accelerating scientific discovery

XSEDE

Extreme Science and Engineering Discovery Environment

XSEDE is a virtual organization that provides a dynamic distributed infrastructure, support services, and technical expertise that enable researchers, engineers, and scholars to address the most important and challenging problems facing the nation and world. **XSEDE** supports a growing collection of advanced computing, high-end visualization, data analysis, and other resources and services.

XSEDE: Man and Machine

XSEDE supports a collection more than 20 supercomputers, high-end visualization and data analysis and storage resources across the country, with the number and type of resources evolving all the time. Researchers, scientists and engineers are able to request allocations, run applications, and transfer and manage both code and data across the breadth of XSEDE—all to facilitate scientific discovery.

Uniting so many resources under the XSEDE banner streamlines processes for XSEDE's users. Victor Hazelwood, XSEDE operations deputy director, points out, "If XSEDE did not exist then there would be 22 different ways to request allocations on the different resources; 22 different formats for user documentation on how to use the resources; 22 different authentication credentials that users would have to keep up with; 22 different customer service desks to use to get answers to resource use questions."

XSEDE is much more than the computational, networking, and data infrastructure it supports. It's also a collection of services and programs that make those resources easier to access and use for a broader range of scientists, engineers, scholars, and students. The XSEDE Champions and Champion Fellows, countless training and educational opportunities, science gateways, Campus Bridging efforts, and Extended Collaborative Support Services all combine to give more people more knowledge of and access to the power of digital knowledge discovery. "Little fish turn into big fish. We can give a chance to that small researcher to go out there and swim around the pond," explains Ken Hackworth, XSEDE allocation manager. "And there are a lot of places to swim. Stampede, Kraken, Blacklight, Gordon, Trestles, Keeneland, whichever XSEDE resources it is, it gives them computation experience, expands their knowledge, and advances their research. That's where big fish are born."

XSEDE lowers technological barriers to the access and use of computing resources, and more and more virtual doors are opening every day for a wide variety of users. XSEDE offers researchers the resources, services, and collaboration support they need to be productive.

EXTENDED COLLABORATIVE SUPPORT SERVICES (ECSS)

ECSS pairs members of the XSEDE user community with expert staff members for an extended period to work together to solve challenging science and engineering problems through the application of cyberinfrastructure. Expertise is available in a wide range of areas, from performance analysis and petascale optimization to the development of community gateways and work and data flow systems.

CAMPUS CHAMPIONS

The Champions program supports campus, regional, and domain representatives as a local source of knowledge about XSEDE's high-performance and high-throughput computing and other digital services, opportunities and resources. The program is 200+ strong and growing with Champions found in 49 of the 50 United States at more than 150 institutions.

CAMPUS CHAMPION FELLOWS

The fellows program partners Champions with ECSS staff and research teams to work side by side on research projects. The Fellows also develop a high level of expertise within a particular aspect of what XSEDE has to offer for them to share with the students, administrators, staff and faculty on their campuses.

CAMPUS BRIDGING

The XSEDE Campus Bridging team is part of an initiative intended to lower the barriers to utilization of XSEDE resources by researchers and to disseminate the best practices of XSEDE resources to campus IT staff. The team focuses on advocating a common means of access, gathering requirements from campus and local users, and providing opportunities to interact with new software for use in XSEDE.

SCHOLARS PROGRAM

The XSEDE Scholars Program (XSP) is a program for U.S. students from underrepresented groups in the area of computational sciences that provides opportunities to learn more about high-performance computing and XSEDE resources, network with cutting-edge researchers and professional leaders, and belong to a cohort of student peers to establish a community of academic leaders.

TRAINING

The training classes focus on systems and software supported by the XSEDE Service Providers, covering programming principles and techniques for using resources and services effectively. Training classes are offered in high-performance computing, visualization, data management, distributed and grid computing, science gateways, and more.

SCIENCE GATEWAYS

A Science Gateway is a community-developed set of tools, applications, and data that are integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a specific community. Gateways enable entire communities of users associated with a common discipline to use national resources through a common interface that is configured for optimal use.

XSEDE USER PORTAL

The XSEDE User Portal provides XSEDE users access to view and manage their accounts and allocations via the web, as well as find information about and access the XSEDE services and resources.

THE MACHINES

name	*location*	*specialty*
Blacklight	**PSC**	**Large shared memory**
Gordon	**SDSC**	**I/O adaptability and large-memory capability**
Keeneland	**NICS**	**Accelerator performance codes**
Kraken	**NICS**	**Massively-parallel processing**
Lonestar	**TACC**	**HTC and remote visualization**
Mason	**IU**	**Large memory, serial processes**
Quarry	**IU**	**Hosts Scientific Gateway and Web Service allocations**
Stampede	**TACC**	**Massively-parallel workflow support**
Trestles	**SDSC**	**Fast local I/O**

The Bridge to Biofabrication

In the near future, medical professionals may be able to send 3D-image data containing patient-specific needs to a lab for the automated manufacture of new tissues to replace diseased, damaged or congenitally defective tissue structures, according to William Mondy, director of Adiana Research and Development and a professor at the Medical University of South Carolina.

Mondy's strategy for tissue biofabrication begins with atomic 3D images of vascular corrosion casts, representations of the geometry of the original vascular system. The images of the casts are converted into computer-aided-design models and combined with developmental chemical-engineering data in a bio-informative blueprint capable of guiding the computer-aided manufacture (CAM) of tissues using 3D robotic bio-printers and multi-photon lasers.

The chemical-engineering data influence how cells move, grow and differentiate based on tissue type. Those activities are responsible for the changes in cell form and structure that take place when new tissue cells are formed and when damaged tissue is restored to good health.

According to Mondy's strategy, the data in the bio-informative blueprint instruct CAM tools to create environments in bioreactors that mimic molecular mechanisms and cellular interactions that would normally take place in the human body when tissue is being developed or regenerated. And then CAM uses 3D robotic bio-printers and multi-photon lasers to produce the desired outcome: authentic 3D tissue structures.

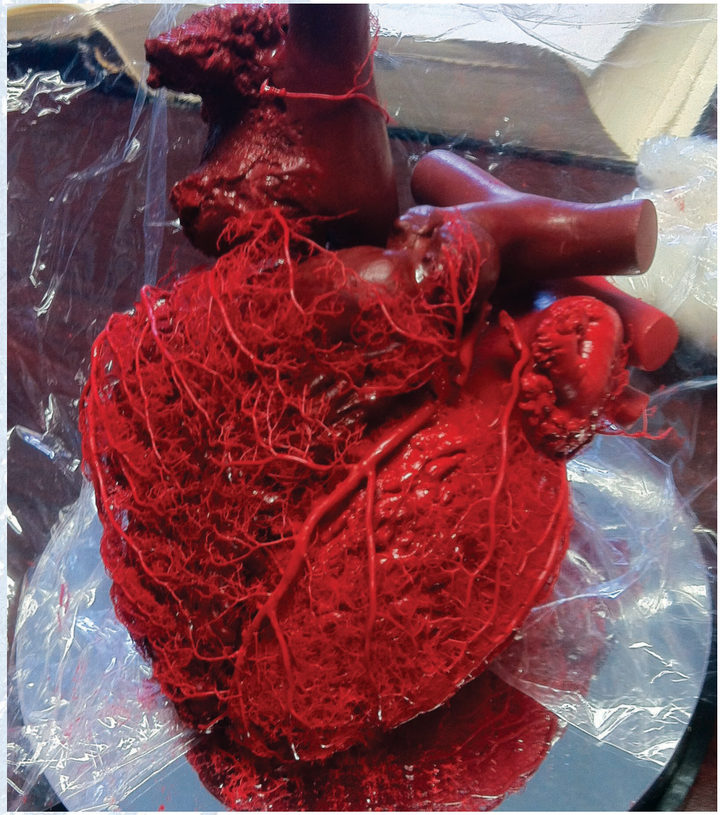
The current focus of the biofabrication project is on overcoming impediments. The limiting factors have been the need for terabytes of random-access memory and greater computing power to process the huge sets of micro-CT data and 3D stereolithographic (STL) files. Through XSEDE's Extended Collaborative Support Services program, a team from the National Institute for Computational Sciences (NICS) is supplying the high-performance-computing (HPC) expertise essential to resolving the project's issues, while XSEDE provides the needed allocation of compute time on the NICS-managed Nautilus supercomputer.

NICS team leader Bhanu Rekepalli explains that currently most of the analysis tools used for biofabrication are not parallelized, and the tools required for generating 3D CAD models are not located on a single resource or at one center but are distributed across multiple resources at various centers. Other major hindrances, he said, are the manual submission of jobs one at a time and large-scale data movement across centers.

"Our team is proposing a new workflow to create a sustainable and robust automated pipeline to enable large-scale 3D model development and downstream analysis at NICS," Rekepalli said. "We're also developing web-enabled science gateways for researchers to use these optimized parallel applications on the NICS HPC resources. That easy access and use will broaden the research community by engaging persons who are not keen on operating a supercomputer. The web-enabled science gateways will also host databases for easy access to generated 3D models by researchers worldwide, along with analyzed datasets for the broader research community."

NICS team member Pragneshkumar Patel adds: "We were able to generate a 3D model–STL file of the complete dermal vascular system of an adult New Zealand white rabbit, which consisted of 3,301 images, each 4,000 x 4,000 pixels in size, and altogether requiring more than 3 TB of memory. We have implemented parallel I/O [input/output] to handle large datasets to generate parallel STL files from CT-scan datasets. We are also working on visualizing and creating STL files for new medical datasets that include human organs."

Rekepalli said that progress in the research so far has shown that the necessary microscopic detail can be accurately captured, and work continues in the development of the new workflow, automated pipeline, parallel visualization, and web-enabled science gateways for implementation of Mondy's strategy for tissue biofabrication.



Corrosion cast of a human heart. Courtesy of: Institute for Plastination, Heidelberg, Germany; Adiana Research and Development; and William L. Mondy.

Building a Better Carbon Trap

In May 2013, a U.S. government lab in Hawaii measured a carbon dioxide concentration of over 400 parts per million. The last time carbon dioxide was that high, *Australopithecus afarensis* walked a world that was more than 2 million years away from seeing its first modern human.

Human generation of carbon dioxide and other greenhouse gasses is altering the global climate. No one can say yet what technologies we'll need to solve this dilemma. But capturing carbon dioxide from smokestacks and other waste streams is attractive because it would allow us to continue using fossil fuels, at least for a while.

To this end, a team led by Patrick Nugent of the University of South Florida's Department of Chemistry and Youssef Belmabkhout of the King Abdullah University of Science and Technology's Advanced Membranes & Porous Materials Center undertook a series of lab experiments and computer simulations to create a better carbon trap—a workable "carbon capture material"—out of a series of materials called metal organic frameworks (MOFs).

DETAILS COUNT: HARNESSING MEMORY TO A SMALL-SCALE MODEL

Brian Space, professor of chemistry at USF, leads the molecular simulation group that performed the simulations. First, they produced precise computer models of individual gas molecules and tested how they interacted in pairs and triplets. Moving from simulating pairs of molecules to triplets—which had never been done before—made a big difference in the computers' predictions. The group performed these initial simulations on PSC's Blacklight, because of that machine's ability to handle computations with extremely large memory demands.

In the three-molecule simulations, they discovered, electrical charges were distorting the gas molecules, attracting them much like a magnet attracts iron filings. This changed the molecules' behavior significantly compared with the earlier, two-molecule simulations.

SCALING UP: TAKING ADVANTAGE OF RAW PARALLEL POWER

The next step—using the results from the detailed models to scale up to a bulk simulation—required a different type of computing.

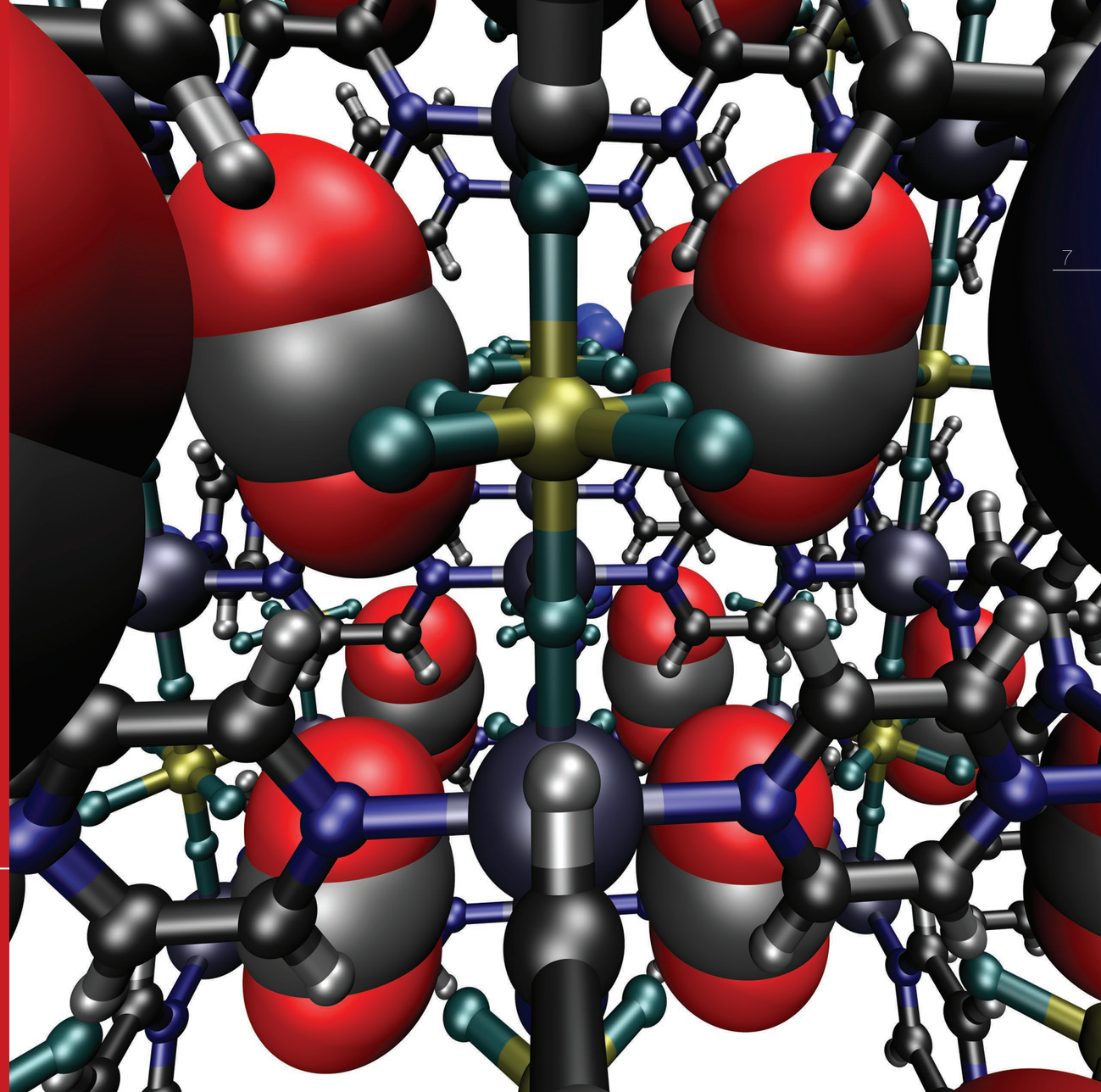
"For the different pieces of the puzzle we needed different machines," Space explained. "After our group develops an in-depth model of the forces between various guest molecules and their host materials, we perform molecular simulations of the sorption-mediated processes on highly parallel resources."

Instead of being limited by memory, the bulk simulations were limited by processing speed. For that, Space and his colleagues turned to another set of XSEDE resources: Highly parallel machines capable of splitting large problems into small, quickly solvable ones using tens of thousands of processors. The availability of such resources—SDSC's Trestles, TACC's Ranger and Georgia Tech's Keeneland—via a single XSEDE account greatly facilitated the work, Space noted.

GIVE-AND-TAKE: TOWARD A BETTER CARBON TRAP

The lab and computational researchers traded results, using the simulations to improve the candidate materials and results with those materials to improve the simulations. The end result was a family of MOF structures containing an electrically charged silicon-fluoride compound that attracts carbon dioxide much more strongly than other gases. They even work in the presence of water vapor, a common exhaust component that prevented efficient carbon capture in earlier materials. This research was published in February 2013 in the journal Nature.

"The new materials show very selective carbon dioxide capture," Space said. "Indeed, carbon dioxide fits [into the MOFs] like a glove; nothing else fits as well."



Carbon dioxide molecules (gray and red) trapped in the MOF matrix (blue). Courtesy of: Katherine Forrest, Tony Pham, Christian Cioce, and Brian Space.

From the Kitchen to Chemotherapy

Food science researchers are exploring ideas for using natural starch and protein molecules as delivery vehicles with a variety of potential uses, from making more tasty gluten-free bread to foods with enhanced health benefits. Outside the kitchen, the technique also might be used to make chemotherapy easier on patients.

Molecular dynamics simulations on multiple XSEDE computational resources, and assistance from the XSEDE Extended Collaborative Support Service (ECSS) team, help make the research possible.

The simulations are rooted in a serendipitous experiment that happened to use old flour in examining how starch, fatty acid and soluble protein molecules self-assemble into nanoparticles with unique functionality.

An unexpected observation led scientists at the Whistler Center for Carbohydrate Research at Purdue University to conclude that a type of large starch molecule with a helical structure was not only trapping smaller hydrophobic molecules, a well-known property of the starch, but also rendering the molecules soluble in the presence of a soluble protein. Water-averse hydrophobic molecules don't normally dissolve well in water.

The finding suggests so-called "nutraceutical" uses, for example loading starch molecules with small hydrophobic molecules that have nutritional health benefits, such as antioxidants. The loaded starch could be incorporated into new fortified foods, said Professor Bruce Hamaker, a Purdue food scientist and director of the Whistler Center for Carbohydrate Research.

In a similar vein, Hamaker and collaborator Osvaldo Campanella, a Purdue agricultural and biological engineering professor, also are examining potential pharmaceutical uses, such as delivery vessels for the small hydrophobic molecules common to popular chemotherapy treatments. Because the molecules resist solubility, large doses of the chemicals are necessary to ensure a high enough concentration to kill tumor cells. Delivered in a soluble form via a nanoparticle formed by self-assembling safe biomolecules, it's possible the dose could be reduced and targeted more precisely, mitigating the side effects chemotherapy patients suffer.

"There's a whole group of small hydrophobic molecules that have value to the body as nutrients, and there are a lot of drugs of the right size and hydrophobic classification, that we have found we can trap inside this nanoparticle," Hamaker said.

This kind of thing can be difficult to explore experimentally, the traditional method employed by food scientists, because of the small scale and the difficulty of strictly controlling variables such as temperature, say Hamaker and Campanella, principal investigators for the research.

High-performance computing is for food scientists a relatively new way to tackle such challenges. With Gromacs, a molecular dynamics simulation program, and enough computing power, the Purdue researchers can employ a supercomputer as, in essence, a super microscope to open a window on the action at a molecular level and likewise control and manipulate conditions.

"It changes the way that I am thinking about these kinds of projects, it really does," Hamaker said.

The simulations are computationally demanding because they involve tracking movements and interactions of numerous molecular particles over significant time scales. On lab computers, simulating small molecules forming complexes with amylose, a starch component, required 10 days just to cover a 2- to 10-nanosecond period. That was not only insufficient to match experimental observations but also too short to capture key elements of the process.

Then, the researchers started working with the Rosen Center for Advanced Computing at Purdue and Feng "Kevin" Chen, a staff member for the Rosen Center's Scientific Solutions Group. Chen is Purdue's XSEDE campus champion and an XSEDE ECSS team member.

Chen helped them get time on—and their code running on—XSEDE-allocated systems Lonestar and Stampede at the Texas Advanced Computing Center, Kraken at the National Institute for Computational Sciences at the University of Tennessee and the now-retired Steele cluster at Purdue. They are now running 500-nanosecond simulations and capturing key details of the amylose binding process that would not have been revealed in simulations shorter than 200 nanoseconds. Their results have been reported in the Journal of Agricultural and Food Chemistry and the Journal of Food Engineering, among other places.

"The problems they would like to address required more resources," Chen said. "Now, they can run at a larger scale over a longer period of time and they get a lot faster feedback."

They're also coming away with new ides for applying the research. Hamaker and Campanella are beginning a project, in collaboration with researchers in Purdue's Botany and Plant Pathology Department, to look at carrying fungicides and herbicides for agricultural use, perhaps, as in the chemotherapy example, allowing lower doses to be more effective and lessening impacts on the environment.

In addition, the researchers are looking at configuring a non-wheat cereal protein to act more like wheat gluten, which might be applied to give gluten-free bread more of the elastic, springy quality of traditional wheat-based bread.

Sculpting Flow

Have you ever noticed the way water flows around boulders in a fast-moving river, creating areas of stillness and intense motion? What if those forces of fluid flow could be controlled at the smallest levels?

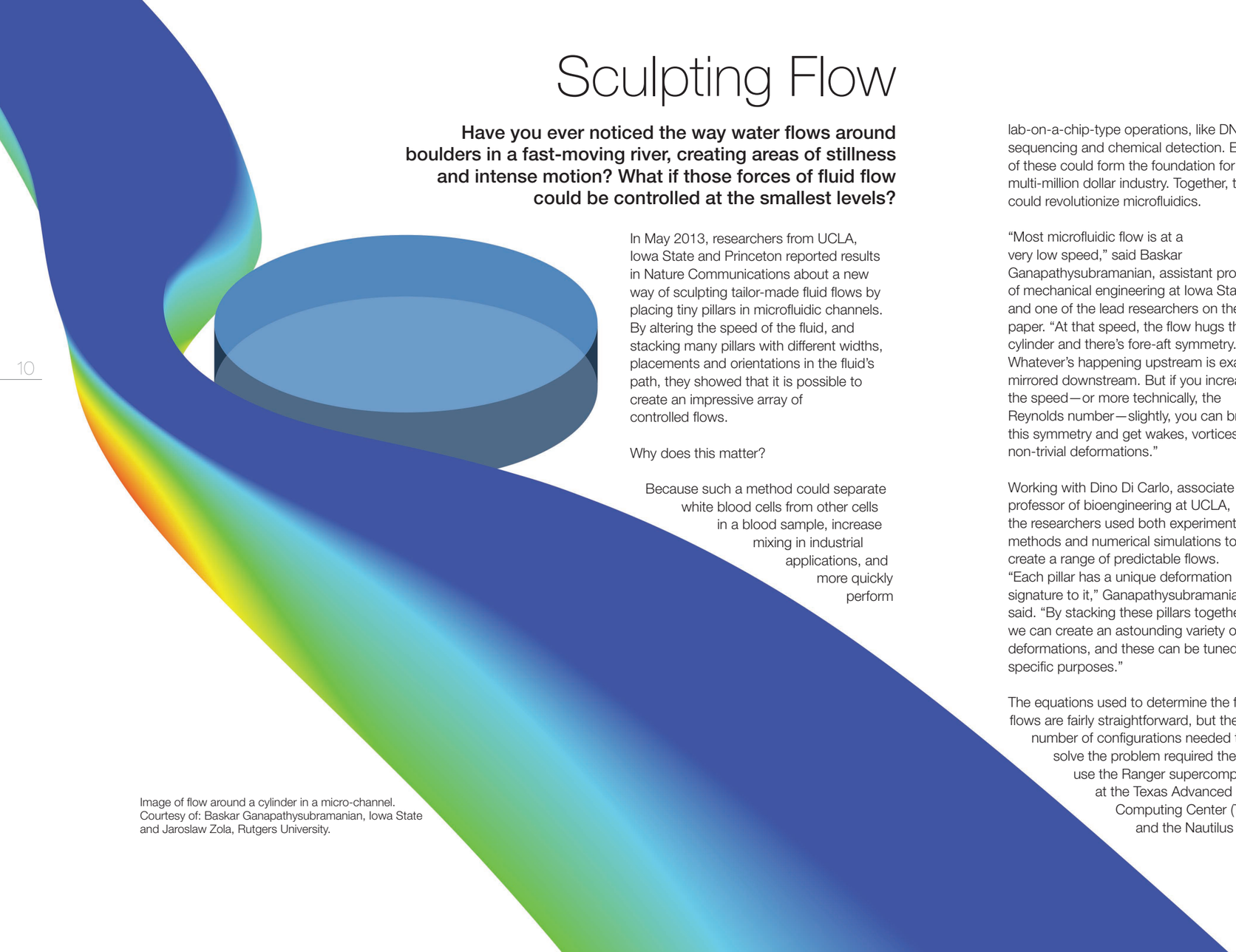


Image of flow around a cylinder in a micro-channel. Courtesy of: Baskar Ganapathysubramanian, Iowa State and Jaroslaw Zola, Rutgers University.

In May 2013, researchers from UCLA, Iowa State and Princeton reported results in Nature Communications about a new way of sculpting tailor-made fluid flows by placing tiny pillars in microfluidic channels. By altering the speed of the fluid, and stacking many pillars with different widths, placements and orientations in the fluid's path, they showed that it is possible to create an impressive array of controlled flows.

Why does this matter?

Because such a method could separate white blood cells from other cells in a blood sample, increase mixing in industrial applications, and more quickly perform lab-on-a-chip-type operations, like DNA sequencing and chemical detection. Each of these could form the foundation for a multi-million dollar industry. Together, they could revolutionize microfluidics.

"Most microfluidic flow is at a very low speed," said Baskar Ganapathysubramanian, assistant professor of mechanical engineering at Iowa State and one of the lead researchers on the paper. "At that speed, the flow hugs the cylinder and there's fore-aft symmetry. Whatever's happening upstream is exactly mirrored downstream. But if you increase the speed—or more technically, the Reynolds number—slightly, you can break this symmetry and get wakes, vortices and non-trivial deformations."

Working with Dino Di Carlo, associate professor of bioengineering at UCLA, the researchers used both experimental methods and numerical simulations to create a range of predictable flows. "Each pillar has a unique deformation signature to it," Ganapathysubramanian said. "By stacking these pillars together, we can create an astounding variety of deformations, and these can be tuned for specific purposes."

The equations used to determine the fluid flows are fairly straightforward, but the number of configurations needed to solve the problem required them to use the Ranger supercomputer at the Texas Advanced Computing Center (TACC) and the Nautilus and Kraken systems at the National Institute for Computational Science (NICS), all of which are XSEDE-allocated resources.

Using several thousand processors concurrently, the researchers ran more than 1,000 problems, each representing a combination of speeds, thicknesses, heights or offsets. "Each of these gives us one transformation and together, they form what we call a library of transformations," he described.

With this method, Ganapathysubramanian said it's possible to create a sequence of pillars that would push white cells to the boundaries of a channel and then return them to the center to be recaptured to aid in blood analysis, and another that would remove heat when fabricating microprocessors.

"Engineering tools like this allow scientists to easily develop and manipulate a flow to a shape of their interest," Di Carlo said. "There hasn't been that platform available in the fluids community."

Recently, partnering with Manish Parashar, director of the Rutgers Discovery Informatics Institute, and with Rutgers University research professor Jaroslaw Zola, Ganapathysubramanian undertook another experiment.

Using a framework called CometCloud that allows researchers to combine diverse computing resources, the team harnessed 10 supercomputers at six high-performance computing centers across three continents to extend Ganapathysubramanian's microfluidics simulations. The consortium included TACC's Stampede system, as well as resources from the Department of Energy, FutureGrid, and international HPC centers.

All told, the researchers ran 12,845 flow simulations, consuming more than 2.5 million core-hours and generating 400 gigabytes of data over the course of 16 days.

"The experiment allowed us to explore an alternate paradigm for doing computational science and demonstrate that we can support applications using this paradigm," Parashar said. "Many applications have a similar workflow so this could be a model for supporting researchers without all of them going to one resource or another."

The computations enabled by XSEDE and CometCloud brought Ganapathysubramanian halfway to his dream of a complete library of microscopic fluid flows. However, completing the entire library will require much more computing power. Fortunately, supercomputers are getting relentlessly faster, and with new technologies come new opportunities for industry, science and medicine.

"We're doing lots of interesting work on this project that people are very excited about," he concluded, "and everything is predicated on HPC resources like those offered by XSEDE."

XSEDE Reaches Out

“In other organizations, I have been much more of a starter and trendsetter because they really didn’t have much for Native American students before me. I hadn’t even really seen many, if any, Native American students at their past events and meetings either.”

Shanadeen Begay, PhD candidate in computational chemistry at Boston University and XSEDE Scholar, spoke about her experiences as a Navajo student in the world of computational science. “I actually took a break from the other organizations because I needed to not feel like the token.”

XSEDE reaches out to people who are under-represented in the fields of science, engineering, and high-performance computing--including women, minorities, and people with disabilities. Through various events and programs, XSEDE aims to increase their involvement with computational science.

“It’s nice to see an organization that is already [doing minority outreach] and succeeding,” said Begay.

Begay first heard of XSEDE and the project’s minority outreach efforts while presenting a student poster session at the ACM Richard Tapia Celebration of Diversity in Computing. There she met Ruth Kravetz, associate director of the Richard Tapia Center for Excellence and Equity and program manager for the XSEDE Scholars Program and ELA Program. According to Richard Tapia—a Hispanic mathematician and professor in the Department of Computational and Applied Mathematics at Rice University in Houston, Texas and champion of under-represented minorities in the sciences—“Ruth Kravetz is actually the main person who makes things happen; she gets into the trenches.”

That was certainly the case in the meeting of Begay and Kravetz back at the Tapia Conference in November 2012. Since then Begay has joined the XSEDE Scholars program, attended another Tapia conference, competed in the XSEDE13 Student Programing Challenge with fellow scholars, and mentored many minority students along the way. And she has no plans to stop.

The XSEDE Scholars program is for U.S. students from underrepresented groups in the area of computational sciences, providing opportunities to learn more about high-performance computing and XSEDE resources as well as to network with cutting-edge researchers and professional leaders.

“She is a strong, intelligent, involved, a most excellent human being who sees herself as PhD faculty leading her field down the road,” Kravetz said, recalling some of her first impressions of Begay. “The focus of the XSEDE Scholars program is to get driven students like Shanadeen, particularly from the African American, Hispanic, female and Native American areas of computational science, to form a cohort of student peers to establish a community of future academic leaders.”

This is all a part of Richard Tapia’s vision for minorities in science and computational fields. His philosophy for success is to get minority students from majority schools—Illinois, Berkeley, Stanford, Rice, Georgia Tech, etc.—involved with research activities.

Under the leadership of Tapia, XSEDE has implemented the Minority Faculty Council (MFC). This team of minority faculty members—researchers and leaders in their respective fields—looks to increase underrepresented minority leadership, mentoring and participation.

Tapia believes XSEDE respects what it means to bring minorities, both students and faculty, into these positions along with understanding of what it will mean for our future as a community. “Getting them internships, getting them working with researchers, letting them see what it is like helps them begin feeling comfortable in that environment,” he said, “and then they start saying ‘I belong here.’”

Mariela Barrerra is a Latino undergraduate studying computer science at West Texas A&M University. Shortly after being accepted into the University, Barrerra learned about XSEDE scholars and is now going into her second year with the program. She credits much of her academic success with her participation in XSEDE.

“I’m from north Texas, small community, and you don’t hear about this stuff,” said Barrerra. At West Texas, she is the only female computer science student and one of three Hispanic students out of a total of 80 within the department. “So being in that environment, I don’t get to see people who look like me,” Barrerra continued, “We all have different talents and views that we can bring to the table, and I think that is really important to share, and that’s why the whole XSEDE minority outreach is great.”

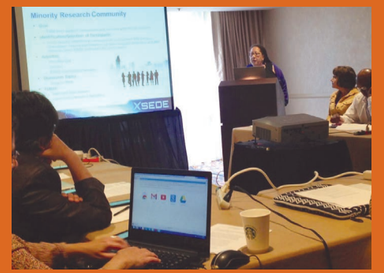
The MFC hopes to implement plans and tactics that will be used broadly. This includes focusing on undergrads; building a community feel; placing more responsibility on faculty members; encouraging post-doctoral fellowships to enhance specialization; integrating this effort into mainstream efforts; and making sure “underrepresentation” doesn’t just refer to “international” students.

With programs like XSEDE Scholars and MFC, XSEDE is set to become a leader in garnering participation from underrepresented minorities and steering the way to a more inclusive, welcoming environment that breeds success for all.

XSEDE Scholars and members of the Minority Faculty Council gathered at XSEDE13.

Mariela Barrerra (back, middle), a Latino undergraduate studying computer science at West Texas A&M University and XSEDE Scholar, participating in peer-to-peer menetoring at XSEDE13. Barrerra said this is one of the more important things an XSEDE Scholar can do.

From left to right: Mariela Barrerra, Richard Tapia, Ruth Kravetz, Shanadeen Begay during a XSEDE Scholars dinner.



Linda Akli, Assistant Director, Training Education & Outreach for SURA, gives a presentation about outreach to fellow members of the Minority Faculty Council.

Shanadeen Begay (middle, right), PhD candidate in computational chemistry at Boston University, member of the Navajo Nation and XSEDE Scholar, as she helps bring her team to victory in the XSEDE13 Student Programming Competition.

Make Mine a Double

Everyone has vertebral spinal endplates. These thin layers of bone are essential to the health of intervertebral discs, the tissues that link the backbones together and can cause major damage if they degenerate or slip out of place.

Recently, researchers at The University of California, San Francisco (UCSF) discovered and confirmed the existence of a tissue that has been overlooked, unnoticed, or inadvertently sliced up in the dissection or sampling process: double-layer vertebral endplates. The findings appeared in the journal Spine in October 2012.

“This feature has a mystical quality because there are only a handful of studies that have reported seeing it; they have reported it only anecdotally since its structure and function are difficult to study with traditional sectioning techniques,” said Aaron Fields, a post-doctoral researcher in the Laboratory for Orthopaedic Bioengineering at UCSF.

Fields and his colleagues discovered the second layer of bone after harvesting small cores from six cadaver spines and scanning them with a micro-CT scanner. The high-resolution scans allowed for exceptionally thin images of the endplate cores to be collected and reconstructed into detailed digital renderings that brought the structure of the double-layer endplate into view for the first time.

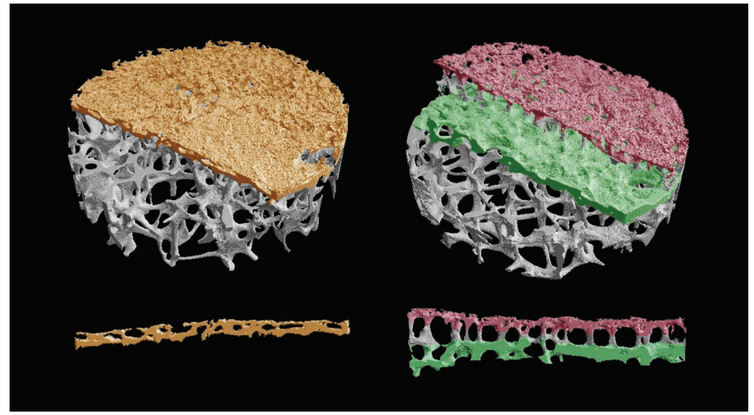
“We were looking to see how different endplate structures affect disc health and we were intrigued by the potential for double-layer endplates to more effectively balance conflicting endplate functions,” Fields said.

To promote disc health, a spinal endplate must balance two opposing functions: distributing weight and facilitating nutrient and waste transport. According to Fields, having a redundant layer of bone means that when there’s damage in the first layer, some of the load can be redistributed to the second layer and proper function is maintained, much in the same way that structural redundancies in a bridge might permit safe function when a part of the structure fails.

Analyzing the high-resolution models required some serious hardware and software. The problem was too big for his local computing cluster, so Fields turned to XSEDE. To accomplish the task, the researchers accessed the processing power of the Ranger supercomputer and the Spur visualization system located at the Texas Advanced Computing Center. In no time, the California-based researchers made their models come to life while using resources halfway across the country.

“These large finite element models contain hundreds of millions of degrees of freedom so they’re far too large for conventional computer systems to handle,” Fields said. “Without the resources at TACC available through XSEDE, this research wouldn’t have been possible.”

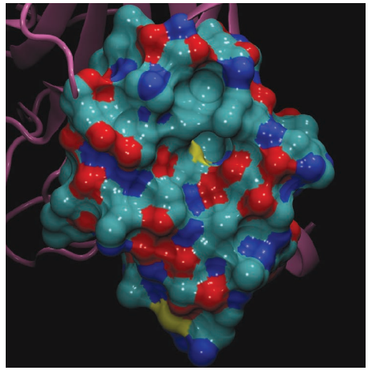
The discovery of the double endplate spinal structure could have an impact on personalized spinal treatments of the future. “As we’re exploring personalized treatments for degenerative disc disease, the double endplate could be a factor that makes one particular treatment more effective than another,” Fields said.

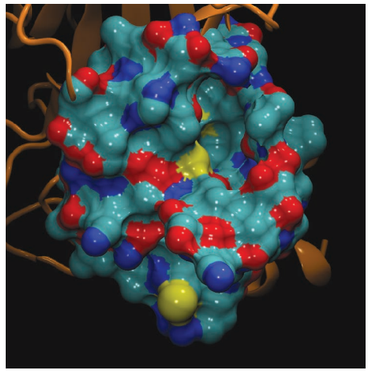


A double layer endplate is made up of a porous, thin top layer (red) and thicker second layer (green). The double and single layer endplates (orange) examined had a very similar overall thickness. Courtesy of: The University of California, San Francisco.

Pocket Full of Promise

Called “the guardian of the genome,” tumor protein 53 (p53) regulates the body’s cell cycle, suppresses tumors, and plays a role in preventing genetic mutations. However, in many cases, p53 itself is mutated, losing its tumor-suppressing function and allowing cancer to proliferate.





Top: crystal structure tumor protein 53 (p53). Bottom: computationally-generated structure of p53 with open pocket. Courtesy of: Rommie Amaro and Özlem Demir, University of California, San Diego.

“The mutant p53 gene is found in almost 40 percent of the most commonly diagnosed cancers today,” said Rommie Amaro, a professor of chemistry and biochemistry at the University of California, San Diego, who studies the protein.

The biology of cancer is complex and the development of effective cures lies in creating more effective, targeted therapies. Amaro and her team of University of California-Irvine biologists, chemists and computer scientists might have found such an approach. In a January 2013 article in Nature Communications, they described the discovery of a promising pocket in the structure of the p53 molecule.

“The discovery of this pocket might ensure that drugs can reactivate the mutant p53 function,” Amaro said. “This is an exciting possibility.”

The discovery was enabled by simulations performed on supercomputing resources at the Texas Advanced Computing Center (Ranger) and the San Diego Supercomputing Center (Trestles). The finding could herald new and better-targeted therapies for cancer treatment.

Proteins are dynamic molecular machines, moving, flexing, and changing shape to accommodate small molecules or other proteins. Researchers are not able to resolve all the structures a particular molecule may adopt in a cell using experimental methods alone.

XSEDE-allocated supercomputers allowed Amaro and her team to simulate many slightly different structures of p53, instead of looking at a single, fixed structure. Starting with a static picture of the enzyme, they used molecular dynamics simulations to add flexibility and motion to the proteins based on the physics of the molecules and the environment in which they exist. The discovery of the pocket furthers the scientists’ understanding of p53’s structure and behavior.

“In a number of drug targets, there is great flexibility in the site where the drugs are binding,” said Heather Carlson, a professor of medicinal chemistry at the University of Michigan. “With the p53 molecules, the real site where drugs can bind wasn’t visible or exposed in the crystal structures.”

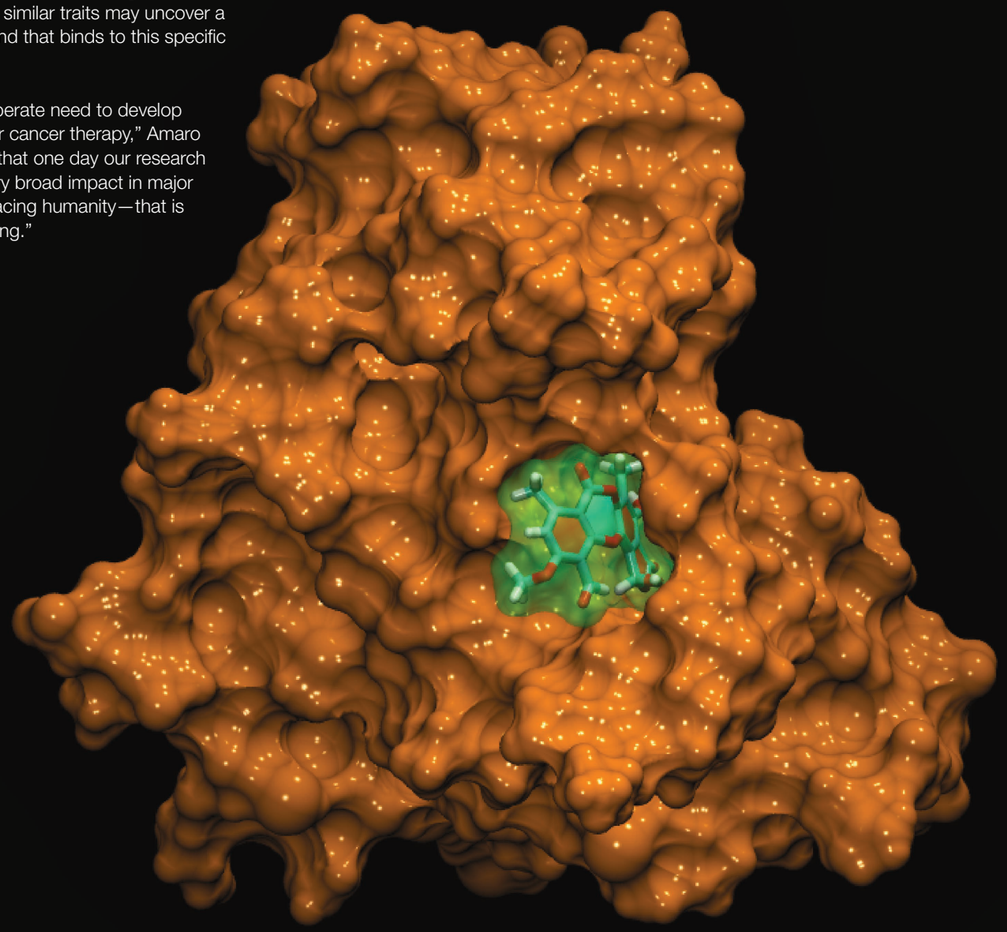
It was computer simulations carried out by Özlem Demir, a post-doctoral researcher in the Amaro lab, that changed the team’s structural understanding of p53 and revealed what wasn’t found in the experimental studies.

“We were able to use the newly-exposed pocket to show where these agents can bind,” Amaro said.

After the discovery of the mutant p53 pocket, the team performed virtual drug screenings to search for a drug molecule that could fit snugly in the protein pocket. Stictic acid, a product found in some species of lichens, emerged as a potential p53 reactivation compound. It docked in the pocket for 60 nanoseconds, and, as the molecule moved deeper into the pocket, it achieved better shape complementarity. In other words, it fit.

Study co-leader Peter Kaiser noted that while stictic acid cannot be developed into a viable drug, the work suggests that a comprehensive screening of small molecules with similar traits may uncover a viable compound that binds to this specific p53 pocket.

“There’s a desperate need to develop better drugs for cancer therapy,” Amaro said. “To think that one day our research may have a very broad impact in major health issues facing humanity—that is hugely motivating.”



Stictic acid (highlighted in green) docked in the open p53 pocket.

New Code Shakes Up Seismic Modeling

A team of seismologists, leveraging the massive computational power available within the XSEDE network, has developed a highly scalable computer code that will allow researchers to better predict earthquake activity, and in turn increase the potential to save lives and minimize property damage when a major temblor strikes.

The breakthrough—using accelerated GPU code in place of CPUs to allow a sustained two petaflop/s performance—also means a dramatic reduction in both research time and energy costs in simulating seismic hazards.

The team, led by Yifeng Cui, a computational scientist at the San Diego Supercomputer Center (SDSC) at the University of California, San Diego, developed the accelerated code for use in earthquake engineering and disaster management through regional earthquake simulations at the petascale level as part of a larger computational effort coordinated by the Southern California Earthquake Center (SCEC).

"The increased capability of GPUs, combined with the high-level GPU programming language CUDA, gives us the tremendous power required for acceleration of numerically intensive 3D simulation of earthquake ground motions," according to Cui.

The accelerated code is based on a widely-used wave propagation code called AWP-ODC, which stands for Anelastic Wave Propagation by Olsen, Day and Cui. It was named after Kim Olsen and Steven Day, geological science professors at San Diego State University, and SDSC's Cui. The research team restructured the code to exploit high performance and throughput, memory locality, and overlapping of computation and communication, which made it possible to scale the code linearly to more than 8,000 NVIDIA Kepler GPU accelerators.

Using the Keeneland system managed by Georgia Tech, Oak Ridge National Laboratory (ORNL), and the National Institute for Computational Sciences (NICS), the team performed GPU-based benchmark simulations of the 5.4 magnitude earthquake that occurred in July 2008 below Chino Hills, near Los Angeles. Researchers were able to achieve a five-fold speedup over a heavily optimized CPU code and a sustained performance of two petaflops per second (or two quadrillion calculations per second). A previous seismic simulation milestone performance peaked in 2010 at only 200 teraflops (trillions of calculations per second) sustained.

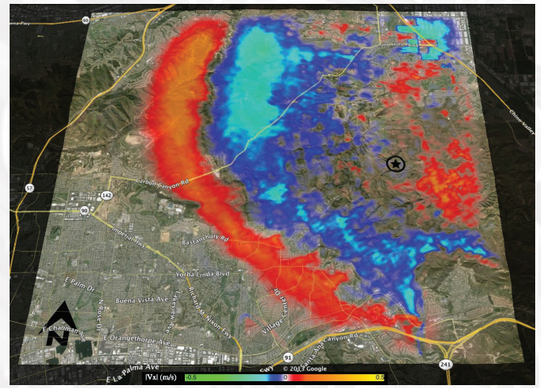
"This is an impressive achievement that has made petascale-level computing a reality for us, opening up some new and really interesting possibilities for earthquake research," said Thomas Jordan, SCEC's director.

NEXT STEPS

While the GPU-based AWP-ODC code is already in research use, further enhancements are being planned. "One goal going forward is to use this code to calculate an improved probabilistic seismic hazard forecast for the earthquake-prone California region," said Cui. "Our ultimate goal is to support development of a CyberShake model that can assimilate information during earthquake cascades so we can improve our operational forecasting and early warning systems."

CyberShake is a SCEC project focused on developing new approaches to performing seismic hazard analyses using 3D waveform modeling. The GPU-based code has potential to save hundreds of millions of CPU-hours required to complete statewide seismic hazard map calculations in planning.

Funding for this research was in part provided through XSEDE's Extended Collaborative Support Service (ECSS) program. "ECSS exists for exactly this reason—to help research teams make significant performance gains and take their simulations to the next level," said Nancy Wilkins-Diehr, co-director of the ECSS program and SDSC's associate director. "ECSS projects are typically conducted over several months to up to one year. This type of targeted support may be requested by anyone through the XSEDE allocations process."



A snapshot of simulated horizontal (east-west) ground motion of 2008's magnitude-5.4 Chino Hills, CA, earthquake. The red-yellow and green-blue colors depict the large and small amplitude of shaking; the star shows the earthquake's location. Courtesy of: Efecan Poyraz/UC San Diego, Kim Olsen/SDSU. Map image courtesy of: Google.

Experiencing Some Turbulence

While most people think of turbulence as the source of unsettling bouts of chaotic airflow during a flight on an airplane, physicists have a much deeper concept of its significance in the world around us. In fact, Nobel Laureate and theoretical physicist Richard Feynman once said, "Turbulence is the most important unsolved problem of classical physics."

Probing the puzzling nature of turbulence is essential because of its impact on the flow physics of liquids and gases, and, by extension, the influence of those fundamental states of matter streaming outside and inside things. For example, turbulence must be considered in designing vehicles and in understanding how particles —such as pollution and volcanic ashes— disperse in the atmosphere. Turbulence also affects the flow inside a jet engine, a combustor or a nuclear reactor.

To convey the essence of the ubiquitous influence of turbulence, researcher Antonino Ferrante of the William E. Boeing Department of Aeronautics and Astronautics of the University of Washington, Seattle, quotes Greek philosopher Heraclitus: "Everything flows and nothing abides; everything gives way and nothing stays fixed."

"As I sit and look around me, I notice several examples of turbulent flows: the smoke flowing out of a chimney, the wind moving between the leaves and branches of trees, massive clouds moving in the atmosphere, the air surrounding a flying bird and an airplane," Ferrante said. "No matter how big or small, or how close or far you look, you will see fluids in motion. As the ratio of the inertial forces to the viscous forces in the flow (that is, what's known as the Reynolds number) increases above a certain threshold, the flow transitions from laminar to turbulent, or from smooth to random. Most flows in nature and engineering applications are turbulent. Thus, understanding turbulent flows is very important for human progress and for a sustainable planet Earth."

XSEDE HPC RESOURCES: ESSENTIAL TO PROJECT SUCCESS

Ferrante is principal investigator for the Computational Fluid Mechanics group, a research team engaged in the study of turbulence modeling and simulations, one of the most challenging areas of fluid dynamics —the natural science of fluids, both gases and liquids, in motion. Turbulence simulations are particularly difficult because of the wide-range scales of motion involved. Resolving all of those scales requires fine computational grids with billions of points that include the tiniest of scales, where viscous dissipation—the heating of the fluid due to resistance—occurs.

The goals of the team's project are to simulate fluid flow using the largest Reynolds number ever reported, perform the first direct numerical simulation (DNS) of fully resolved droplet-laden isotropic (without a preferential direction) turbulence and share the generated DNS database with the scientific community. Accomplishing their objectives required the team to use the XSEDE-allocated Kraken supercomputer at the National Institute for Computational Sciences (NICS).

The researchers also availed themselves of the XSEDE Extended Collaborative Support Services program, through which they consulted with staff at NICS and NCSA.

The assistance of NICS enabled Ferrante's team to simulate turbulent flows as well as turbulence coupled with other phenomena—including chemically reactive and multiphase (gas and liquid combined) turbulent flows—and get results in a reasonable time.

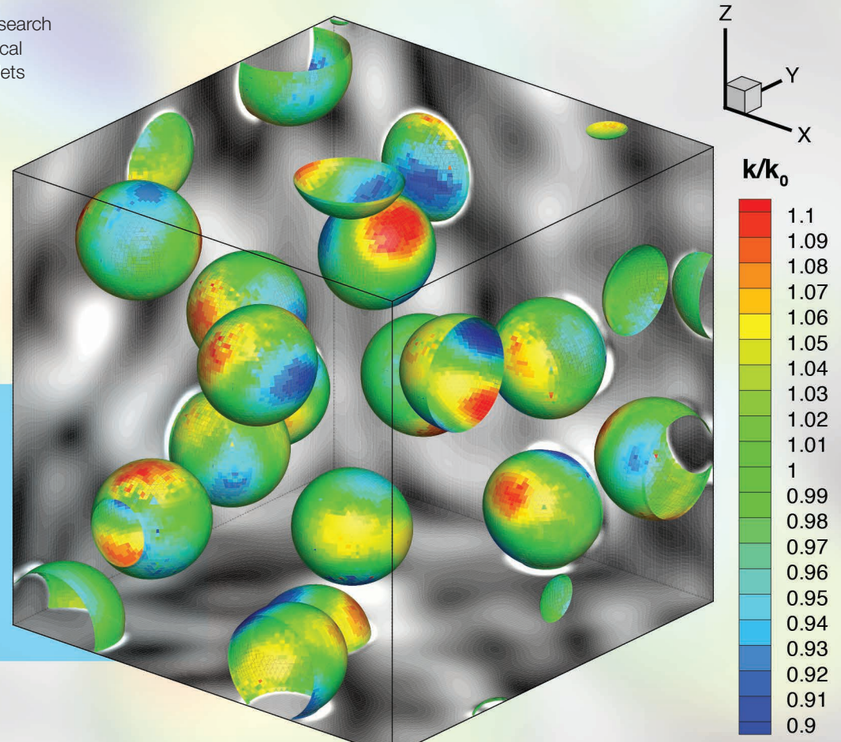
NCSA provided petascale scaling and deployment, as well as the development of a high-level HDF5 (Hierarchical Data Format Release 5) parallel input and output library (called H5DNS). HDF5 is the name of a set of file formats and libraries designed to store and organize large amounts of numerical data. NCSA also contributed custom visualization of simulation results to the project.

DNS modeling, Ferrante explains, is particularly HPC dependent because it requires that several flow variables, such as velocity and pressure, be advanced in time in billions of grid points. The problem the researchers were investigating required a minimum resolution of 10,243 data points.

"Using Kraken, we were able to solve the problem involving billions of unknowns accurately," Ferrante said. "For the first time, we are simulating the effects of droplets on turbulence by performing fully resolved DNS. In practice, we are running a real-life experiment on a supercomputer rather than a wind tunnel."

AHEAD: LARGER-SCALE SIMULATIONS AND MORE-COMPLEX FLOWS

Ferrante said the outcome of the research was an accurate and robust numerical methodology that couples the droplets and the turbulent flow to better understand droplet-laden turbulent flows, and that research in turbulence simulations will continue on larger scales and solve flows of greater complexity. He presented a talk on the results of the project in May 2013 to the International Conference on Multiphase Flow in Jeju, Korea.



Simulation of freely moving, initially spherical droplets in a turbulent flow. The droplets and the turbulent flow interact, causing the droplets to deform. The color contours represent droplet curvature, which is directly proportional to the droplet surface tension. The black and white contours show enstrophy, a measure of the rotation rate of the flow. Courtesy of: Antonino Ferrante, University of Washington.

Experts Found Across XSEDE

XSEDE's scientific community and technological resources are supported by an important group of staff called the Extended Collaborative Support Service (ECSS). See how ECSS helps accelerate scientific discovery.

DES MOINES, IOWA

Klaus Bartschat and his research team faced a quandary. They wanted to know how a process called the two-photon double ionization of hydrogen could be used to perform superfast laser imaging.

Short laser pulses already allow physicists and chemists to watch fast molecular motion as it happens, but many atomic processes are even faster and require the shortest pulses ever created, prompting researchers to investigate new modes of laser imaging.

Unfortunately, previous calculations showed discrepancies in their description of the fundamental dynamics of the process, whereby electrons are removed from a hydrogen atom by photons.

Bartschat and team wanted to resolve these discrepancies, in part because free electron laser experiments are expensive. Understanding the basic dynamics of the process beforehand is critical when the time comes to perform the experiment in the lab.

They turned to XSEDE's ECSS, a unique program that pairs members of the XSEDE user community with expert staff to solve challenging science and engineering problems. Expertise is available in a wide range of areas from performance analysis to petascale optimization to the development of community gateways.

With the help of Lars Koesterke, an ECSS staff member from the Texas Advanced Computing Center, the researchers improved their code for the Stampede supercomputer, making it more than 1.5 times faster. As a result, they were able to sample different photon energies and pulse durations to solve new and bigger problems.

"We would never have been able to get such beautiful results if it hadn't been for Lars' help with the optimization," Bartschat said. "A fifty percent increase is a lot when it comes to using millions of hours on a supercomputer."

KNOXVILLE, TENNESSEE

Besides being the subject of everyday conversations about diet and nutrition, proteins are responsible for many of the most important functions in the body, from building tissue to producing hormones. The more we know about proteins, the better.

One way scientists learn about proteins is through genome studies, said University of Tennessee, Knoxville professor Michael Gilchrist. Gilchrist is leading a research team that's using computational methods to extract meaningful information from the genome of brewer's yeast. Specifically, they are seeking to better understand protein translation, one of the most fundamental and universal biological processes.

Gilchrist's team developed software in the R programming language to model the efficiency of a ribosome translating mRNA. R is a popular statistical analysis language used in bioinformatics to extract useful information from large amounts of genomic data.

Gilchrist consulted with the data analysis team at the Remote Data Analysis and Visualization Center (RDAV), including ECSS staff members Pragneshkumar Patel and George Ostrouchov. Working together, they parallelized the parts of the code that used R and converted the most computationally-intensive parts to C. This allowed the software to run efficiently on the Nautilus supercomputer, an XSEDE-allocated system managed by the National Institute for Computational Sciences and housed at Oak Ridge National Laboratory.

Preliminary results showed that calculations that took 24 hours on the desktop computer could be executed in less than an hour on Nautilus. Gilchrist said that while NICS provides his team with access to powerful computers, they offer another essential element—researchers who can assist in the use of these systems. "Being able to get the results in hours instead of weeks greatly increases our ability to develop our methods and test alternative hypotheses," Gilchrist said.

SAN DIEGO, CALIFORNIA

Soon, scientists hope to place chemical sensors (also known as e-noses) in the homes of the elderly in order to monitor their living environments. Ramon Huerta from the University of California, San Diego serves as the lead researcher on the project. He believes sensor data can detect anomalous behavior that may suggest a change in health status. Such monitoring systems would allow seniors to stay longer in their own homes.

In preparation for a pilot study, ECSS staffer Bob Sinkovits, a researcher at the San Diego Supercomputer Center, adapted the software developed by the Huerta lab so it could be used on the Gordon supercomputer, which is specifically designed for data-intensive applications. After optimizing their code, calculations that had taken 15.5 hours on a local workstation can now be completed in four minutes on a single Gordon compute node.

"ECSS staff think obsessively about compilers and instruction set architectures, but most of our users do not," Sinkovits said. "We want them to focus on their science impact."

Huerta's project no longer needs supercomputers to analyze a single data set, but the new and improved code now makes it practical to simultaneously analyze large numbers of data sets. Sinkovits said that this is exactly how he wants the ECSS program to work: "Tasks that were previously hard become routine and the researchers can turn their attention to problems that had been considered impossible."

Learn more about the ECSS program and ECSS projects: https://www.xsede.org/ecss.

Crunching Collider Data

One of XSEDE's newest supercomputers recently completed one of its most data-intensive tasks to-date: rapidly processing raw data from almost one billion particle collisions as part of a project to help define the future research agenda for the Large Hadron Collider.

As part of a collaboration with the Open Science Grid, researchers at the University of California, San Diego turned to the Gordon supercomputer launched last year by the San Diego Supercomputer Center (SDSC) to process massive data sets generated by the Compact Muon Solenoid, or CMS, one of two large general-purpose particle detectors at the Large Hadron Collider used by researchers to find the elusive Higgs particle.

The around-the-clock data processing run was completed in about four weeks' time, making the data available for analysis several months ahead of schedule. About 1.7 million core hours – or about 15% of Gordon's total compute capacity – were dedicated to the task, with more than 125 terabytes of data streaming through Gordon's nodes and into SDSC's Data Oasis storage system for further analysis.

"With only a few weeks' notice, and in part thanks this data-intensive computing and storage system that is part of the XSEDE network, we were able to gain access to Gordon and complete the runs, making the data available for analysis in time to provide crucial input toward international planning meetings on the future of particle physics," said Frank Wuerthwein, a professor of physics at UC San Diego and member of the CMS project.

"Giving us access to Gordon effectively doubled the data processing compute power available to us," added Lothar Bauerdick, OSG's executive director and the U.S. software and computing manager for the CMS project. "This gave CMS scientists precious months to get to their science analysis of the data reconstructed at SDSC."

The collaboration came as the Large Hadron Collider was shut down last February to make numerous upgrades during the next two years. One major activity during the shutdown includes the development of plans for efficient, effective searches once the collider is back in operation. To do that – and to have time enough to upgrade equipment – researchers must also sift through massive amounts of stockpiled data to help define future research agendas.

"Unfortunately, the shutdown schedule meant that the parked data would not be available for analysis this summer, and possibly not even for deriving meaningful contributions to planning documents for future upgrades of the experiment that were due this fall," explained Wuerthwein.

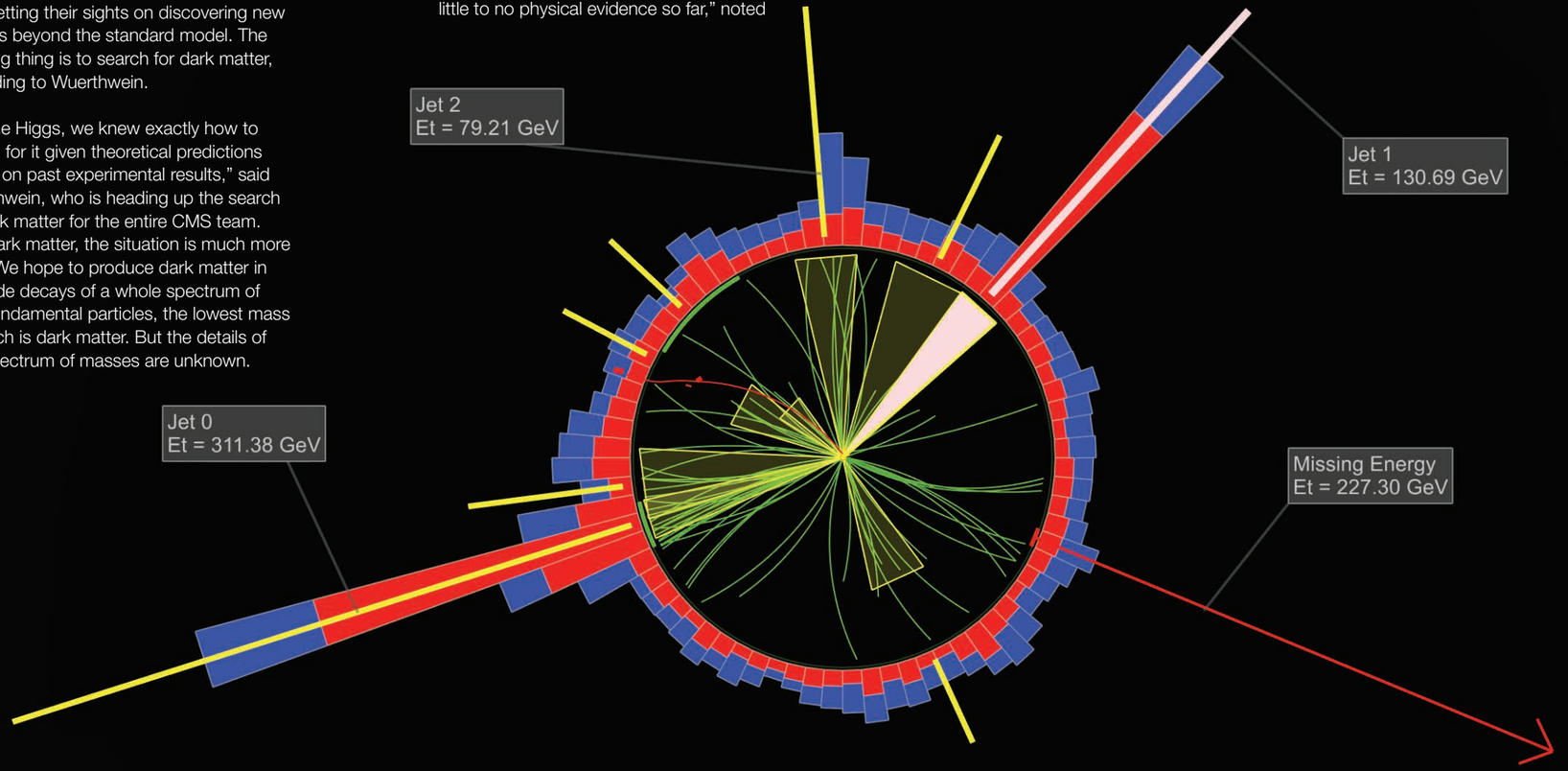
WHAT'S NEXT: THE HUNT FOR DARK MATTER

With the recent discovery and later confirmation last March of the Higgs boson – the last missing piece of the standard model of particle physics – scientists are now setting their sights on discovering new physics beyond the standard model. The next big thing is to search for dark matter, according to Wuerthwein.

"For the Higgs, we knew exactly how to search for it given theoretical predictions based on past experimental results," said Wuerthwein, who is heading up the search for dark matter for the entire CMS team. "For dark matter, the situation is much more hazy. We hope to produce dark matter in cascade decays of a whole spectrum of new fundamental particles, the lowest mass of which is dark matter. But the details of this spectrum of masses are unknown.

To have sensitivity to a larger range of possible mass spectra, we needed to write more data to tape so we can carefully analyze it."

The origin of this spectrum of new fundamental particles is a new kind of symmetry of nature called Supersymmetry, or SUSY. "Underlying this symmetry is a fascinating but theoretical conjecture with little to no physical evidence so far," noted Wuerthwein. "It's fascinating because it could provide an ordering principle that allows for all known physical forces to be unified during the earliest times of the 'Big Bang' or birth of the universe, while providing an explanation for dark matter, and resolving some of the outstanding questions about details of the Higgs mechanism and mass."



This image of a supersymmetry event shows the transverse momentum imbalance due to dark matter particles escaping the detector (direction indicated by red arrow). Red and blue rectangles indicate energy deposited in the electromagnetic and hadronic calorimeter respectively; green tracks in the center show charged particles with transverse momentum larger than 2GeV. Yellow-outlined triangles indicate jet cones or the presence of subatomic particles called quarks. Courtesy of: Matevz Tadel, University of California, San Diego/CMS.

XSEDE PARTNER INSTITUTIONS

National Center for Supercomputing Applications
(University of Illinois at Urbana-Champaign)

Cornell University Center for Advanced Computing

Indiana University

Jülich Supercomputing Centre

Michigan State University

National Center for Atmospheric Research

National Institute for Computational Sciences
(University of Tennessee-Knoxville/Oak Ridge National Laboratory)

Ohio Supercomputer Center
(The Ohio State University)

Pittsburgh Supercomputing Center
(Carnegie Mellon University/University of Pittsburgh)

Purdue University

Rice University

San Diego Supercomputer Center
(University of California, San Diego)

Shodor Education Foundation

Southeastern Universities Research Association

Texas Advanced Computing Center
(The University of Texas at Austin)

University of California, Berkeley

University of Chicago

University of Southern California

University of Virginia

XSEDE

Extreme Science and Engineering
Discovery Environment

XSEDE is supported by the
National Science Foundation.

xsede.org | info@xsede.org