Artificial Intelligence in Studying Catalytic Antibodies

John Wensink

ENG102-4 Composition II

Colorado State University-Global Campus

Dr. Rose Bunch

May 10, 2020

Artificial Intelligence in Studying Catalytic Antibodies

In late 2019 a novel coronavirus (SARS-CoV-2) appeared in Hubei Provence, China, causing widespread pneumonia and respiratory distress to individuals who become infected (Zheng, 2020). The virus quickly jumped national borders, and at the time of this publication has been responsible for 137,184 American deaths, with a peak of 2,215 deaths per day on April 17, 2020 (IHME, 2020). Researchers including Hifumi et al. (2020) are examining the role of monoclonal antibodies (mAbs), specifically the effects that come from the deletion of the genetic expression of PRO95, a light-chain antibody residue that has been observed to have properties which inhibit the catalytic function of mAbs. The research presented by Hifumi et al. is supported by parallel research conducted by other teams including Usui et al. (2017), Walls et al. (2019), as well as Buneva & Nevinsky (2017) who have been able to demonstrate that the absence of PRO95 is a necessary part of antigen-binding as well as multispecificity of monoclonal antibodies. These research efforts could be aided by the application of artificial intelligence algorithms to help address human researcher’s cognitive limitations in making sense of the vast collections of data that such research entails. The processing power required to perform big data analytics can be provided by everyday home users with their own PCs, Mac’s, and Laptops. The researchers Bowman, Chodera, & Voelz (2020a) have created the web application Folding@Home, which allows individuals to donate their computer’s processing power to fighting a variety of viral illnesses including Ebola and SARS-CoV-2 (Cruz et al. 2020).

Bowman, Chodera, & Volez (2020b) describe viral protein foldings as biological nanomachines that are critical to the normal functioning of microbiological activity. The web application not only allows users to donate their processing power, but it also provides open-source reference databases for researchers studying the disease. With much of the country experiencing devastating economic impacts from the COVID-19 pandemic, it should be the obligation of every American to donate their computer device’s surplus processing power to researchers studying viral protein folding structures in the fight against COVID-19.

**Abzymes**

In the late ‘80s research into the structure and function of antibodies began to show that it was possible to program antibodies to display properties of novel enzymatic catalysts. These ‘abzymes’ (from the words antibodies and enzymes) give humans and other animals the ability to hydrolyze DNA and RNA by lowering the chemical transition energy required for a biochemical reaction to take place with a specific type of molecule. The effects of these novel enzymatic catalysts show a tremendous possibility for the treatment of a variety of ailments including retroviruses like HIV and T-cell leukemia virus (Bowen, Wear, & Casadevall, 2017). Is it possible for these catalytic monoclonal antibodies to be programmed to attack the novel coronavirus that is responsible for COVID-19? Recently published research suggests the answer to this question is yes. “Considering the relatively high identity of receptor-binding domains (RBD) in 2019-nCoV and SARS-CoV, it is urgent to assess the cross-reactivity of anti-SARS-CoV antibodies with 2019-nCoV spike protein, which could have important implications for rapid development of vaccines and therapeutic antibodies against 2019-nCoV.” Research in this field is ongoing with new developments occurring daily or weekly. One trend that has emerged across these efforts is that researchers need processing power and lots of it.

The country’s supercomputing infrastructure has been a tremendous asset in the fight against COVID-19, with computational resources being widely diverted from other areas of research. “Typical supercomputing research projects related to predicting hurricanes and earthquakes are running more slowly in order to prioritize COVID-19 projects” (Castellanos, 2020). Although these tasks have taken a back seat to the coronavirus pandemic for now, the need hasn’t gone away, and it is in American’s best interest to relieve the burdens placed on the nation’s supercomputer infrastructure so that it can return to its previously tasked priorities. By using the Folding@Home web application, everyday American citizens can help mitigate the opportunity costs associated with the emergency prioritization of COVID-19-related computational tasks.

**Cloud Computing**

Although many states currently have restrictions on freedom of physical movent, for now, the telecommunications industry is still functional, and up to the task of providing the capability of assisting researchers with distributed data analytics that have traditionally used computational resources provided through prominent cloud services such as Microsoft’s Azure, Amazon’s AWS, and Google’s Cloud Computing Services. While the argument can certainly be made that because of the fact that cloud-service giants are able to provide unmatched Infrastructure as a Service (IaaS) capabilities for computational resources, therefore these services should be leased. However, these companies do not provide their services for free. The cost associated with leasing a powerful IaaS cloud platform can be staggering and takes capital resources away from other cost-intensive areas of the project. By utilizing the Folding@Home app, the processing power required for the project’s big data analytics tasks can be provided to these research efforts, without cost, by individual home users with their own devices (Bowman, Chodera, & Volze, 2020a). The web application not only allows users to donate their processing power, but it also provides open-source reference databases for researchers studying the disease

For an American considering donating their computer’s processing power, some common objections include the energy cost associated with running this web application, security concerns individuals have with downloading untrusted software that is meant to be a parasite on the system’s computational resources, and questions on whether or not running the software could damage a computer system by over-leveraging its resources over a prolonged period of time. Bowman, Chodera, & Volze (2020b) address these questions in-depth on the site’s miscellaneous frequently asked questions page. It turns out, running a standard consumer Computer Processing Unit (CPU) costs about as many watts of power as a typical light bulb. The creators of this service estimate an individual computer’s power usage for a 24 hour period to be approximately $0.36. Privacy concerns are addressed by the service utilizing a 2048 bit digital encryption signature for all inbound and outbound connections, rejecting connections that fail authentication. Beyond the encrypted handshake, however, individuals must inherently trust the service to which they are subscribing. To date, there have not been any widespread accusations of malintent on the researchers behind Folding@Home. In addressing the potential for damage to the host system, Folding@Home is substantially less forthcoming about the possibility that an individual’s computer system could be damaged from using the software. In the ‘Rules & Policies’ section of the website, the following liability disclaimer is stated. “Folding@home was developed by Stanford University. Folding@home and Stanford University assume no liability for damage to your computer, loss of data, or any other event or condition that may occur as a result of participating in Folding@home” (Bowman, Chodera, & Volez, 2020c). This disclaimer is unsettling to read, however, one would be hard-pressed to find a distributed computing service that would accept liability for any hardware failures that their service may cause. In the end, it may be prudent to only use old devices that are no longer in primary usage to support this project. Best practices for running computationally intensive operations include keeping the device under adequate airflow for heat mitigation, in a dust-free room where the device can perform the operation uninterrupted for several hours or days.

**Machine Learning**

The donation of raw processing power is not the only capability ordinary citizens can bring to the fight against COVID-19. Cloud-based Platform as a Service (PaaS) capabilities are at such a developed level that individual users can start their own research into machine learning algorithms of their own to help classify the substantial databases published within the Folding@Home page. Supervised techniques such as linear regression, logistic regression, Bayesian classifiers, and artificial neural networks are widely available to use across multiple cloud provider’s PaaS services, as well as unsupervised techniques such as clustering, dimensional reduction, and deep learning (Hwang & Chen, 2017). Supervised machine learning algorithms are models that are tuned using pre-labeled training data. Supervised models can be useful in the fight against COVID-19 in calculating regressions in epidemiological variables such as mobility, infection rate, hospital resources, and projected deaths. Supervised models are useful for making projections about the future based on the information that is available today. What’s interesting about unsupervised machine learning is that the algorithms will start correlating different data in ways that humans do not. The correlations discovered by unsupervised machine learning algorithms can be useful in finding relationships in the dataset that would have an unlikely probability of being found by a human researcher. The downside of these unsupervised algorithms is that they do not necessarily explain how the connections were discovered, only that they exist. While a useful tool, these machine learning techniques will still need highly skilled human researchers to exploit the insights that these tools can provide.

**Conclusion**

In the wake of COVID-19, it is important for everyone to do their part. For some, this carries more meaning than it does for others. Those that have the education and training in artificial intelligence and machine learning have a unique capability to assist researches working on advances in the understanding of this and other infectious diseases. It is imperative to lend not just a computer’s processing power, but also the unique skills possessed by individuals in order to contribute to the research effort. Professional researchers, academics, and hobbyists alike are able to run their own machine-learning algorithms in hopes of discovering previously unknown correlations relating to protein folding structures of the novel coronavirus. The potential role of PRO95 in the inhibition of enzymatic catalysis is currently a critical area of research in the fight against COVID-19. Researchers like Hifumi et al. could be substantially assisted by the donation of computational processing power in their endeavor to understand the properties of abzymes in creating catalytic monoclonal antibodies that are able to effectively fight the pathogen that causes COVID-19. The ends of this effort can be achieved through the widespread adoption of the Folding@Home web application. It is in the best interest of the individuals donating their computer, smartphone, or machine-learning algorithms to participate in this research project, not only to discover a vaccine against the virus which has caused substantial death and suffering across the globe but also so that the nation can restore priority back to the supercomputing projects that demanded attention before the pandemic began. Research suggests that widespread adoption of distributed supercomputing applications can make a meaningful impact in the fight against COVID-19 and could hasten our return to a normal life. Every American should feel personally compelled to set up the Folding@Home application on their unused computing devices, so that we may hasten the development of a vaccine, and so that the country’s economy can begin to heal. Before the pandemic, America’s economy had been the strongest it has ever been and with a small sacrifice from a large population, we can cooperate and provide a meaningful contribution to researchers who dedicate their lives to defeating the virus and restoring our great society.

References

Bowen, A., Wear, M., & Casadevall, A. (2017). Antibody-Mediated Catalysis in Infection and Immunity. *Infection and Immunity*, *85*(9). Retrieved from <https://iai.asm.org/content/iai/85/9/e00202-17.full-text.pdf>

Bowman, G, Chodera, J, & Volez, V. (2020a). Folding@Home. Retrieved May 10, 2020, from <https://foldingathome.org/>

Bowman, G., Chodera, J., & Volez, V. (2020b). Miscellaneous FAQ. Retrieved May 10, 2020, from <https://foldingathome.org/support/faq/miscellaneous/>

Bowman, G., Chodera, J., & Valez, V. (2020c). Rules & Policies. Retrieved May 10, 2020, from <https://foldingathome.org/support/faq/rules-policies/>

Buneva, V. N., & Nevinsky, G. A. (2017). Exceptional Diversity of Catalytic Antibodies with Varying Activity in the Blood of Autoimmune and Viral Disease Patients. *Molecular Biology*, *51*(6), 840–854. doi: 10.1134/s0026893317060036

Cruz, M. A., Frederick, T. E., Singh, S. I., Vithani, N. R., Zimmerman, M. E., Porter, J. K., … Bowman, G. undefined. (2020). Discovery of a cryptic allosteric site in Ebola’s ‘undruggable’ VP35 protein using simulations and experiments. Retrieved from <https://www.biorxiv.org/content/10.1101/2020.02.09.940510v1/>

Hifumi, E., Taguchi, H., Tsuda, H., Minagawa, T., Nonaka, T., & Uda, T. (2020). A new algorithm to convert a normal antibody into the corresponding catalytic antibody. Science Advances, 6(13). Retrieved from <https://advances.sciencemag.org/content/6/13/eaay6441>

Hwang, K., & Chen, M. (2017). *Big-data analytics for cloud, IoT and cognitive computing*. Hoboken, NJ: John Wiley & Sons Ltd.

The Institute for Health Metrics and Evaluation (IHME). (2020, May 10). IHME: COVID-19 Projections. Retrieved May 10, 2020, from <https://covid19.healthdata.org/united-states-of-america>

Tian, X., Li, C., Huang, A., Xia, S., Lu, S., Shi, Z., … Ying, T. (2020). Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerging Microbes & Infections*, *9*(1), 382–385. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/22221751.2020.1729069>

Usui, D., Inaba, S., Kamatari, Y. O., Ishiguro, N., & Oda, M. (2017). Light-chain residue 95 is critical for antigen binding and multispecificity of monoclonal antibody G2. *Biochemical and Biophysical Research Communications*, *490*(4), 1205–1209. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/28669727>

Walls, A., Xiong, X., Park, Y., Tortorici, M., Snijder, S., Quispe, J., … Veesler, D. (2019). SARS-CoV complex with human neutralizing S230 antibody Fab fragment (state 2). *World Protein Data Bank*. Retrieved from <https://www.wwpdb.org/pdb?id=pdb_00006nb7>

Zheng, J. (2020, March 15). SARS-CoV-2: an Emerging Coronavirus that Causes a Global Threat. Retrieved May 10, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7098030/>

**Appendix**

Liu, X., Zhang, B., Jin, Z., Yang, H., & Rao, Z. (2020). The crystal structure of COVID-19 main protease in complex with an inhibitor N3. *Protein Data Bank in Europe*. Retrieved from <https://www.ebi.ac.uk/pdbe/entry/pdb/6lu7>

Tian, X., Li, C., Huang, A., Xia, S., Lu, S., Shi, Z., … Ying, T. (2020). Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Cold Spring Harbor Labratory*. Retrieved from <https://www.biorxiv.org/content/10.1101/2020.01.28.923011v1>