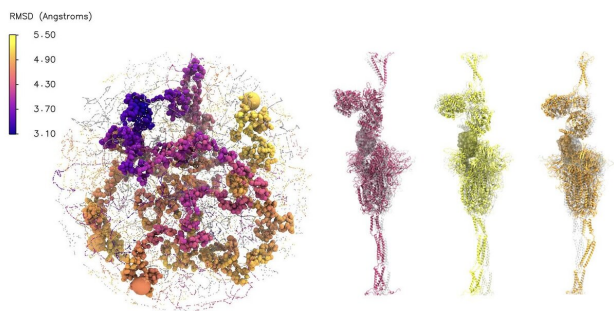


Collaborative AI effort unraveling SARS-CoV-2 mysteries wins Gordon Bell Special Prize

20 November 2020



The latent space learned by the AI model provides a means to understand the conformational changes in the spike protein complex with the ACE2 receptor. Credit: Image produced by Anda Trifan, Argonne/University of Illinois at Urbana Champaign (UIUC); John Stone, UIUC; Lorenzo Casalino and Rommie Amaro, University of California San Diego; Alex Brace and Arvind Ramanathan, Argonne.

The Association for Computing Machinery (ACM) awarded its first ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research to a multi-institution research team that included the U.S. Department of Energy's (DOE) Argonne National Laboratory.

The team was singled out for its work, "AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics," which shines light how the SARS-CoV-2 virus infiltrates the [human immune system](#), setting off a viral chain reaction throughout the body. The award was announced Nov. 19 at SC20, the International Conference for High Performance Computing, Networking, Storage, and Analysis, held virtually this year.

"We are excited to have won this prestigious

award," said Arvind Ramanathan, an Argonne computational biologist and co-principal investigator on the project. "The whole point is to push the boundaries of what we can do with AI. The ability to scale such a huge set of simulations and use AI to drive some factors was key to this work."

Supporting a large collaboration of research organizations and scientific disciplines, Argonne exploring the use of artificial intelligence and [high-performance computing](#) resources to study, in great detail, the complex dynamics of the spike [protein](#), one of the key proteins in the SARS-CoV-2 virus. The research was supported in part by the the DOE's National Virtual Biotechnology Laboratory with funding from the Coronavirus CARES Act.

The team, comprised of nearly 30 researchers across 10 organizations, is trying to understand how that protein binds to and interacts with one of the first point of contacts with the human cell, the ACE2-receptor protein. That binding begins a cascade of events that eventually lets the viral and human cell membranes fuse, allowing the SARS-CoV-2 virus to enter and infect the host.

Proteins aren't static, they have a wide range of motions that span multiple length- and timescales and it's not always understood which motions are important, notes Arvind Ramanathan, an Argonne computational biologist and co-principal investigator on the project. To understand and simulate those actions requires a huge amount of data and computing resources.

Developing a reasonable simulation of the spike protein alone can create a huge system consisting of approximately 1.8 million atoms and simulations can consist of enormous datasets that tax the resources of even the largest supercomputers. In order to make that data more accessible for

interpretation, the team developed a [machine learning method](#) that can summarize large volumes of data.

"One of the key things that this method allowed us to do was to determine what was interesting, what was important, even those things that were not obvious to the human eye," said Ramanathan. "So, when you look deeper using the simulations, you start seeing significant changes in the protein structure, which told us something about how the spike protein opens up such that it can interact with the ACE2 receptor."

As the size of the systems they were working on grew, the team faced challenges of scaling all of the data to run fluidly on today's biggest and best supercomputing systems, as well as their key components.

Because many of the machine learning models they were training on these large simulations needed to be efficiently scaled for use on supercomputers, they partnered with NVIDIA, a leader in GPU and artificial intelligence design, to effectively run the models on Summit, at the DOE's Oak Ridge National Laboratory. The team also utilized many of the top U.S. supercomputers, including Theta at Argonne; Frontera/Longhorn at Texas Advanced Computing Center; Comet at San Diego Supercomputing Center; and Lassen at DOE's Lawrence Livermore National Laboratory, to uncover alternate ways to handle the deluge of data.

"Given the complexity of the data, trying to understand the ACE2 receptor-spike interaction seemed almost impossible at this scale," Ramanathan confided. "One of the things that we clearly showed was that we could actuate a sampling of these dynamical configurations, pushing the idea that we could use AI to bridge these different scales."

The data generated, so far, is providing new insights into how the stalk region of the spike protein changes its overall motions when it interacts with the ACE2 receptor, he said. Eventually, these kinds of insights derived from the highly conjoined combination of machine learning and [simulation](#) will

help facilitate antibody or vaccine discoveries.

The team's article, "AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics," will appear in the *International Journal of High Performance Computing Applications*, 2020.

Provided by Argonne National Laboratory

APA citation: Collaborative AI effort unraveling SARS-CoV-2 mysteries wins Gordon Bell Special Prize (2020, November 20) retrieved 29 November 2020 from <https://phys.org/news/2020-11-collaborative-ai-effort-unraveling-sars-cov-.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.