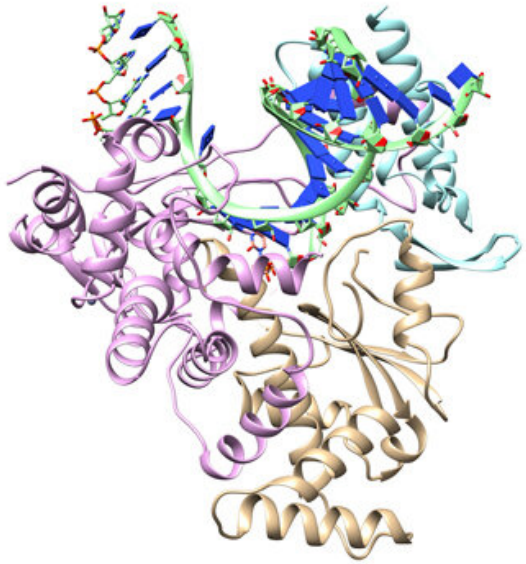
Pulling the plug on the coronavirus copy machine

September 19 2020, by Jorge Salazar



Antiviral drug remdesivir forms the main line of FDA-approved therapeutic defense against the COVID-19 virus. Researchers at the University of North Texas are using the Frontera supercomputer to model how remdesivir blocks coronavirus reproduction, in the hopes of developing improvements on the drug. Shown here are the crystal structures of the RNA-dependent RNA Polymerase ternary complex model with double stranded RNA and incoming remdesivir triphosphate. Credit: Cisneros Research Group, UNT

Key proteins used by coronavirus for its reproduction being modeled on NSF-funded Frontera supercomputer by Andres Cisneros research group

of the University of North Texas. Research goals include finding ways to improve on COVID-19 therapeutic remdesivir. NSF-funded Frontera allocation awarded to Cisneros through the COVID-19 High Performance Computing Consortium.

In May 2020, the U.S. Food and Drug Administration authorized the antiviral drug remdesivir for emergency treatment of COVID-19, one of only four therapeutics currently with this status. Remdesivir stops the chemical machinery that the coronavirus uses to copy itself, binding to an enzyme that does the assembly. While remdesivir has shown promise in helping patients recover from COVID-19, scientists are investigating ways to improve its effectiveness.

A team of scientists led by G. Andres Cisneros of the University of North Texas is modeling the key parts of the coronavirus that it uses to copy itself. The simulations are being done on the Stampede2 and Frontera supercomputers at the Texas Advanced Computing Center (TACC).

"We were very fortunate to be granted an allocation on Frontera to be able to work on investigating the mechanism of drugs that target two specific proteins in COVID-19," Cisneros said. His work investigates how remdesivir and other available drugs inhibit the proteins NSP-12 and the main protease, both enzymes the coronavirus needs for replication. "By looking at how these drugs do their work, perhaps this information can be used to improve upon them."

The NSP-12 protein puts together the nucleotides that make up viral RNA, abbreviated as A, U, G, and C, building complete sets of genetic material for new coronavirus copies. NSP-12 is actually part of a larger structure called the RNA-dependent RNA polymerase (RDRP) that copies the complete RNA. Remdesivir binds with RDRP, plugging up the machinery.

"We're investigating how this process happens," said Cisneros. "By doing this, perhaps there might be a way for us and other scientists to come up with ideas on whether and how remdesivir can be improved."

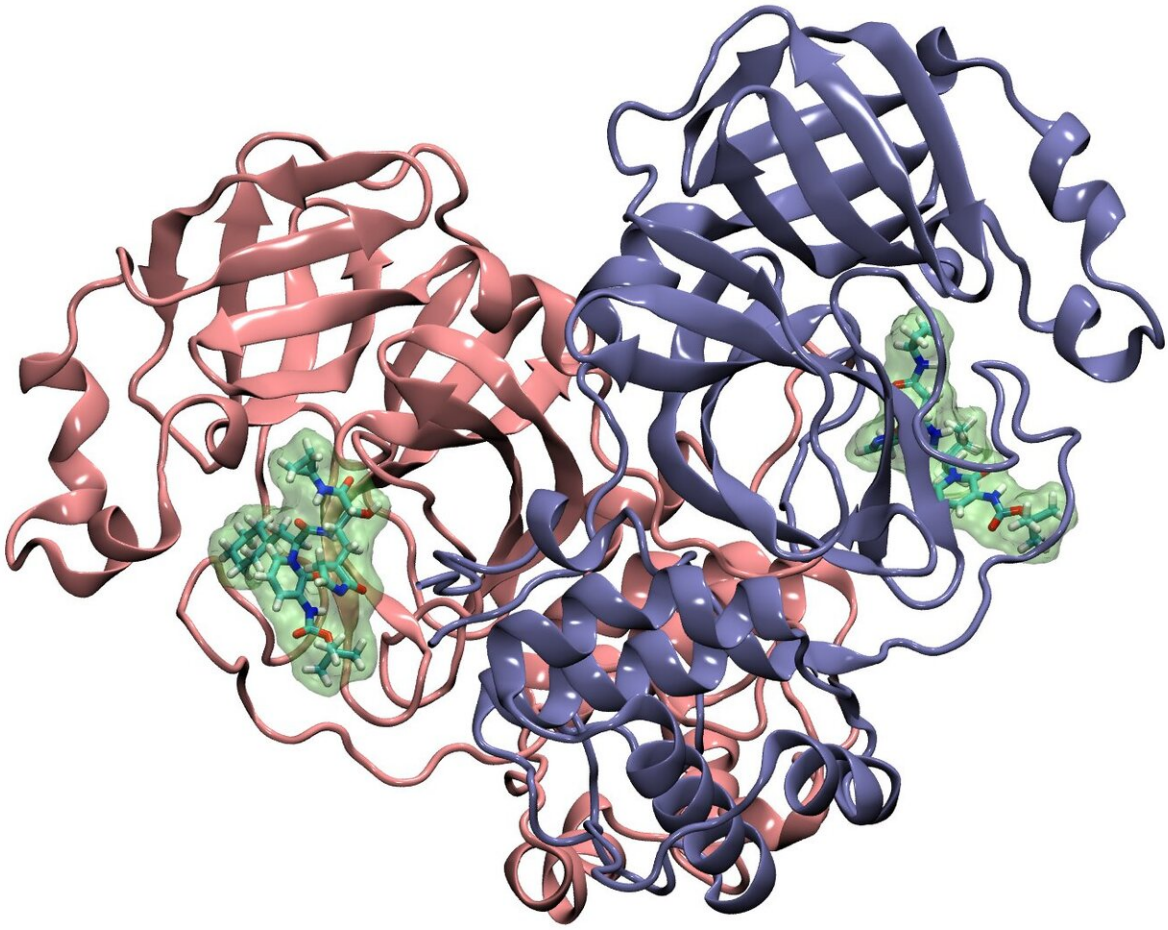
The other protein Cisneros is studying is called the main protease. It separates a polyprotein produced by SARS-CoV-2 translated from viral RNA into functional proteins that put 'meat' on its viral bones. Stop the protease, and you stop the virus from forming. This makes it a great drug target.

Cisneros explained that he uses the basic math and physics of Newton's equations and quantum mechanics to calculate the properties of the proteins, including everything relevant to its functioning, such as the RNA and water. An approach called classical molecular dynamics uses Newton's equations to simulate how the proteins move and interact dynamically in time. "We're talking about systems that we simulate that are in the hundreds of thousands of atoms," Cisneros said.

He also simulates the chemical reactions inside the proteins to investigate how the drugs stop RDRP or the protease. A hybrid method called QM/MM (quantum mechanics/molecular mechanics) saves computational time and money by focusing more intently on interactions at the active site, using the more approximate straight molecular dynamics for everything else.

The Cisneros group developed and maintains a program called LICHEM that lets them use the QM/MM approach. "One of the features of LICHEM is that it allows us to use approaches for the classical mechanics part that include a better description of the physics that are happening between the molecules in the classical environment, specifically, the AMOEBA potential" Cisneros said. AMOEBA is developed by Pengyu Ren of UT Austin; Jay Ponder of the University of Washington; and Jean-Philip Piquemal at Sorbonne University in Paris

with contributions from the Cisneros group for ionic liquids.



Another research target of the Cisneros group being modeled on Frontera is a protein called the main protease. It cleaves a polyprotein produced by the virus that build up the functional proteins of the copies of itself it generates. Crystal structure of the coronavirus main protease with bound Inhibitor shown here. Credit: Cisneros Research Group, UNT

"Frontera, with not only compute power but the intercommunication between the nodes, allows us to run these QM/MM calculations with much higher, not only speed, but also throughput," Cisneros said. Frontera freed them to run multiple systems at a time. "In my group, I have five different scientists, graduate students and post docs, that are working on both of these systems, but in different pieces of the puzzle. All of them are able to access these resources. It's definitely very useful, and we very much appreciate the allocation."

What got Cisneros going was news in April of 2020 of the crystal structure of the SARS-CoV-2 RDRP being reported. "I contacted my group and told them that with this information, there's something we can do to help with the pandemic," he said.

Within two days of the news, Cisneros successfully proposed his research on coronavirus drug targets to the COVID-19 High Performance Computing Consortium. Dozens of national and international supercomputing facilities, industry, and organizations including TACC have volunteered their resources to the consortium in support of scientists' effort to combat the coronavirus.

The allocation was initially awarded just on TACC's Stampede2, the supercomputing flagship of the National Science Foundation (NSF) that's ranked 21st fastest in the world and #2 for academic systems according to the Top500. "Then we were contacted by TACC and grateful that we were granted access to Frontera. Now we have access to both systems, which is really great," Cisneros said.

The Frontera supercomputer is the #1 fastest academic supercomputer and #8 fastest worldwide. Both Frontera and Stampede2 are funded by the NSF.

"We're very happy with this system. We were able to transfer some of

the knowledge that we had from Stampede2 to Frontera," Cisneros said. One of his recently graduated students, Erik Vazquez Montelongo, set up all of the calculations for LICHEM on Frontera based on what he learned on Stampede2. "That really has been a boon. Frontera for our calculations has been running really well. We're really happy with it."

One of the postdocs in The Cisneros Group, Sehr Nazeem-Kahn, generated the model for RDRP, the remdesivir and other drug candidates, all in the active site. With that in hand, they started running simulations.

"We were very happy to see that her model was actually very close to the experimental structure. That's really useful for us, because it validates the model that has been built by the group and shows that we are on the right track," he added.

Currently, Dr. Naseem-Khan is running molecular dynamics simulations of this model with remdesivir on Frontera. "We are also starting with our QM/MM calculations for RDRP. In the case for the main protease, there were structures that also needed to be modeled and subsequently were confirmed. That was also very satisfying," Cisneros said.

With that structure data, they're looking at six different inhibitor molecules. "One of those, we're already starting QM/MM calculations on Frontera, and another one on Stampede2," Cisneros said. If all goes well, he's hoping to get results in the next five to six months. "These are very expensive calculations," he added. "Also, running the analysis takes time. If we were to use just the resources at home, it would take several years."

Provided by Texas Advanced Computing Center