCATCTCGTTGACTTTCAGGTTACTATAGCAGAGATATTACTAATTATTATGAGGACTTTTAAAGTTTCATTGGAATCTTGATTACATCATAAACCTCATAATTAATAATTTATCTAAGTCACTAACTGAGAATAAATATTCTCAATTAGATGAAGAGCAACCAATGGAGATTGATTAA

Even though that's really NOT that much DNA in the grand scheme of things, for the sake of our sanity, we're just going to look at the first 60 nucleotides.

Fill in the corresponding nucleotides to complete the template strand from the coding strand.

<b>Coding Strand</b>	ATG TTT CAT CTC GTT GAC TTT CAG GTT ACT ATA GCA GAG ATA TTA CTA ATT ATT ATG AGG
<b>Template Strand</b>	



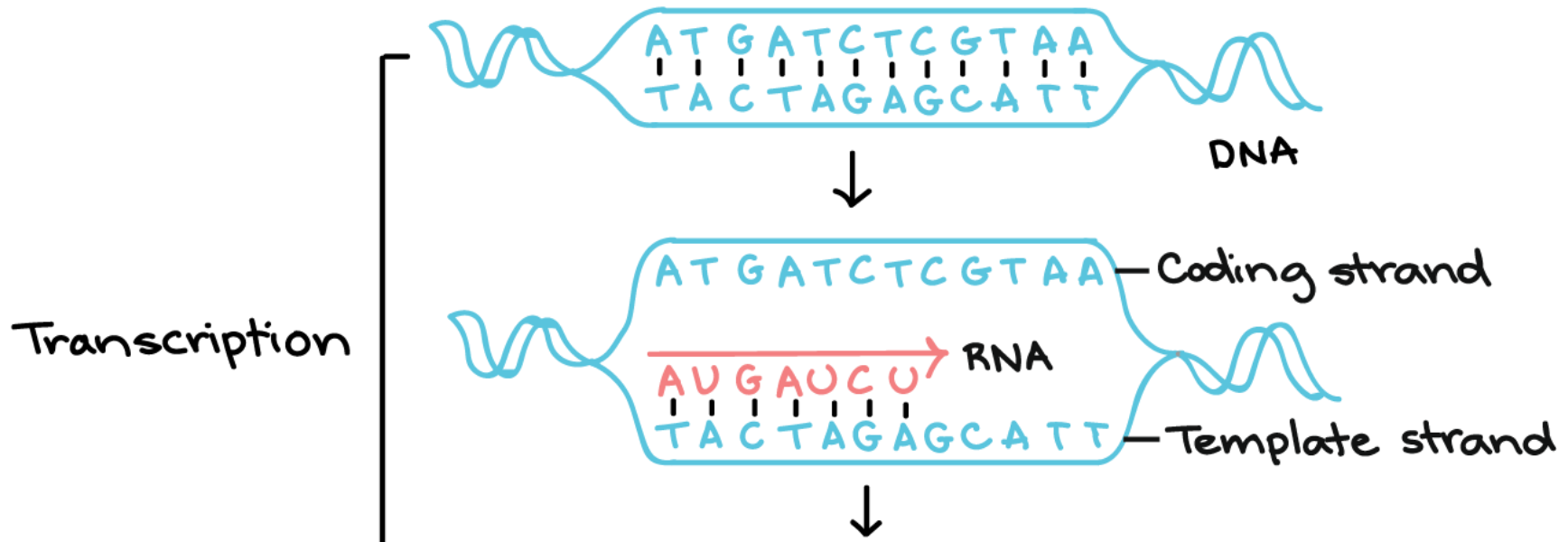
*A small quiet aside... SARS-CoV-2 is a positive-sense single-stranded RNA virus...meaning that viral particles enter into the cell with their genetic information already ONE STEP further along than the cell's genetic info... making it that much easier to use to produce the proteins that will help the virus replicate faster.*

*The DNA that I used for this activity is the DNA sequence that corresponds to the RNA sequence of DNA that was actually injected into the host cell.*

## Part 2: RNA

**Copy and paste** your template strand DNA into the box below. Next, transcription happens, where the letters of DNA are transcribed into mRNA. We will go into greater detail when we get to this unit.

Template Strand	
mRNA strand	



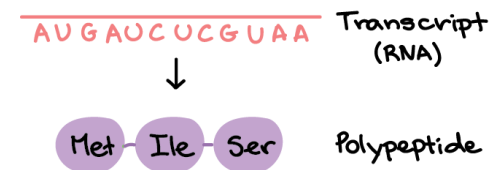
## Part 3: Amino Acids

Copy and paste your mRNA strand into the box below. Next, translation happens, where the mRNA is matched up with an Amino Acid. Use the codon table to find the corresponding Amino Acid for each codon.

<b>mRNA strand</b>	
<b>Amino Acid sequence</b> (just use the 3 letter abbreviation of the AA name)	

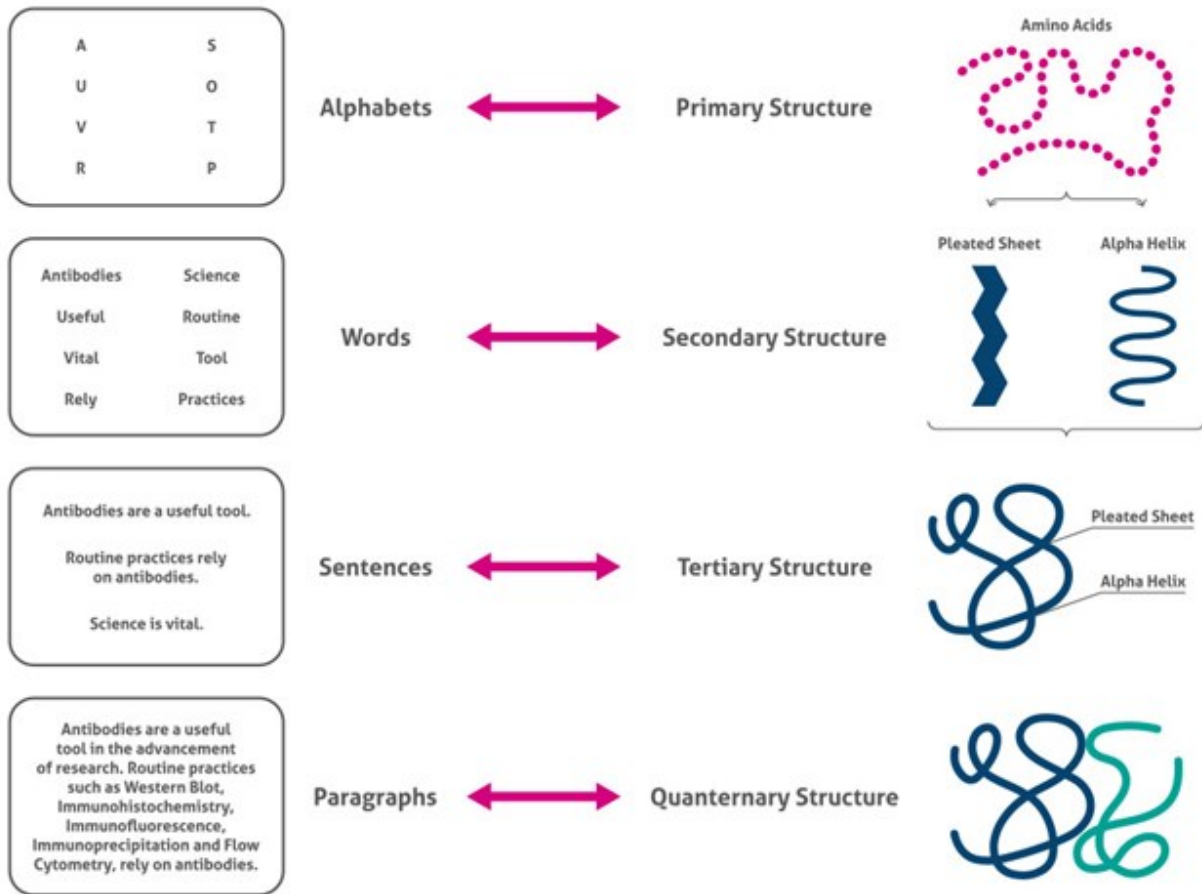
		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G
		Third letter				

Translation



## Part 4: Protein

Your amino acid sequence is just the start of protein structure. There are four levels of protein structure and folding. **Use the image below to describe** the 4 levels of protein structure in your own words.



1st:

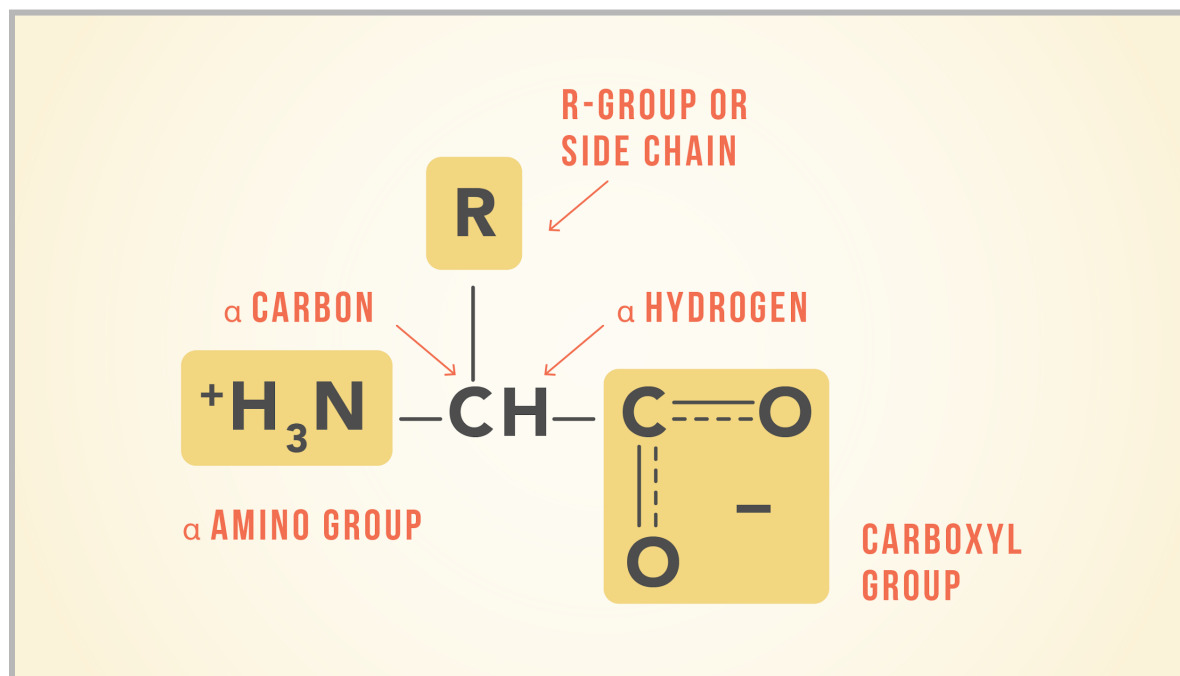
2nd:

3rd:

4th:

You are now going to make a 3D structure of this protein. First, **copy and paste** your amino acid sequence from above into the box. Next, using the categories in the image below, **highlight your amino acids** accordingly. You may choose the colors.

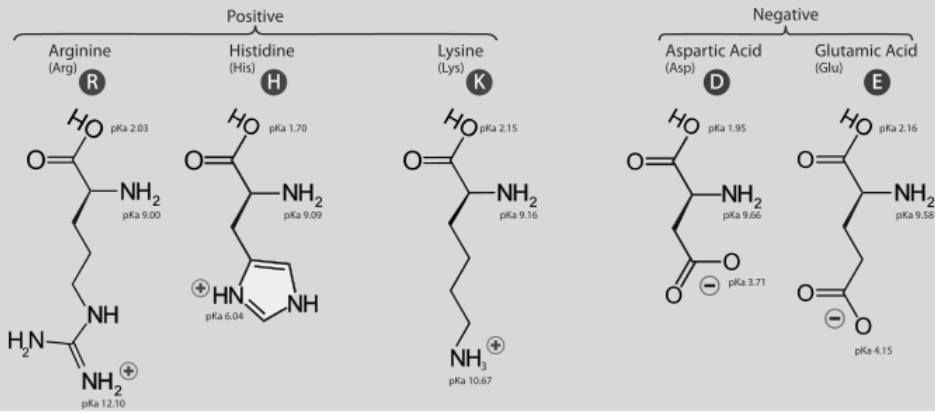
Color KEY	Non-polar/Hydrophobic	Polar uncharged/Hydrophilic	Positive charge/Hydrophilic
	Negative charge/Hydrophilic		Special cases
Amino Acid sequence			



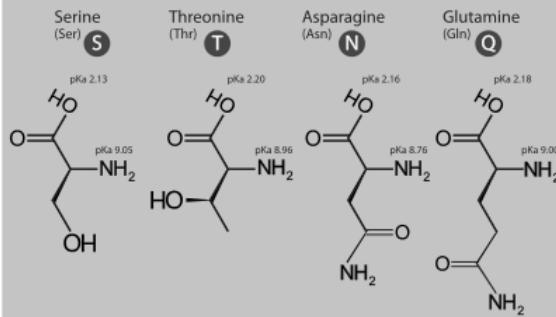
## Twenty-One Amino Acids

⊕ Positive ⊖ Negative  
• Side chain charge at physiological pH 7.4

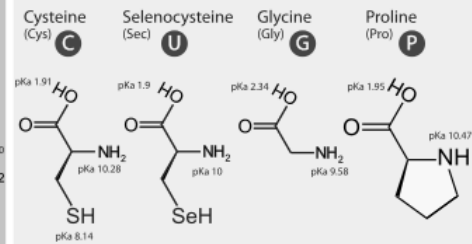
### A. Amino Acids with Electrically Charged Side Chains



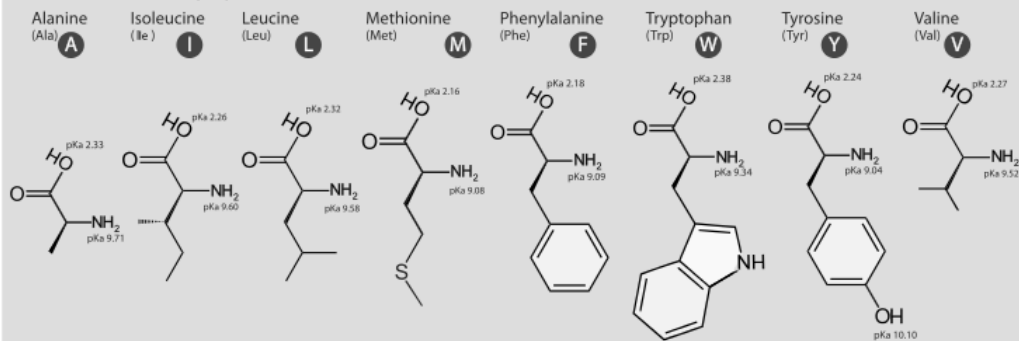
### B. Amino Acids with Polar Uncharged Side Chains



### C. Special Cases



### D. Amino Acids with Hydrophobic Side Chain



pKa Data: CRC Handbook of Chemistry, v2010

Den Copcari, Department of Medical Biophysics, University of Toronto, 2010



As you can see in the diagram of the amino acids, the R group of each determines the *personality* of that individual amino acid. I like to think of them as students in a class. Some students may have very similar personalities (amino acids of the similar side chains). Some may have attractive forces between them (positive and negatively charged side chains) and some may repel each other (positive + positive, negative + negative). Some students like to sit on the outside of the room (polar) and some may like to cluster in the middle of the room (non-polar).

Why do you think some amino acids (nonpolar) cluster on the inside of the protein and some are more often found on the outside (polar)? Think about the aqueous environment that proteins usually are in.

Now, it's time to fold your own ORF6 protein!

### Primary Structure

Take two pipe cleaners and fold their ends together to get one long pipe cleaner. Put your amino acids sequence on your pipe cleaner.

**Take a picture** of your protein at this point and **insert** it into the box below.

### Secondary Structure

Fold your pipe cleaner into both zig zag formations and spiral formations. These represent the formation of alpha helices or beta pleated sheets. These shapes form because of the hydrogen bonding of the carboxyl and amino groups of the amino acids.

**Take a picture** of your protein at this point and **insert** it into the box below.

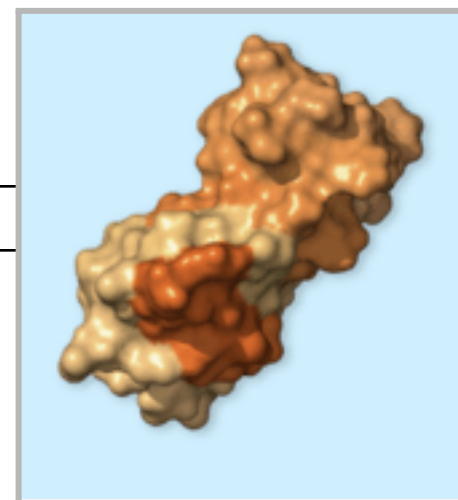
## Tertiary Structure

Amino acid/bead color (copy from your key above)	Chemical properties of that R group	How it affects protein folding
	Non-polar/Hydrophobic	Hydrophobic amino acids remain in the interior of the protein and attract to other hydrophobic amino acids.
	Polar uncharged/Hydrophilic	Polar amino acids are the outermost amino acids.
	Positive charge/Hydrophilic	Attracted to negatively charged amino acids. Are the outermost amino acids.
	Negative charge/Hydrophilic	Attracted to positively charged amino acids. Are the outermost amino acids.
	Special cases	Paper clip the beads together.

Fold your protein according to the following order:

1. Add your paper clips for your special case AAs
2. Fold your positive and negative AAs near each other
3. Fold your hydrophobic AA to the inside.
4. Fold your hydrophilic AA to the outside.

**Take a picture** of your protein at this point and **insert** it into the box below.



## Quaternary Structure

The ORF6 protein does not have quaternary structure. Your protein is complete at this point.

## Explore Current Research

❑ Listen to the TED Talk 5 challenges we could solve by designing new proteins | David Baker via Edpuzzle  
[3rd hr Link](#) [5th hr Link](#)

Feel free to put this video on and do other things (*not TikTok*) while you listen to it like a podcast - his visuals are not that interesting.

Why should we continue to study proteins, including their shapes and the DNA that codes for them? (5 sentences minimum)

Read [this article](#) about some current SARS-CoV-2 research.

How is protein research contributing to ending this pandemic? (5 sentences minimum)

## Conclusion

What is the connection between DNA, RNA, and proteins? (5 sentences minimum)

How do proteins get their shape? (5 sentences minimum)