Meeting the Ethanol Challenge

Scientists use SDSC supercomputer to target cellulose bottleneck and speed ethanol production

by Paul Tooby

Termites and fungi already know how to digest cellulose, but the human process of producing ethanol from cellulose remains slow and expensive. The central bottleneck is the sluggish rate at which the cellulose enzyme complex breaks down tightly bound cellulose into simple sugars, which are then fermented into ethanol.

To help unlock the cellulose bottleneck, a team of scientists has conducted molecular simulations at the San Diego Supercomputer Center (SDSC) at UC San Diego. By using "virtual molecules," they have discovered key steps in the intricate dance in which the enzyme acts as a molecular machine—attaching to bundles of cellulose, pulling up a single strand of sugar, and putting it onto a molecular conveyor belt where it is chopped into smaller sugar pieces.

"By learning how the cellulase enzyme complex breaks down cellulose, we can develop protein engineering strategies targeted to speed up this key reaction," said Mike Cleary, who is coordinating SDSC's role in the project. "This is important in making ethanol from plant biomass a realistic 'carbon neutral' alternative to the fossil petroleum used today for transportation fuels."

The results of the research were reported in the April 12 online edition of the *Protein Engineering, Design and Selection* journal, which featured visualizations of the results on the cover.

A convergence of factors from looming global warming to unstable international oil supplies is driving a surge in renewable biofuels such as ethanol, with worldwide ethanol production more than doubling between 2000 and 2005.

To date, corn has been the favorite ethanol source. While good news for farmers, corn prices have doubled in the past two years, and consumers worldwide are feeling the pinch as food prices climb. In poorer countries this can be more than an inconvenience, giving rise to real hardships.

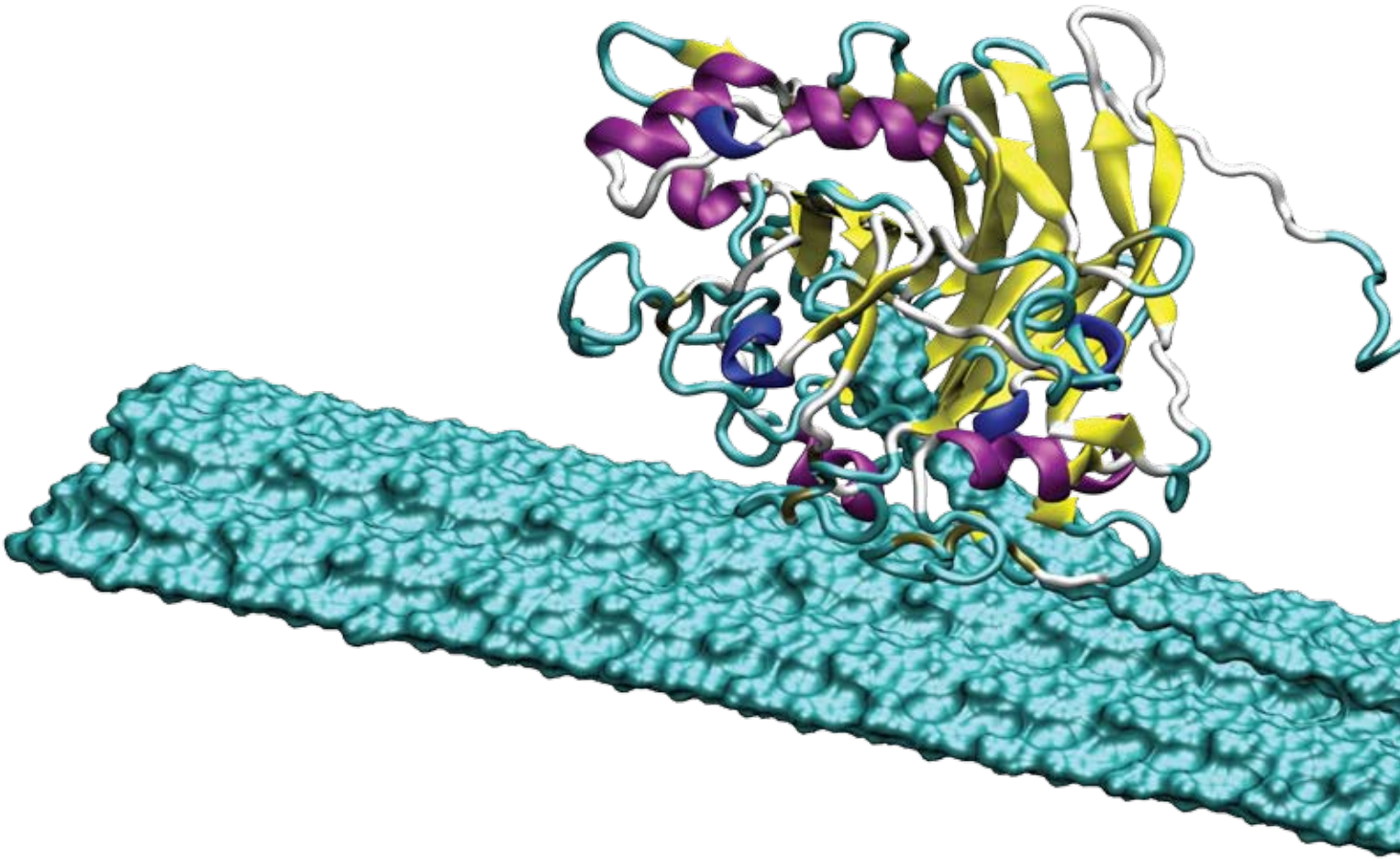
A far better source is to produce ethanol from cellulose, easing pressure on food supplies and yielding greater greenhouse gas benefits. The fibrous part that makes up the bulk of plants, cellulose is the cheapest and most abundant plant material, found in everything from corn stalks left after harvest to wood chips from papermills and fast-growing weeds.

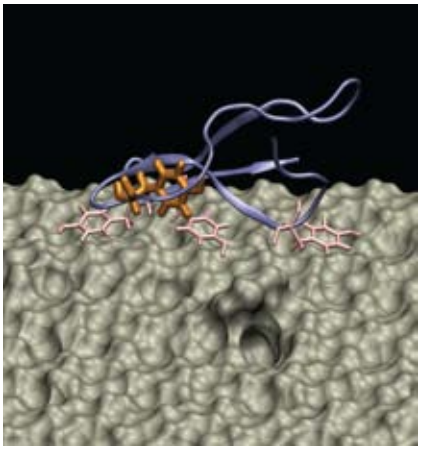
Virtual Molecules

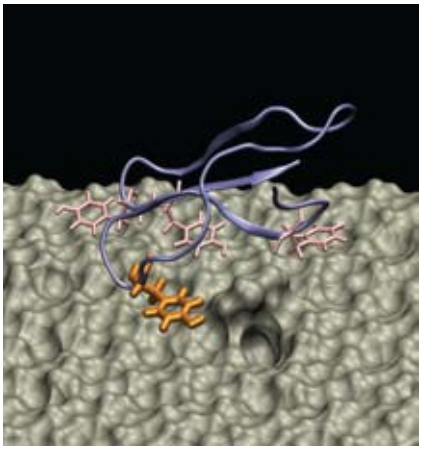
"Our simulations have given us a better understanding of the interactions between the enzyme complex and cellulose at the molecular level—the computer model showed us how the binding portion of this enzyme changes shape, which hadn't been anticipated by the scientific community," said first author Mark Nimlos, a senior scientist at NREL. "These results are important because they can provide crucial guidance as scientists formulate selective experiments to modify the enzyme complex for improved efficiency."

The simulations—a "virtual microscope" that lets the scientists zoom in on previously invisible details—showed that initially the binding module of the enzyme complex moves freely and randomly across the cellulose surface, searching for a broken cellulose chain. When the binding module encounters an available chain, the cellulose seems to prompt a change in

Converting Biomass to Ethanol
Scientists used SDSC's DataStar supercomputer to explore ways to improve cellulose conversion to ethanol for cheap and clean energy. The researchers created a "virtual molecule" to model the smaller binding module (left) as it searches for and attaches to a broken strand of cellulose. The larger catalytic module then digests the cellulose strand, breaking it into smaller sugars which can be fermented into ethanol.
Image: R. Walker and A. Chourasia SDSC/UCSD.







Finding a Foothold on Cellulose
Two snapshots from a simulation of the binding module of the cellulase enzyme complex attaching to a broken strand of cellulose. The earlier image (left) shows the module "searching" randomly for a broken strand. The later image (right) shows the module bound to the broken strand, with a fourth amino acid (orange) unfolding from inside the protein and coordinating with the surface. In a new insight, the simulations suggest that identifying the broken strand occurs via an interactive, induced fit mechanism. Image: R. Walker & A. Chourasia, SDSC/UCSD.

the shape of the binding module so that it straddles the broken end of the cellulose chain. Binding to the broken chain gives the enzyme complex a crucial foothold to begin the process of digesting or "unzipping" the cellulose into sugar molecules.

Finding this induced-fit binding mechanism was a key insight from the computations, according to SDSC computational scientist Ross Walker, who worked on the research in an SDSC Strategic Applications Collaboration.

Massive Simulations

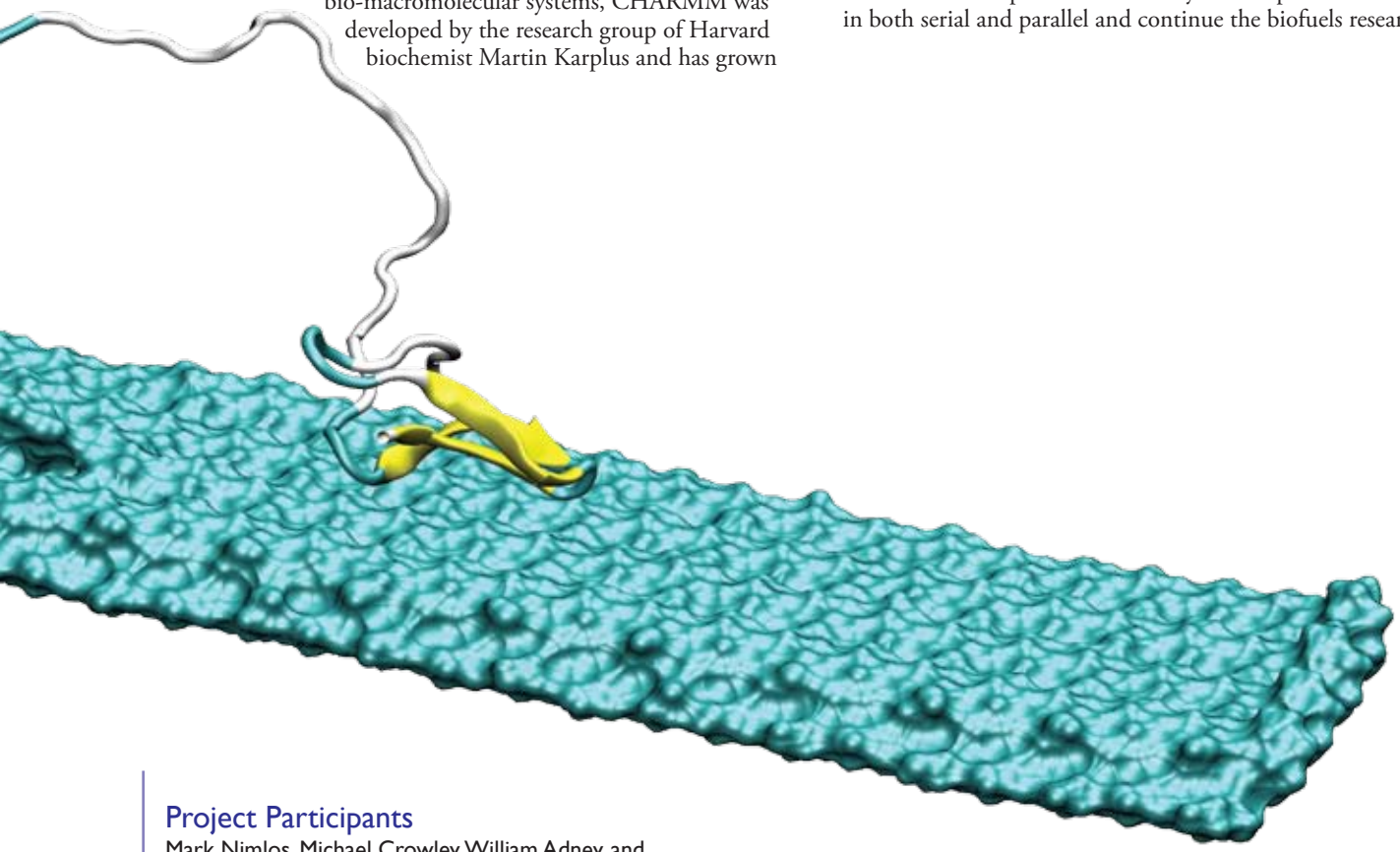
To the scientists, the simulation is like a stop-motion film of a baseball pitcher throwing a curveball. In real-life this process occurs far too quickly to evaluate visually, but by using the supercomputer simulations to break the throw down into a step-by-step process, the scientists can see the precise details of the role of velocity, trajectory, movement, and arm angle.

To undertake the large-scale simulations, the researchers used the CHARMM (Chemistry at HARvard Molecular Mechanics) suite of modeling software. One of the first programs developed for modeling the motions and mechanics of bio-macromolecular systems, CHARMM was developed by the research group of Harvard biochemist Martin Karplus and has grown into a major software suite under continued development by a large, distributed group of investigators, including Charles Brooks and colleagues at The Scripps Research Institute (TSRI).

According to the researchers, an accurate understanding of the key molecular events required the simulations to run for some six million time steps to span 12 nanoseconds (billionths of a second), allowing them to capture enough of the motion and shape changes of the enzyme as it interacted with the cellulose surface.

Twelve nanoseconds is an extremely long time in molecular terms, and the computation-hungry simulations consumed some 80,000 processor-hours running on SDSC's DataStar supercomputer.

The research was partially funded by the Department of Energy's Biomass Program and the National Science Foundation. To continue their work, the researchers have been awarded a five year collaborative DOE grant, headed by NREL and incorporating Charles Brooks at The Scripps Research Institute, John Brady at Cornell, and Ross Walker at SDSC. The work will improve the efficiency and scope of CHARMM in both serial and parallel and continue the biofuels research.



Project Participants

Mark Nimlos, Michael Crowley, William Adney, and Michael Himmel, DOE/NREL; James Matthews and John Brady, Cornell; Linghao Zhong, Penn State; Ross Walker and Mike Cleary, SDSC/UCSD; and Giridhar Chukkapalli, Sun Microsystems

Related Link

National Renewable Energy Laboratory http://www.nrel.gov/