Modelling COVID-19 Spike Protein Interactions

Building a better computer virus

Michael Kuiper and Tim Ho

In this study, we built computational models of the COVID-19 spike protein utilising CSIRO's Bracewell GPU cluster to help better understand how it infects our cells and evades our immune system.

Background

The coronavirus SARS-CoV-2 (COVID-19) attaches to human cells through its spike (S) protein receptor binding domain (RBD) to ACE2 proteins to begin the infection process. Because the S protein is the most prominent surface protein of the virus, it is used by many of the COVID vaccines currently in development to help prime our immune systems. As the virus mutates over time, it is important to evaluate changes we observe in human populations to ensure vaccines are still effective as well as check the new strain's virulence.

The Models

Homology models of COVID-19 S protein, based on the original SARS structures, were built and simulated using NAMD [1] to interact with various human proteins relevant to the infection process and immune response.

Our first model investigated the RBD bindng to ACE2 (Angeotensin converting enzyme 2) which is found on the surface of human cells such as certain epithelial lung cells. This interaction is crucial to understanding why COVID-19 can infect humans.

Our second model included most of the S protein, which is actually a trimer covered in glycans (types of sugar molecules). The glycosylation of the protein gives the virus some protection from our immune system by hiding the S protein underneath. By mapping mutations on the model we can evaluate if the changes are likely to be seen by our immune cells or effect binding to ACE2 (Figure 1).

A third model investigated the binding of antibody fragments recognizing the RBD of the virus. This helps us understand how vaccines might potentially fail as the virus evolves (Figure 2).

Bracewell

The Bracewell GPU cluster managed by CSIRO's Scientific Computing group provides access to over 400 Nvidia P100 GPUs and a BeeGFS file system connected by an InfiniBand network for fast data processing. The cluster is used by a range of HPC applications and workflows, such as computer vision processing, finite element analysis and molecular dynamics simulations. The use of Bracewell has allowed us to parallelise our NAMD simulations by taking advantage of multiple compute nodes and GPUs. In addition, we also used VMD [2] on the remote visualisation nodes attached to Bracewell for interacting with the models and visualising simulation results.

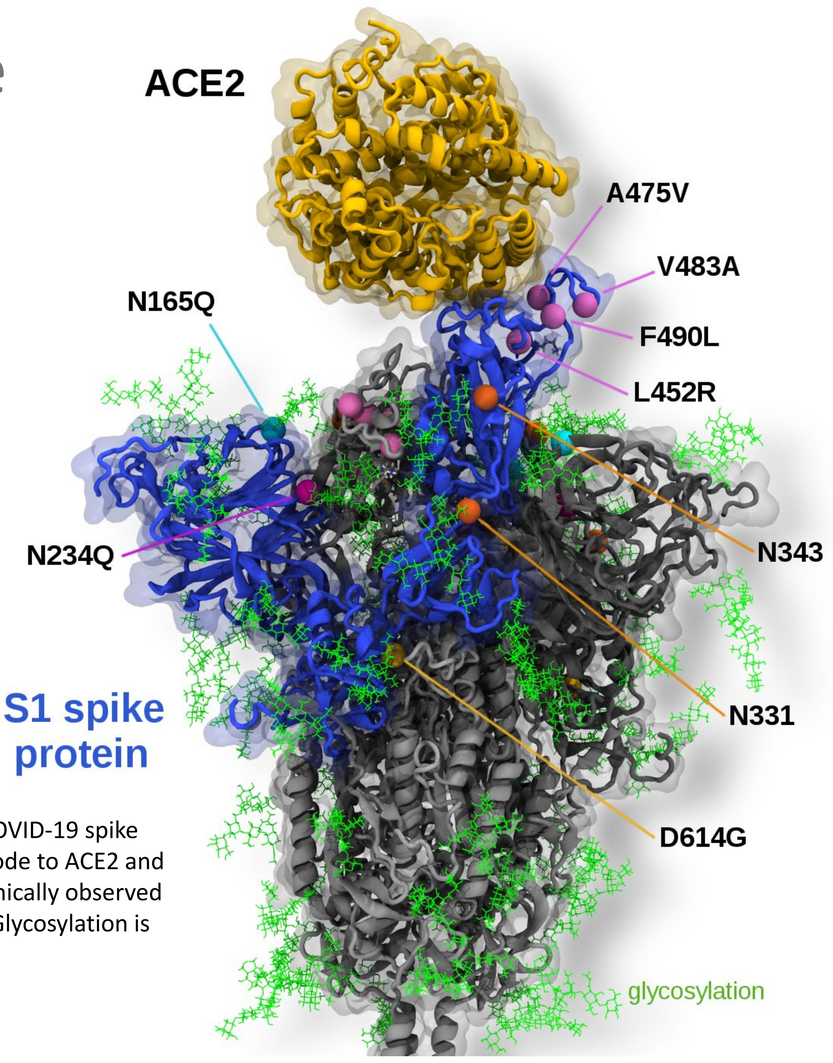

Figure 1: A Model of the COVID-19 spike protein showing binding mode to ACE2 and positions of some of the clinically observed mutations such as D614G. Glycosylation is depicted as green sticks.

Result

The Bracewell GPU cluster was key to quickly initiating and performing this work. Having rapid access allows for quick prototyping and troubleshooting with the simulations. Earlier on in the pandemic little structural information was available, but as the worldwide research effort gained pace, we were able to incorporate new findings into our developing models.

As a new dominant strain was being recorded (D614G), we were able to incorporate this into our models and with experimental validation from ACDP (Australian Centre for Disease Prevention) we were able to assess the effects of the mutation. Fortunately, this strain does not appear to resist vaccine responses, however work by others indicates it is a more efficient spreader [3].

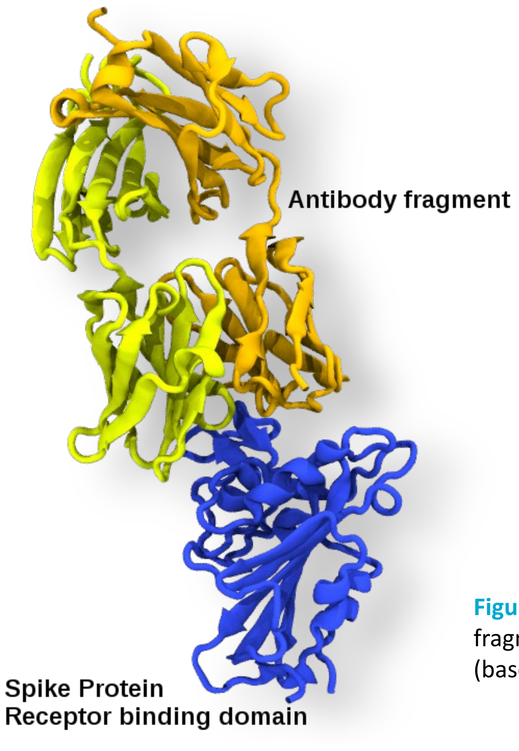

Figure 2: Modelling of a COVID-19 specific antibody fragment against the Receptor binding domain (based on pdb structure: 7bz5)

Figure 3: CSIRO's Bracewell GPU cluster

FOR FURTHER INFORMATION

Michael Kuiper
CSIRO Data61
Michael.Kuiper@csiro.au
https://data61.csiro.au/

REFERENCES

[1] James C. Phillips, *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. Journal of Chemical Physics, 153:044130, 2020. doi:10.1063/5.0014475

[2] Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", J. Molec. Graphics, 1996, vol. 14, pp. 33-38.

[3] McAuley A, *et al*. Experimental and in silico evidence suggests vaccines are unlikely to be affected by D614G mutation in SARS-CoV-2 spike protein. Npj Vaccines (2020) (in print) 10.1038/s41541-020-00246-8

ACKNOWLEDGEMENTS

This study was supported by resources and expertise provided by CSIRO IMT Scientific Computing.