# Homework 2 Write-Up

4. Making the `startEndPenalty` a `0` really helped the `score` that the algorithm gave. For the run without the `startEndPenalty` defined, I got a score of `1075` and with the `startEndPentaly` set to `0` I got a score of `1711`. By making this change we are allowing for `SARS-Cov-2_reference_genome_spike_protein` to have more tolerance in the end without imposing the large `-2` penalty. Since the Pfizer vaccine starts with Putative 5' UTR and ends with 3' UTR, the `startEndPenalty` of `0` ensures that we skip those portions since they don't match the `SARS-Cov-2_reference_genome_spike_protein`.

5. As we can see from the figure, the start of the vaccine contains Putative 5'UTR (cyan) and the ending contains 3' UTR (purple). Without using the `startEndPenalty`, it would force the spike protein to match with the cyan and the purple parts of untranslated region which doesn't correlate with the spike protein. Moreover, since the vaccine is bigger than the spike protein, there are a lot of bits which don't correspond to the spike protein which is what the `startEndPenalty` aims to alleviate.

6. There are a total of `1051` mismatches.

9. There are a total of `2` mismatches between the amino acid sequence. Towards the center of the protein, there is a Lysine and Valine, whereas the vaccine has two Prolines instead.

10. This makes sense because Proline ends up being the most rigid amino acid which allows for better stability during the perfusion state. If this were not the case, the human body would not adequately be able to make antibodies since fusion with human cells would be lack luster.

11. By having so many synonymous mutations, we are able to increase the GC content. This increase the stability of the vaccine and also increases the melting point of the vaccine compared to the spike protein.