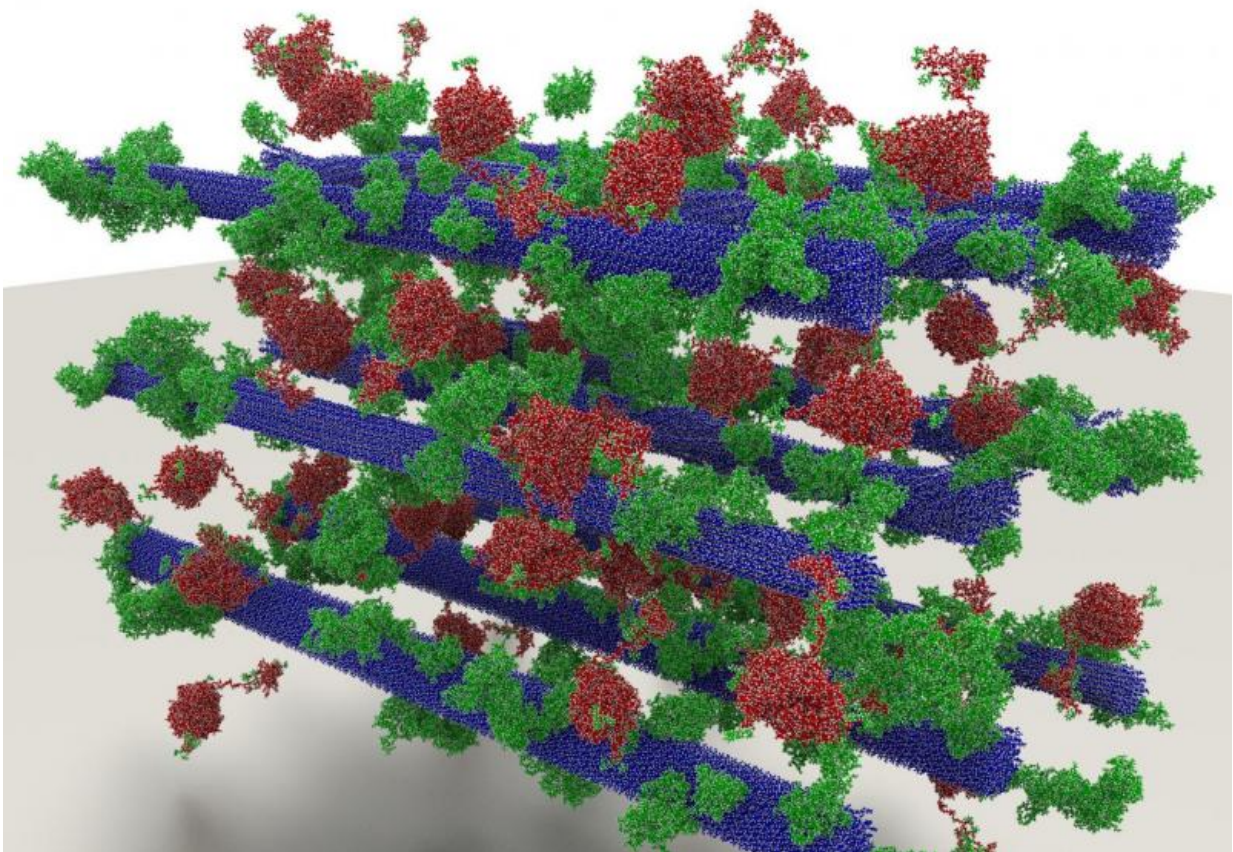


# Supercomputer probes depths of biofuel's biggest barrier

February 16 2016

---



Researchers used experimental data to create a 23.7-million atom biomass model featuring cellulose (purple), lignin (brown), and enzymes (green). Credit: Mike Matheson, ORNL

Ask a biofuel researcher to name the single greatest technical barrier to cost-effective ethanol, and you're likely to receive a one-word response: lignin.

Cellulosic ethanol—fuel derived from woody plants and waste biomass—has the potential to become an affordable, renewable transportation fuel that rivals gasoline, but [lignin](#), one of the most ubiquitous components of the plant cell wall, gets in the way.

In nature, the resilient lignin polymer helps provide the scaffolding for plants, reinforcing slender cellulosic fibers—the primary raw ingredient of cellulosic ethanol—and serving as a protective barrier against disease and predators. Lignin's protective characteristics persist during biofuel processing, where it's a big hindrance, surviving expensive pretreatments designed to remove it and blocking enzymes from breaking down cellulose into simple sugars for fermentation into bioethanol.

To better understand exactly how lignin persists, researchers at the US Department of Energy's (DOE's) Oak Ridge National Laboratory (ORNL) created one of the largest biomolecular simulations to date—a 23.7-million atom system representing pretreated biomass (cellulose and lignin) in the presence of enzymes. The size of the simulation required Titan, the flagship supercomputer at the Oak Ridge Leadership Computing Facility (OLCF), a DOE Office of Science User Facility, to track and analyze the interaction of millions of atoms.

The research, led by Jeremy Smith, a Governor's Chair at the University of Tennessee (UT) and director of the UT-ORNL Center for Molecular Biophysics, revealed in atomistic detail why lignin is such a problem: Not only does it bind to cellulose in the preferred locations sought by enzymes, but lignin also attracts and occupies the cellulose-binding domain of the enzymes themselves.

"That impedes the mechanism the [enzyme](#) has to anchor to cellulose. Thus lignin binds exactly where it is least desired for industrial purposes," said ORNL staff scientist Loukas Petridis. "This detailed knowledge of lignin behavior can guide genetic engineering of enzymes that bind less to lignin and therefore produce bioethanol more efficiently."

Beyond the scientific knowledge obtained from the simulation, the team's biomass system advances computational biophysics' shift toward complex, multicomponent systems, a move enabled by leadership-class supercomputers.

## **Building a Biomass Model**

During pretreatment, acid, water, and heat work to remove non-cellulosic biomass from plant material. Lignin, however, sticks around, clustering into aggregates around the cellulose and impeding enzymes from reaching cellulose.

To accurately model this crowded environment, Smith's team used experimental data to create a representative sample of pretreated biomass and enzymes. The model took into account details such as the ratio of cellulose to lignin, type of lignin, and relative amount of enzymes. In total, the simulation tracked nine cellulose fibers, 468 lignin molecules, and 54 enzyme molecules in a rectangular water box.

The team built the model using a molecular dynamics code called GROMACS under an allocation awarded through the Innovative and Novel Computational Impact on Theory and Experiment, or INCITE, program. With a complete model, the team turned to the Cray XK7 Titan, America's fastest supercomputer, to supply the necessary computing power to observe the system in action.

During its largest runs, the biomass simulation scaled to nearly 4,000 of Titan's 18,666 nodes, producing roughly 45 nanoseconds of simulation time in one day. Over the course of a year, the team amassed 1.3 microseconds of simulation time, a significant length of time in the world of computational biophysics.

"There's nowhere else in the world where we could have run this simulation," Petridis said.

In addition to lending insight to the challenges of next-generation biofuels, the team's simulation pointed toward potential pathways that could help mitigate lignin's impact. Specifically, the simulation demonstrated that lignin does not bind as much to less-ordered, or amorphous, [cellulose fibers](#), meaning it competes less with the enzymes there.

"Industrialists knew amorphous cellulose is more easily broken down by enzymes, but what we show is that it's not only the inherent properties of amorphous cellulose that makes it easier for the enzymes but also that lignin is less of a pest," Petridis said.

## **Analysis in Parallel**

To maximize their time on the OLCF's flagship supercomputer, Smith's team tweaked GROMACS to streamline communication across thousands of Titan's CPU cores. Additionally, the team doubled the time interval GROMACS used to calculate the motion of the biomass system. By implementing a more computationally efficient method to track long-range interactions between atoms, the team was able to increase its timestep from 2 femtoseconds to 4 femtoseconds, or 4,000 trillionths of a second, without losing accuracy.

The resulting data was transferred to the OLCF's High-Performance

Storage System until it could be analyzed. Typically, analysis is carried out in serial, or one event a time, but growth in computing power and simulation size has created an analysis bottleneck—it just takes too much time.

To get around this constraint, Smith's team worked to equip GROMACS with the capability to conduct analysis in parallel, meaning thousands of Titan's processors could work in tandem to carry out analysis tasks. For example, running parallel analyses on 2,000 CPU cores, the researchers could obtain results 2,000 times faster than conventional methods. In collaboration with the ORNL team, Josh Vermaas, a graduate student at the University of Illinois at Urbana-Champaign, contributed significantly to this effort as a DOE Computational Science Graduate Fellow at ORNL.

The new capability not only helped the team reduce its time to solution, but it also paves the way for analyzing similar large-scale simulations in the future. "Analysis was one of the stumbling blocks for simulations at this scale," said team member Roland Schulz, a UT postdoctoral researcher. "With parallel analysis, it's now more feasible and will make leadership-class simulations easier."

As supercomputers allow for larger and more realistic systems, the ambitions of researchers and the realism of their biological systems continue to rise. Summit, the OLCF's next leadership-class supercomputer, will offer at least five times the [computing power](#) of Titan. For Smith's team, that means its biomass models have room to grow in complexity to further probe biofuel's challenges.

"We're trying to reach the complexity that is found in nature and industrial conditions," Petridis said. "Eventually, we would like to construct a simple model of a plant cell wall that we could process in silico, or via computer simulation, and see how it changes during

pretreatment."

**More information:** Josh V. Vermaas, Loukas Petridis, Xianghong Qi, Roland Schulz, Benjamin Linder, and Jeremy C. Smith, "Mechanism of lignin inhibition of enzymatic biomass deconstruction." *Biotechnology for Biofuels* 8, no. 1 (2015): 1.

Provided by Oak Ridge National Laboratory

Citation: Supercomputer probes depths of biofuel's biggest barrier (2016, February 16) retrieved 26 June 2024 from

<https://phys.org/news/2016-02-supercomputer-probes-depths-biofuel-biggest.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.