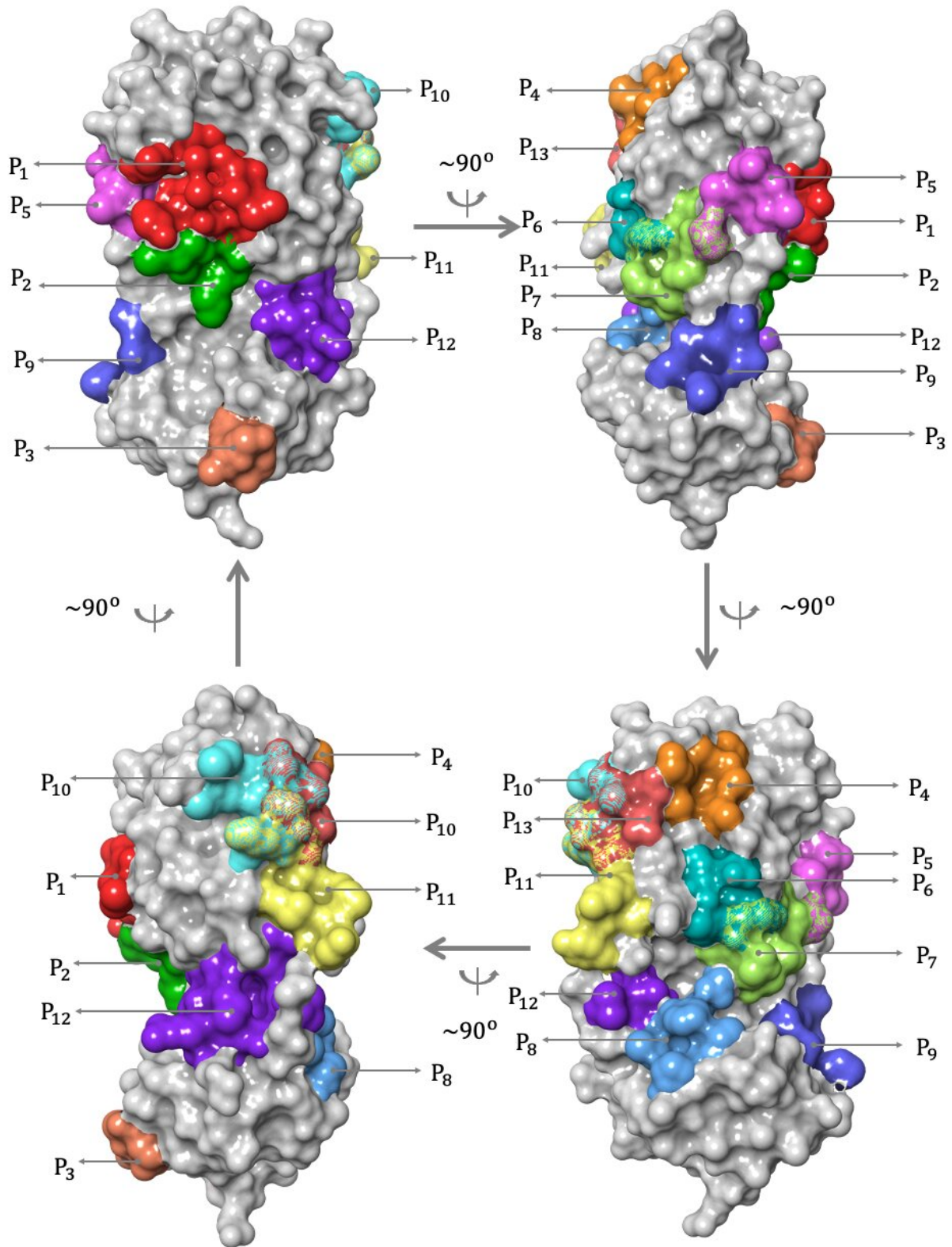# Machine learning helps hunt for COVID-19 therapies

October 28 2020, by Matt Davenport

The colored regions mark the regions where potential drugs bind to the

coronavirus's main protease, shown in white, as predicted by MSU deep learning models. Credit: Guowei Wei's research team

Michigan State University Foundation Professor Guowei Wei wasn't preparing machine learning techniques for a global health crisis. Still, when one broke out, he and his team were ready to help.

The group already has one machine learning model at work in the pandemic, predicting consequences of mutations to SARS-CoV-2. Now, Wei's team has deployed another to help drug developers on their most promising leads for attacking one of the virus' most compelling targets. The researchers shared their intel Oct. 21 in the peer-reviewed journal *Chemical Science*.

Prior to the pandemic, Wei and his team were already developing machine learning computer models—specifically, models that use what's known as deep learning—to help save drug developers time and money. The researchers "train" their deep learning models with datasets filled with information about proteins that drug developers want to target with therapeutics. The models can then make predictions about unknown quantities of interest to help guide drug design and testing.

Over the past three years, the Spartans' models have been among the top performers in a worldwide competition series for computer-aided drug design known as the Drug Design Data Resource, or D3R, Grand Challenge. Then COVID-19 came.

"We knew this was going to be bad. China shut down an entire city with 10 million people," said Wei, who is a professor in the Departments of Mathematics and Electrical and Computer Engineering. "We had a technique at hand, and we knew this was important."

Wei and his team have repurposed their deep learning models to focus on a specific SARS-CoV-2 protein called its main protease. The main protease is a cog in the coronavirus's protein machinery that's critical to how the pathogen makes copies of itself. Drugs that disable that cog could thus stop the virus from replicating.

What makes the main protease an even more attractive target is that it's distinct from all known human proteases, which isn't always the case. Drugs that attack the viral protease are thus less likely to disrupt people's natural biochemistry.

Another advantage of the SARS-CoV-2 main protease is that's it's nearly identical to that of the coronavirus responsible for the 2003 SARS outbreak. This means that drug developers and Wei's team weren't starting completely from scratch. They had information about the structure of the main protease and chemical compounds called protease inhibitors that interfere with the protein's function.

Still, gaps remained in understanding where those protease inhibitors latch onto the viral protein and how tightly. That's where the Spartans' deep learning models came in.

Wei's team used its models to predict those details for over 100 known protease inhibitors. That data also let the team rank those inhibitors and highlight the most promising ones, which can be very valuable information for labs and companies developing new drugs, Wei said.

"In the early days of a drug discovery campaign, you might have 1,000 candidates," Wei said. Typically, all those candidates would move to preclinical tests in animals, then maybe the most promising 10 or so can safely advance to clinical trials in humans, Wei explained.

By focusing on drugs that are most attracted to the protease's most

vulnerable spots, drug developers can whittle down that list of 1,000 from the start, saving money and months, if not years, Wei said.

"This is a way to help drug developers prioritize. They don't have to waste resources to check every single candidate," he said.

But Wei also had a reminder. The team's models do not replace the need for experimental validation, preclinical or clinical trials. Drug developers still need to prove their products are safe before providing them for patients, which can take many years.

For that reason, Wei said, antibody treatments that resemble what immune systems produce naturally to fight the coronavirus will be most likely the first therapies approved during the pandemic. These antibodies, however, target the virus's spike protein, rather than its main protease. Developing protease inhibitors would thus provide a welcome addition in an arsenal to fight a deadly and constantly evolving enemy.

"If developers want to design a new set of drugs, we've shown basically what they need to do," Wei said.