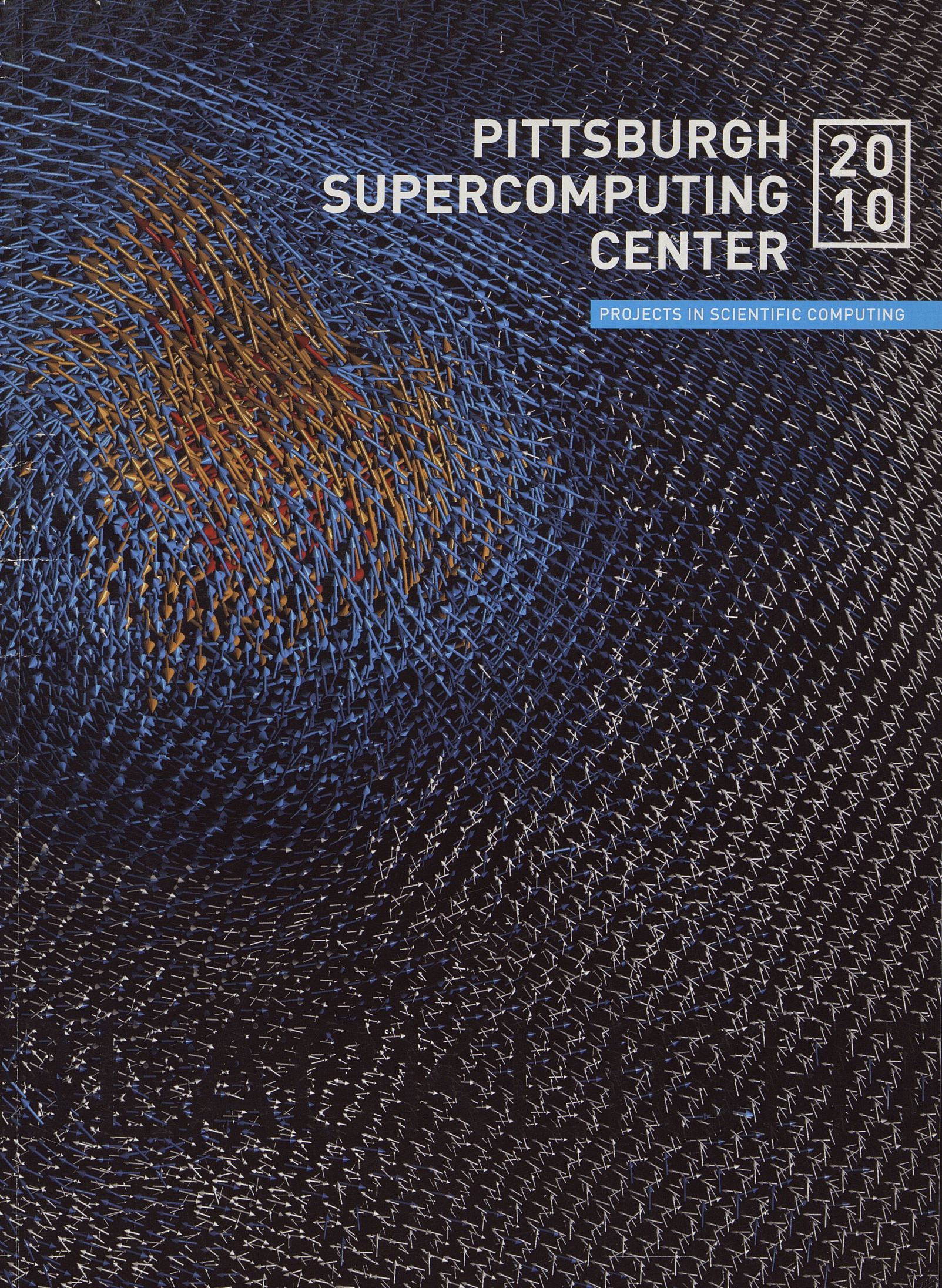
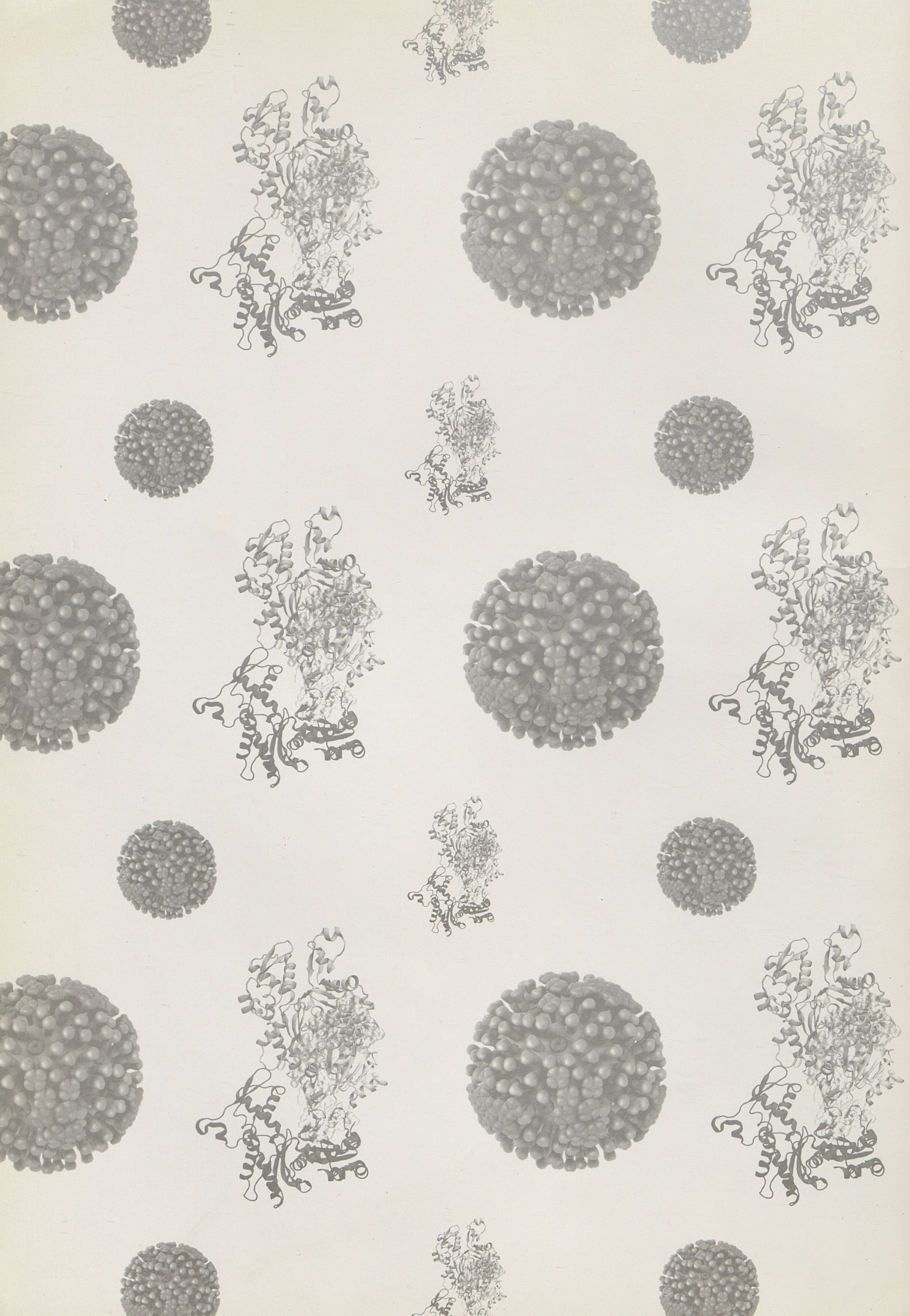
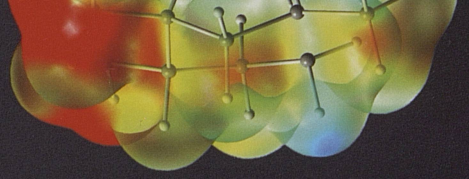


PITTSBURGH SUPERCOMPUTING CENTER

2010

PROJECTS IN SCIENTIFIC COMPUTING





PSC.EDU/10

The Pittsburgh Supercomputing Center provides university, government and industrial researchers with access to several of the most powerful systems for high-performance computing, communications and data-handling available to scientists and engineers nationwide for unclassified research. PSC advances the state-of-the-art in high-performance computing, communications and informatics and offers a flexible environment for solving the largest and most challenging problems in computational science.

As a leading partner in the TeraGrid, the National Science Foundation's program to provide a coordinated national cyberinfrastructure for education and research, PSC works with other TeraGrid partners to harness the full range of information technologies to enable discovery in U.S. science and engineering.

WWW.PSC.EDU | 412-268-4960

Foreword from the Directors

This has been a year of achievement for the Pittsburgh Supercomputing Center. As we have done many times over the past 24 years, we are making two new supercomputing systems available to the national research community. Blacklight (see p. 4) is the world's largest shared-memory computing system. Anton (p. 12) is the world's most effective system for simulation of proteins and other biomolecules. Furthermore, PSC's network group, led by Wendy Huntoon, received two major grants from the National Science Foundation (p. 14) and helped to lead a Pennsylvania coalition in securing $100 million in federal stimulus funds to create a statewide broadband infrastructure.

Blacklight's unprecedented technology will make possible advances in many fields. One of them is epidemiological modeling. Working through the University of Pittsburgh MIDAS Center of Excellence, Shawn Brown and colleagues (p. 20) showed the ability of agent-based modeling to ask and answer questions about the spread of infectious disease in a timely fashion during last year's H1N1 epidemic. With Blacklight, it now becomes feasible for the first time to do ABM modeling on a national or international scale.

Shared memory also helped with a major step forward in the innovative work (p. 24) of Tom Mitchell and colleagues at Carnegie Mellon. Their studies in machine learning made national news in 2009, and Indra Rustandi's algorithmic breakthrough — facilitated by PSC consulting, software and hardware — to enable integrated datasets has significantly improved the accuracy of their predictive model.

Atrial fibrillation doesn't have the almost immediate fatal consequences of ventricular fibrillation, but it's a serious heart condition that increases risk for stroke and heart failure. With a new approach to low-energy defibrillation (p. 28), Elizabeth Cherry and colleagues have shown the feasibility of treating atrial fibrillation much more effectively than is currently possible.

Actin is one of the most important proteins in the body, involved in metabolic pathways that when disrupted can lead to cancer. The computational work of Greg Voth and Jim Pfaendtner (p. 32), in close collaboration with their laboratory partners, shows how slight structural changes in actin can have profound implications, insights that give new understanding to experimental findings.

Terpenes are fascinating hydrocarbon compounds that have provided biomedicine with one of the most effective cancer-fighting drugs so far developed, taxol. Dean Tantillo is doing quantum chemistry studies at PSC (p. 36) that could lead to much faster, cheaper ways to produce other useful terpene-derived compounds.

Understanding the universe, how it got to be the way it is, is an ongoing project of physics, and supercomputing made a major contribution with the work of University of Washington physicists in modeling dwarf galaxies (p. 40). With improvements in their software and the availability of new powerful TeraGrid systems, they've solved one of the major problems in the reigning cold dark matter model of cosmological evolution. In astrophysics too, Blacklight's shared memory will make possible new kinds of data-analysis, as this work has demonstrated (p. 43).

We're pleased that PSC resources could enable these advances, and also that we can be a vital resource for research and education in Pennsylvania (p. 6). Along with new scientific insights, this publication also highlights (pp. 9-11) our important work to help educate the next generation of scientists and science-literate citizens.

None of this would have been possible without PSC's staff, an unparalleled assemblage of talent and experience in high-performance computing. We are also grateful for support from the National Science Foundation, the U.S. Department of Energy, the National Institutes of Health, the Commonwealth of Pennsylvania and many others.

Michael Levine Scientific Director

Ralph Roskies Scientific Director

CONTENTS

BLACKLIGHT COMES TO PSC

With a $2.8M award from the National Science Foundation, PSC introduces the world's largest coherent shared-memory system.

In July, the National Science Foundation (NSF) awarded $2.8 million to help PSC acquire a system that features Silicon Graphics' newest scalable, shared-memory computing system and associated disks. Called Blacklight, the new system became available for research through the NSF TeraGrid in October; its extremely large, coherent shared-memory opens a new computational capability for U.S. scientists and engineers.

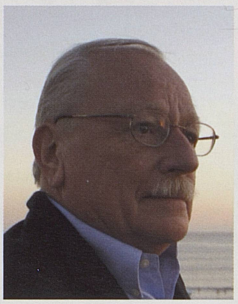
Featuring 512 eight-core Intel Xeon (Nehalem) processors (4,096 cores) with 32 terabytes of memory, Blacklight is partitioned into two connected 16-terabyte coherent shared-memory systems — creating the two largest coherent shared-memory systems in the world.

"Because of the extraordinary memory size and relative ease of programming made possible by this system's shared-memory structure, scientists and engineers will be able to solve problems that were heretofore intractable," said PSC scientific directors Michael Levine and Ralph Roskies. "For many research communities — including data-analysis and many areas of computer science — it will open the door to use of high-performance computation and thereby expand the abilities of scientists to ask and answer questions."

In computer terms, "shared memory" means that a system's memory can be directly accessed from all of its processors, as opposed to distributed memory (in which each processor's memory is directly accessed only by that processor). Because all processors share a single view of data, a shared memory system is, relatively speaking, easy to program and use. Coherence, a feature related to the synchrony of read-write operations by different processors within the system, is an important feature in many large data-analysis tasks.

"We are extremely excited about extending Silicon Graphics' long-standing history of working with PSC on systems that deliver breakthrough scientific research and engineering results," said Mark Barranchea, president and CEO of Silicon Graphics. "For over 15 years, PSC has used Silicon Graphics Data Migration Facility for hierarchical storage management, and over the past years they have used two of our current SGI Altix 4700 scalable shared-memory systems to conduct breakthrough research in chemistry and life sciences. We view PSC as a special partner in the development and delivery of our next generation, scalable shared-memory systems, which will bring new performance, flexibility and power efficiency to the supercomputing industry."

The 4,096 processor cores and 32 terabytes of shared memory are interconnected using Silicon Graphics' next-generation high-bandwidth, low-latency NUMAlink 5 interconnect. This interconnect enables scalable shared-memory or message-passing applications to run with higher levels of parallel efficiency, which means researchers can assign more processor cores simultaneously to the same task — to address larger problems and solve them more quickly.



JIM KASDORF
PSC director of special projects

CREATING NATIONAL CYBERINFRASTRUCTURE

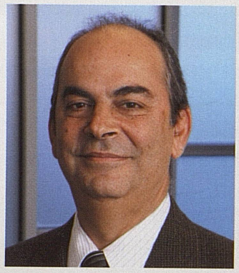
TeraGrid™

The TeraGrid is the world's most comprehensive distributed cyberinfrastructure for open scientific research. As a partner in this National Science Foundation program, PSC helps to shape the vision and progress of the TeraGrid.

PSC AND TERAGRID PSC is actively involved in TeraGrid leadership. Co-scientific director Michael Levine is PSC's representative to the TeraGrid Forum — TeraGrid's principal decision-making group. Other PSC staff with TeraGrid leadership roles include Sergiu Sanielevici, Area Director for User Support and Jim Marsteller, head of PSC's security, who chairs the TeraGrid Security Working Group. Laura McGinnis plays a lead role in TeraGrid education, outreach and training activities. PSC director of systems and operations, J. Ray Scott, leads the TeraGrid effort in Data Movement, and PSC director of strategic applications, Nick Nystrom, leads the TeraGrid Extreme Scalability Working Group, which fosters planning to meet the challenges of deploying extreme-scale resources into the TeraGrid. PSC staff members serve on all of TeraGrid's working groups.

TERAGRID RESOURCE PROVIDERS

Indiana University
Louisiana Optical Network Initiative
National Center for Supercomputing Applications
National Center for Atmospheric Research
National Institute for Computational Sciences
Oak Ridge National Laboratory
Pittsburgh Supercomputing Center
Purdue University
San Diego Supercomputer Center
Texas Advanced Computing Center
The University of Chicago/Argonne National Laboratory



DAVID MOSES
PSC Executive Director

THE PSC STAFF, WHOSE WORK THROUGH TERAGRID AND OTHER PROGRAMS MAKES A LASTING CONTRIBUTION TO U.S. CYBERINFRASTRUCTURE AND TO KNOWLEDGE OPENED UP THROUGH THE REMARKABLE ACHIEVEMENTS OF COMPUTATIONAL SCIENCE:

Row 1 (l to r): R. Reddy, Ken Hackworth, Derek Simmel, Anjana Kar, Pallavi Ishwad, Ralph Roskies, Rich Raymond, Jack Chang, Boris Kaminsky, Markus Dittrich, Marcela Madrid. **Row 2:** Christal Banks (in white), Patricia Sudac, Art Wetzel, Deb Nigra, Bob Stock, Lucille Jarzynka, Robin Flaus-Scibek, Shandra Williams.**Row 3:** Kathy Benninger, J. Ray Scott, Andy Adams, Jenda Domaracki, Michael Lambert, Jason Sommerfield, Dustin Sorge, David Moses, Brian Johanson, Ed Berger, Ryan Omecene, Rick Costa. **Row 4:** Adam Fest, Burt Cubbison, Greg Foss, Hugh Nicholas, Ken Goodwin, Dave Graham, Michael Schneider, Zhihui Zhang, Jared Yanovich, Roberto Gomez, Yang Wang, Joel Welling. **Row 5:** John Kochmar, Ed Wozniak, Shane Filus, Steve Cunningham, Ray Nardozzi, Dave Kapcin, Chad Vizino, Nathan Stone, James Marsteller, Kevin Sullivan, Nick Nystrom, Clint Perrone. (Not pictured: Beth Albert, Erica Anderson, Cheryl Begandy, Vivian Benton, Phillip Blood, Janet Brown, Shawn T. Brown, Robert Budden, Brian Chen, Brian Gill, Ed Hanna, Greg Hood, Wendy Huntoon, Anirban Jana, James Keener, Joe Lappa, Michael Levine, Rob Light, Susan Litzinger, Jeremy Lipson, Mahin Mahmoodi, Tom Maiden, Laura McGinnis, Paul Nowoczynski, David O' Neal, Josephine Palencia, Stephen Petko, Chris Rapier, Alex Ropelewski, Sergiu Sanielevici, Joel Stiles, John Urbanic, Bryan Webb, Troy Wymore.)

SUPERCOMPUTING IN PENNSYLVANIA

With Commonwealth of Pennsylvania support, PSC provides education, consulting, advanced network access and computational resources to scientists and engineers, teachers and students across the state.

OUR REGION'S BUSINESS: TO OUTCOMPETE YOU MUST OUTCOMPUTE In March, Cheryl Begandy, PSC director of education, outreach and training, who coordinates PSC's programs in Pennsylvania, appeared on the TV show "Our Region's Business" with Pittsburgh journalist Bill Flanagan (left). Mike Makowski (third from left) of PPG and Jay McClatchey of Applied Computational Technologies (ACT) in Windber, Pennsylvania joined Begandy. Makowski brought along PPG transition lens sunglasses and self-washing windowglass, products developed through quantum-theory based computing at PSC. McClatchey discussed ACT's reliance on PSC intellectual resources in their work developing next-generation radiation dose calculation for cancer therapy (see p. 44). Drawing from studies by the Council on Competitiveness, a Washington, DC organization of business, labor, academic and governmental leaders, Begandy summed up her message as "To outcompete you must outcompute."

DISCOVER 10: PSC OPEN HOUSE In March PSC held its second annual Open House. With the theme "Innovation through Collaboration," the event drew more than 80 attendees, including student groups from regional schools and universities, and representatives from industry, government and foundations. In the photo, PSC executive director Dave Moses is talking with a group of students from La Roche College in Pittsburgh.

GIGAPANORAMA Art Wetzel, PSC principal computer scientist, collaborated with David Bear of Carnegie Mellon's Studio for Creative Inquiry in creating the first "Pittsburgh Gigapanorama." This interactive, 360-degree portrait of southwest Pennsylvania as seen from the roof of the U.S. Steel Tower was unveiled on April 22. At print resolution, the image would be 50 feet high by 285 feet long, bigger than any screen could accommodate. Bear and his colleagues enlisted Wetzel for his expertise in data-intensive visualization, and Wetzel used his 8-processor 64-Gig RAM system to help manage, adjust and assemble the full image files.

Art Wetzell (left) & David Bear

K-12 SCIENCE EDUCATION With K-12 educational and outreach programs (see pp. 7-9), PSC helps to prepare technology-ready workers and a science-literate populace. PSC workshops this year trained 144 science teachers in 40 Pennsylvania school districts.

Along with its workshops, PSC in January hosted 15 juniors and their mentors from McKeesport High School, providing an enrichment session for the McKeesport Administrative Academy, a program to help students improve their math scores. PSC staff introduced computational modeling and explored concepts such as probability, random numbers and data analysis. "The dedication of the PSC staff to opening a new world to both students and teachers through their computational tools and simulations," said Timothy M. Gabauer, district superintendent of the McKeesport Area School District, "has made an indelible mark on the teaching and learning process."

Two PSC staff members now serve on Occupational Advisory Committees (OACs) for the Pittsburgh Public Schools' new Science and Technology Academy. The OACs recommend improvements and provide assistance to assure up-to-date curriculum. At the 2010 Pittsburgh S.T.E.M. Summit organized by the Pittsburgh Technology Council, PSC staff participated in a panel titled "Connecting Today's Students with Tomorrow's Careers." This half-day event highlighted ways that the region's technology industry can affect educational change to address workforce shortages in STEM fields.

COMMUNITY OUTREACH PSC staff take part in numerous outreach programs, locally and nationally. In March at the Carnegie Science Center, PSC exhibited at the 2010 SciTech Initiative. PSC's booth — visited by more than 400 attendees — presented materials on the theme "What is supercomputing and how is it used?" PSC staff also supported local science education by serving as judges for regional science and technology fairs. In March, four staff members judged projects for the Pennsylvania Computer Fair. Later in March, three staff members helped judge the annual Pittsburgh Regional Science and Engineering Fair in which over 1000 students took part.

PRIVATE-SECTOR & UNIVERSITY RESEARCH

Research by scientists in Pennsylvania supported by PSC is exemplified by several projects in this booklet:

* **Asking "What If?" About H1N1:** PSC scientist Shawn Brown and the University of Pittsburgh Graduate School of Public Health MIDAS Center of Excellence (p. 20).
* **It Knows What You're Thinking:** fMRI and Noun Meaning, Tom Mitchell and colleagues, Carnegie Mellon University (p. 24).
* **The Joy of Giving:** Modeling Private Donation to Charitable Organizations, Holger Sieg, University of Pennsylvania (p. 45).
* **The Right Dose:** New Technologies for Dose Calculation in Radiation Treatment of Cancer, Applied Computational Technologies (p. 44).

RESEARCH & TRAINING AT PENNSYLVANIA COMPANIES, COLLEGES & UNIVERSITIES, 2009-2010

From July 2009 through June 2010, PSC workshops in high-performance computing reached 122 Pennsylvania grad and undergrad students, and PSC provided nearly 12.5 million processor hours to 864 individual Pennsylvania researchers from 35 institutions. The following Pennsylvania corporations, universities and colleges used PSC resources during this period:

Air Products and Chemicals, Inc.
Allegheny General Hospital
Allegheny-Singer Research Institute
Bloomsburg University of Pennsylvania
Bryn Mawr College
Bucknell University
Cabrini College
Carnegie Mellon University
Cedar Crest College
Chatham College
Cheyney University of Pennsylvania
Drexel University
Duquesne University
Fund Science
Haverford College
Indiana University of PA, all campuses
Lehigh University
Lock Haven University
Pennsylvania State University, all campuses
Philips Respironics
PPG, Inc.
Robert Morris University
Shippensburg University of Pennsylvania
Slippery Rock University
St. Vincent College
Swarthmore College
Temple University
Thomas Jefferson University
University of Pennsylvania
University of Pittsburgh, all campuses
Upper St. Clair High School
Ursinus College
Villanova University
Waynesburg College
Wilkes University

THE SUPER COMPUTING SCIENCE CONSORTIUM

Pennsylvania-West Virginia partners in development of clean power technologies

Formed in 1999 and supported by the U.S. Department of Energy, the Super Computing Science Consortium is a regional partnership of research and educational institutions in Pennsylvania and West Virginia. $(SC)^2$ provides intellectual leadership and advanced computing and communications resources to solve problems in energy and the environment and to stimulate regional high-technology development and education.

Through $(SC)^2$, Evergreene Technology Park in Greene County provides a resource that supports and encourages companies to collaborate with local universities in southwest Pennsylvania and West Virginia and to have access to PSC.

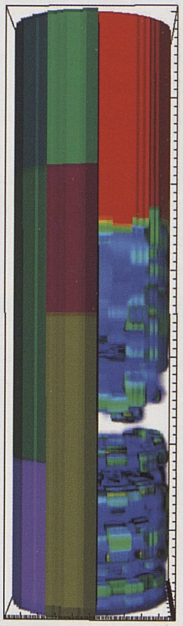
Since the spring of 2000, a high-speed network — the first fiber-optic service to Morgantown, West Virginia — has linked the National Energy Technology Laboratory (NETL) campuses in Morgantown and Pittsburgh with PSC, facilitating NETL collaborations. Researchers at NETL and WVU have actively used this link to tap PSC computational resources. In 2011 this link will be upgraded (as part of a contract between 3ROX, PSC's network exchange, and the National Oceanic and Atmospheric Administration Environmental Security Computing Center in Fairmont, West Virginia) from its current bandwidth (155 megabits per second) to 10 gigabits per second.

$(SC)^2$ co-chairs Lynn Layman, PSC (right) & Bob Romanowsky, NETL

RESEARCH COLLABORATION: PSC & NETL

During 2010, PSC collaborated with NETL staff on three research projects to improve NETL's ability to use the tools of computational science to advance its work in developing clean, affordable fossil-fuel technologies. One of these projects helped to accelerate the processing speed of MFIX (Multiphase Flow with Interphase Exchanges), NETL's award-winning software for simulating coal gasification and other clean-coal technologies. A second project improved the ability of NETL researchers to visualize large datasets quickly and easily, which included implementing VisIt, a software package for scalable visual analysis, on NETL's computing cluster and graphics accelerators. The third effort improved NETL's ability to handle large, complex, four-dimensional datasets.

PSC and NETL reviewed the first phase of this work in July. Among the highlights to date is that improved visualization capability with VisIt allows NETL to do analyses in minutes that used to take a full day. "Animations that took up to three days will now be possible in near-real time," says Nick Nystrom, PSC director of strategic applications, who coordinated the collaborative effort, "allowing engineers to save time and to communicate their results more effectively."



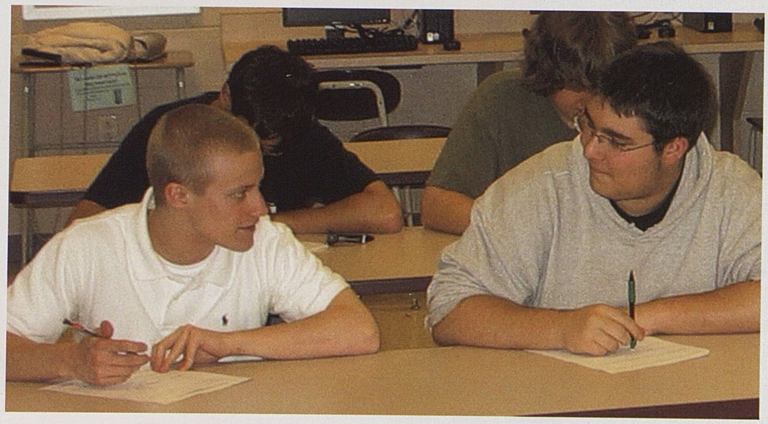
(Right) Deploying VisIt on NETL's accelerated cluster allows interactive visualization of results from complex MFIX simulations. This image reveals (color coding on left side) the processors on which different sections of the geometry of a simulation are being rendered.

ENERGIZING SCIENCE LEARNING

With Commonwealth of Pennsylvania support, PSC provides education, consulting, advanced network access and computational resources to scientists and engineers, teachers and students across the state.

"It's the difference between reading a textbook and visually experiencing a topic such as diffusion or osmosis," says Pallavi Ishwad, education program director of the National Resource for Biomedical Supercomputing (NRBSC), PSC's biomedical program. She's speaking of CMIST (Computational Modules in Science Teaching), one of several programs in secondary science education that PSC has introduced to teachers and to science classrooms in the Pittsburgh region and beyond over the past five years. A former high-school biology teacher, Ishwad has seen first-hand the difference that computational tools, such as vivid 3D animations produced by supercomputer simulations, can make in science learning.

"Introducing 'cool' technology into the classroom engages students," says PSC's director of outreach and education, Cheryl Begandy, "and increases their willingness to stay with subjects they may otherwise find too complicated or just uninteresting." For Begandy and Ishwad along with other PSC staff the goal is to help in re-imagining high-school science instruction so that it better prepares future scientists, engineers and educators. Ultimately the goal is to create the cyber-savvy workforce demanded by the 21st century marketplace.



PSC & (SC)²: RESEARCH FOR CLEAN ENERGY

Since the 1999 founding of (SC)², 51 (SC)² researchers have used PSC systems for a range of clean-energy related projects, using more than six million hours of computing time, over 376,000 hours within the past year.

THIS WORK INCLUDES:

High-Fidelity Simulation of Turbulent Combustion
www.psc.edu/science/2008/sc2/

Clean Liquid Fuel from "Syngas"
www.psc.edu/science/2006/sc2/

Fuel-Quality Hydrogen from Fossil Fuels
www.psc.edu/science/2005/sc2

Gas from Black Liquor
www.psc.edu/science/2004/sc2/

Fluidized-Bed Combustion of Silane
www.psc.edu/publicinfo/netl/

Lean-Fuel Mixes in Next-Generation Power-Generating Turbines
www.psc.edu/science/Richards/clean_power.html

A New Design for a Power-Generating Turbine
www.psc.edu/science/cizmas2002.html

(SC)² Super Computing Science Consortium

(SC)² PARTNERS

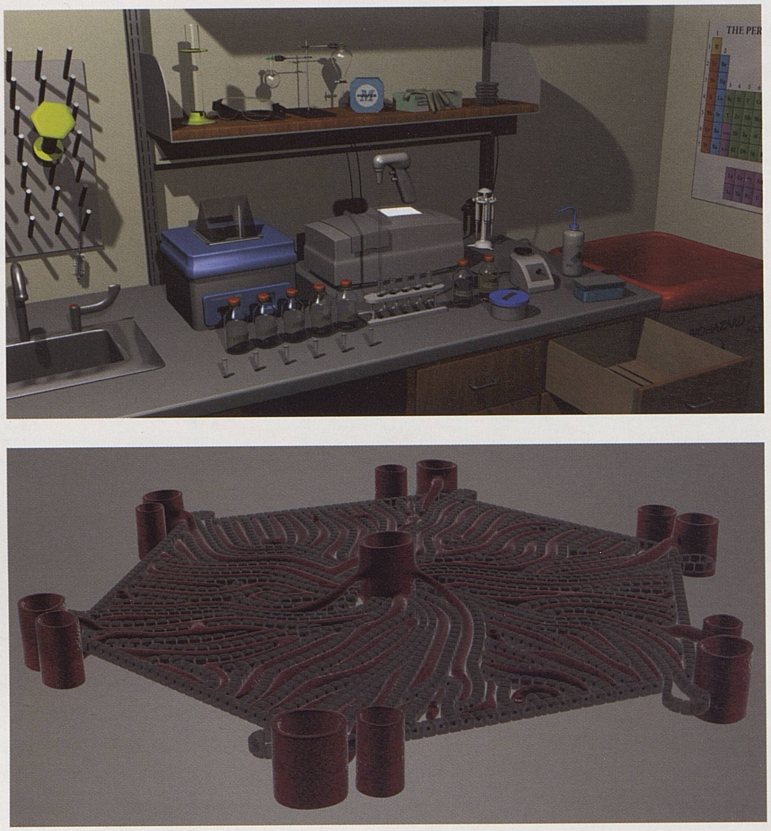
National Energy Technology Laboratory
Pittsburgh Supercomputing Center
Carnegie Mellon University
University of Pittsburgh
Waynesburg University
West Virginia University

More information: *www.sc-2.psc.edu*

CMIST Introduced in 2007, CMIST provides multi-disciplinary teaching materials — including lecture slides, animations and lesson plans — as ready-to-use on-line modules and DVDs. Produced with high-quality, biologically realistic 3-D animations, the modules are geared to lead students at many levels, from high school to grad school, toward an integrated understanding of biology, chemistry, physics, math and computation.

The initial CMIST module, "Molecular Transport in Cells" — produced with software called MCell and DReAMM, co-authored by NRBSC director Joel Stiles — presents important principles of osmosis and diffusion. It prompted enthusiastic feedback and created interest nationally, and to date over 700 people locally and nationally have attended CMIST presentations. A second module, "Big Numbers in Small Spaces: Simulating Atoms, Molecules, and Brownian Motion," premiered in 2009. During 2010, NRBSC developed a third CMIST module, "Enzyme Structure and Function," and Ishwad introduced it at the Supercomputing 2010 conference in New Orleans.

FREE CMIST MODULES: WWW.NRBSC.ORG/CMIST



From "Enzyme Structure and Function," the CMIST virtual laboratory (top) includes a spectrophotometer, scale, micropipettes and other common lab equipment used in MCell simulations of an experiment illustrating enzyme kinetics. This CMIST module also depicts a cross section of a hepatic lobule from the liver (bottom). Blood vessels (red) radiate outward from the central vein toward the hepatic arteries and hepatic portal veins at the periphery. Amongst the vasculature are hepatocytes (brown), the primary cells of the liver.

BEST Begun in 2007 by Ishwad, Better Educators of Science for Tomorrow (BEST) introduces high-school teachers to a bioinformatics curriculum adapted from an NRBSC program called MARC (Minority Access to Research Careers) for undergrad and graduate science students. Drafted and improvised through classroom usage over three years by an interdisciplinary group of high school teachers representing physics, chemistry, biology, mathematics and technology, the BEST curriculum offers ready-to-use lesson plans for single-subject trained educators to teach a multidisciplinary subject like bioinformatics in high schools.

During the 2009-10 academic year, Ishwad piloted BEST as a permanent course offering in three southwest Pennsylvania high schools. Four high school students from these piloting schools completed a first-ever Summer Research Internship Program in Bioinformatics at PSC guided by senior scientists from PSC. Their teachers presented the results of these internships in November at the 2010 SC Education Resource Fair in New Orleans.

PSC Education Program Director Pallavi Ishwad (front row, second from right) and PSC staff member Alex Ropelewski (back right) with participants in the BEST Summer Workshop at PSC (June 14-24, 2010).

Two additional schools participated in the 2010 BEST summer workshop and will offer the BEST curriculum in the upcoming school year. "I am excited to use what I learned here," said Dean Walker, a physics teacher at Seneca Valley High School outside Pittsburgh, "to both inspire students who love biology but hate physics, as well as provide students with a better feel for contemporary positions in the science job market."

MORE INFORMATION ON BEST
www.psc.edu/eot/k12/best.php

Pittsburgh-area high school students (l to r) Annie Kayser, Creg Milko and Shilo Crook present their project at the BEST Summer Workshop.

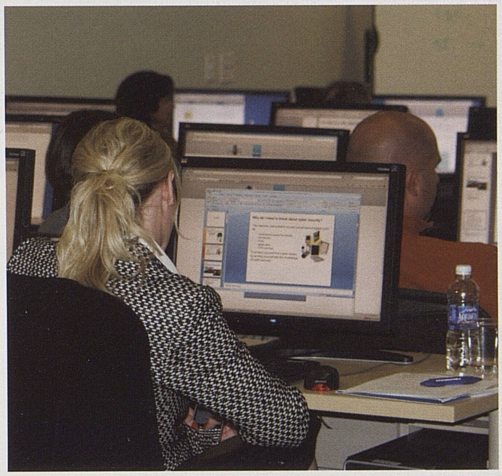
SAFE-NET Do Internet passwords protect personal information from unwanted intrusion? How can you be sure if someone on-line is who they say they are? Does anti-virus software really protect your hard-drive?

To help parents, educators, students and individuals with these questions and many others associated with wide usage of the Internet, PSC in 2010 introduced SAFE-Net, a program funded by an NSF grant for Cyber Safety Awareness. Through SAFE-Net, PSC presents workshops that train educators and provide materials for classroom learning — developed in collaboration with the CERT (Computer Emergency Response Team) Program at CMU's Software Engineering Institute. These materials address cyber threats, measures of protection, and questions of cyber ethics that arise as a result of social networking and other wide uses of the Internet.

"Many Internet users lack an understanding of common threats they may face online," says Begandy. "Among parents, many lack confidence that their child is safe when using the Internet."

In 2010, PSC held two Train-the-Teacher workshops introducing SAFE-Net to 18 Pittsburgh-area teachers. Through a videoconference, arranged via PSC networking infrastructure, Wiam Younes, training and awareness coordinator with CMU's Information Security Office, consulted with 75 educators at six school districts in eastern Pennsylvania. Titled "The Educator's Role in Safe Computing," this three-hour interactive program highlighted how K-12 educators can raise awareness about cyber security and safety in the schools.

The SAFE-Net website provides free information, including classroom and parent materials about cyber-security issues, with lessons geared to grade levels 1-3, 4-6, and 7-12.



SAFE-NET WEBSITE (FREE MATERIALS)
http://csa.3rox.net

Funding for these PSC education and outreach programs has come from Pittsburgh-area foundations — The Buhl Foundation, Grable Foundation and Heinz Endowments — and from the NIH along with additional support from the Commonwealth of Pennsylvania and the NSF TeraGrid program.

THE NATIONAL RESOURCE FOR BIOMEDICAL SUPERCOMPUTING

National Leadership in High-Performance Computing for Biomedical Research

Established in 1987, PSC's National Resource for Biomedical Supercomputing (NRBSC) was the first external biomedical supercomputing program funded by the National Institutes of Health (NIH). Along with core research at the interface of supercomputing and the life sciences, NRBSC scientists develop collaborations with biomedical researchers around the country, fostering exchange among experts in computational science and biomedicine and providing computational resources, outreach and training.

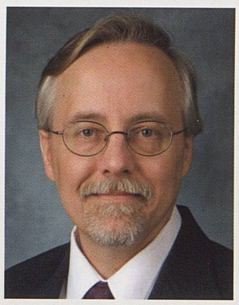
ANTON COMES TO PSC In September 2009, the National Institute of General Medical Sciences, part of NIH, awarded $2.7 million to NRBSC to support a partnership with D. E. Shaw Research to make an innovative new computing system, called Anton, available to U.S. biomedical scientists. This system, with hardware and software specialized to run molecular dynamics (MD) for the study of biomolecular systems such as proteins and nucleic acids, runs MD up to 100 times faster than conventional supercomputers, making it possible for the first time to extend MD simulations into the millisecond range of biological time.

While there are several Anton systems used internally at D. E. Shaw Research, NRBSC will host the first and thus far only one available to the general scientific community. A review committee convened by the National Research Council at the National Academies of Science reviewed and approved 47 proposals from scientists around the country to use Anton, and production research commenced in October 2010.

"This is an incredibly exciting project in many ways," says NRBSC director Joel Stiles. "With this very generous gift from D. E. Shaw Research and the funding provided by NIH, we are deploying a tool of unprecedented power for the benefit of biomedical researchers nationally. We expect that this project will help to significantly advance our understanding of biomolecular structure and function, and to spur ongoing scientific and technological development in MD research and in other areas of computational biology."

MORE INFORMATION

www.nrbsc.org



THE NRBSC TEAM

(Above, left to right) Jenda Domaracki, Jun Ma, Christal Banks, Markus Dittrich, Nikolay Simakov, Boris Kaminsky, Hugh Nicholas, Pallavi Ishwad, Art Wetzel, Greg Hood, Troy Wymore, Jack Chang, Gary Blumenthal

(Left) Joel Stiles, director of NRBSC

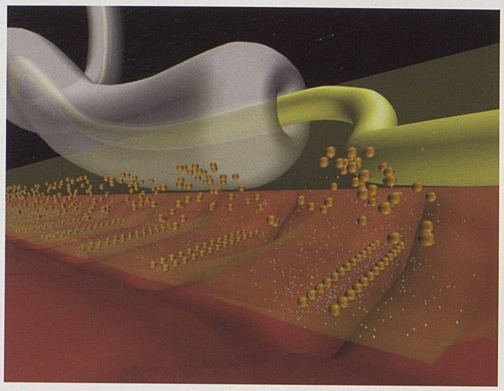
(Not pictured: Jacob Czech, James Keener, Alex Ropelewski)

NRBSC BIOMEDICAL COLLABORATIONS

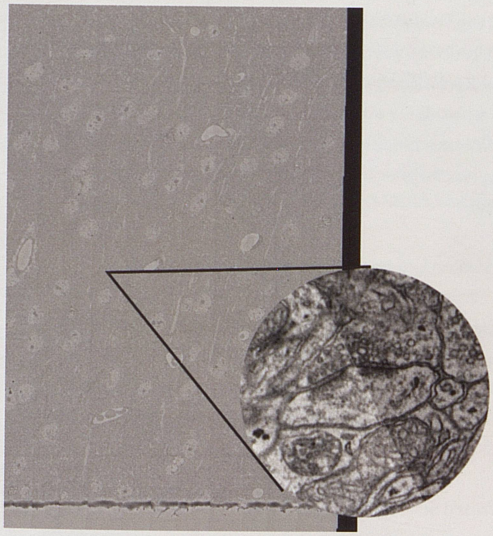
Albert Einstein College of Medicine
Carnegie Mellon University
Duke University
Harvard University
Howard University
Marine Biological Laboratory, Woods Hole
The Salk Institute
University of California at Davis
University of California at San Diego
University of Pittsburgh
University of Pittsburgh School of Medicine
University of Puerto Rico, Medical Sciences Campus
University of Michigan

RESEARCH NRBSC research focuses on three areas of biomedicine that span many scales of space and time: spatially realistic cell modeling, large-scale volumetric visualization and analysis, and computational structural biology.

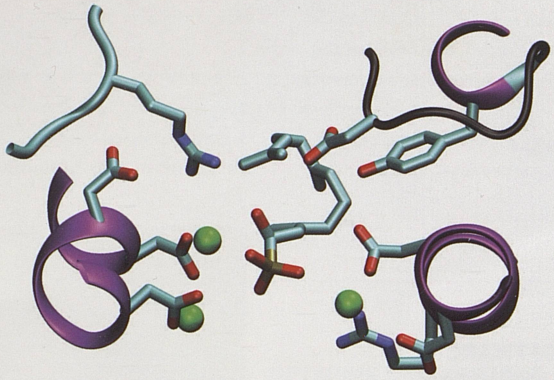
Spatially realistic cell modeling centers on realistic 3-D simulations of movements and reactions of molecules within and between cells, to better understand physiological function and disease. *MCell, DReAMM* and *PSC_DX* software is developed at the NRBSC and used to model and visualize events such as (shown in this image) neurotransmission between a nerve and muscle cell.



Volumetric visualization using the NRBSC's PSC_VB software enables multiple users to share, view and analyze extremely large datasets and time series obtained from light and electron microscopes, CAT and MRI scanners, etc. This reconstructed section of a mouse brain (119,600 x 88,400 pixels) includes a full resolution inset showing alignment at a frame boundary (indicated by artificial contrast).



NRBSC structural biology focuses on computational tools used to determine the structure of proteins from their amino acid sequence and development of quantum-mechanical simulation methods for biomolecules such as enzymes. This image shows conserved residues in the "active site" of an enzyme, 5-epi-artistolocholene synthase, which is involved in the production of compounds called terpenes. PSC-developed software enables researchers to simulate enzyme reactions, to reproduce experimental reaction rates and gain new insight into enzyme function, which facilitates design of new therapeutic drugs.



COMPUTATIONAL SERVICE & TRAINING Since NRBSC's inception, PSC and NRBSC together have provided access to computing resources for more than 1,500 biomedical research projects involving more than 4,600 researchers at 285 research institutions in 46 states and two territories. Among these are several projects featured in this booklet (pp. 20 & 24).

NRBSC training activities reach hundreds of scientists each year. More than 4,600 researchers have participated in NRBSC workshops in such areas as spatially realistic cell modeling, volumetric data visualization and analysis, protein and DNA structure, genome sequence analysis and biological fluid dynamics

NRBSC participates in a range of undergraduate and graduate training programs. These include:

A joint Carnegie Mellon and University of Pittsburgh Ph.D. program in computational biology: *www.compbio.cmu.edu*

The Ray and Stephanie Lane Center for Computational Biology at Carnegie Mellon: *lane.compbio.cmu.edu*

The University of Pittsburgh Department of Computational Biology: *www.ccbb.pitt.edu*

The NRBSC and PSC have developed educational programs, CMIST and BEST (see pp. 9-11), that have provided training to high-school and undergrad students and educators in the Pittsburgh region and nationally.

NETWORKING THE FUTURE

One of the leading resources in the world for network know-how

PSC's Advanced Networking group is one of the leading resources in the world for knowledge about networking. Through 3ROX (Three Rivers Optical Exchange), a high-speed network hub, they operate and manage network infrastructure that connects many universities and schools in Pennsylvania and West Virginia to research and education networks, such as Internet2 and National LambdaRail, that link to universities, corporations and research agencies nationally. Their research on network performance and analysis — in previous projects such as Web100 and the NPAD diagnostic server — has created valuable tools for improving network performance.

MORE INFORMATION
www.psc.edu/networking/

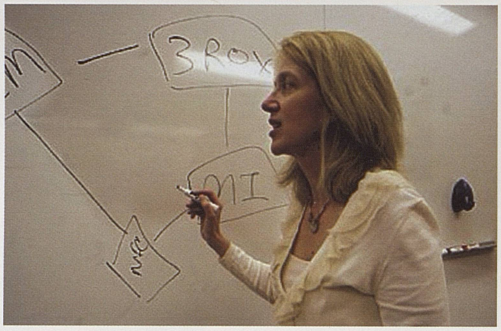
NSF AWARDS FOR SOFTWARE DEVELOPMENT AND NETWORK INFRASTRUCTURE Two grants from the National Science Foundation this year — one from the Software Development for Cyberinfrastructure (SDCI) program and another from the Academic Research Infrastructure (ARI) program — support, respectively, 3ROX research to build on Web100 and a major upgrade to the 3ROX network environment.

A $980,000 SDCI award supports a three-year project called Web10Gig. As a follow-up to Web100, Web10Gig will develop software to enable non-expert users to effectively exploit the bandwidth availability of advanced networks. Web100 ended in 2003 and produced prototype software still heavily used. PSC partnered on Web100 with the National Center for Atmospheric Research (in Colorado) and with the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign, and on Web10Gig is again partnering with NCSA.

"The potential broader impact of Web10Gig is huge," says PSC director of networking Wendy Huntoon. "It can make it easy for users from the broadest range of fields and technical abilities to use the network to its full capacity. Eliminating many common network problems will have a transformative effect for researchers in many disciplines."

An ARI award of $1.5 million for four years will enable 3ROX to renovate its network infrastructure, including upgrading its fiber-based optical capability to increase high-end transmission rates tenfold, from 10 to 100 gigabits per second. The upgrade will also include a pool of circuits (called transponders) that will enable researchers to experiment with bandwidth without having to include cost of the equipment in their grants, allowing them to be less encumbered by limitations that affect data-intensive research.

"We're happy about these new and faster technologies," says Huntoon, "as well as the institution of a pool of transponders for the 3ROX research community. We expect this to make bandwidth more accessible."



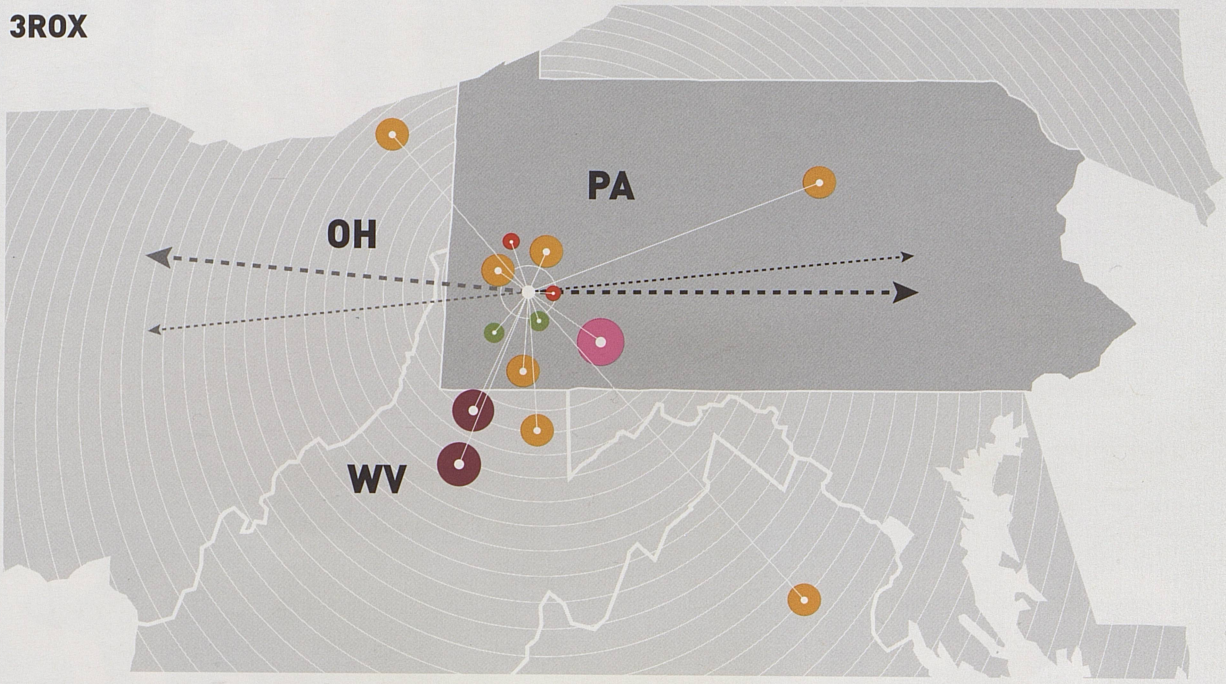
Wendy Huntoon, PSC director of networking. In September, the board of National LambdaRail (NLR) appointed Huntoon, NLR's chief technical officer, as acting CEO.

PENNSYLVANIA STATEWIDE BROADBAND NETWORK Last year 3ROX joined with a coalition of Pennsylvania colleges and universities, heathcare and economic development organizations to form the Pennsylvania Research and Education Network (PennREN). In February, PennREN received $100 million in federal stimulus money through the American Recovery and Revitalization Act and $29 million in private funds to build and maintain a broadband network for expanded educational opportunity and healthcare services across the Commonwealth of Pennsylvania. This network will reach the most rural areas of Pennsylvania, providing access to more than two million households and 200,000 businesses.

At a news conference (at Carnegie Mellon University) announcing the award, Pennsylvania governor Ed Rendell estimated that the project could generate 1,000 new jobs over the next two years. Building the network will be managed by the Keystone Initiative for Network Based Education and Research (KINBER). When complete, the PennREN network will allow teleconferencing between regional hospitals and urban hubs, necessitating fewer patient transfers. The project calls for an initial bandwidth of 40 gigabits per second by January 2012.

"This system will rival any in the United States," says PSC network director Huntoon, "and will provide the capability to connect regional networks across the Commonwealth."

3ROX



3ROX MEMBERS

UNIVERSITIES
Carnegie Mellon University, Pennsylvania State University, University of Pittsburgh, Waynesburg University, West Virginia University, Norfolk State University.

NLR MEMBER INSTITUTIONS
3ROX (PSC), Carnegie Mellon University, Case Western Reserve University, OneCommunity, University of Pittsburgh, Pennsylvania State University, Indiana University.

K-12 INSTITUTIONS
Allegheny Intermediate Unit (AIU3), Arin Intermediate Unit (IU28), Beaver Valley Intermediate Unit (IU27), Intermediate Unit One, Northwest Tri-County Intermediate Unit (IU5), Riverview Intermediate Unit (IU6), City of Pittsburgh School District (IU2), Seneca Highlands (IU9), Central IU (IU10).

GOVERNMENT LABORATORIES AND FACILITIES
The National Energy Technology Laboratory; NOAA Environmental Security Computing Center.

BUSINESS
Comcast, Westinghouse Electric Co.

OTHER
Computer Emergency Response Team.

NETWORK CONNECTIONS

NATIONAL RESEARCH NETWORKS
Internet2 — 5 Gbps, ESnet — 1 Gbps, National LambdaRail PacketNet — 10 Gbps, TeraGrid Extensible Backplane Network — 10 Gbps.

NATIONAL COMMODITY INTERNET NETWORKS
Global Crossing — 1 Gbps; Cogent — 1 Gbps.

PITTSBURGH LOCAL EXCHANGE NETWORKS
Comcast, MetNet & Cavalier.

OTHER NETWORK CONNECTIONS
Southern Crossroads (SOX) — 1Gbps, TransitRail — 1Gbps, OARnet — 1 Gbps, FrameNet — 10 Gbps.

Note: Gbps: a billion (Giga) bits per second.

PITTSBURGH SUPERCOMPUTING CENTER WORKSHOPS

2009 - 2010

Hands-On Workshop on Computational Biophysics Using NAMD and VMD

Molecular Dynamics with Desmond

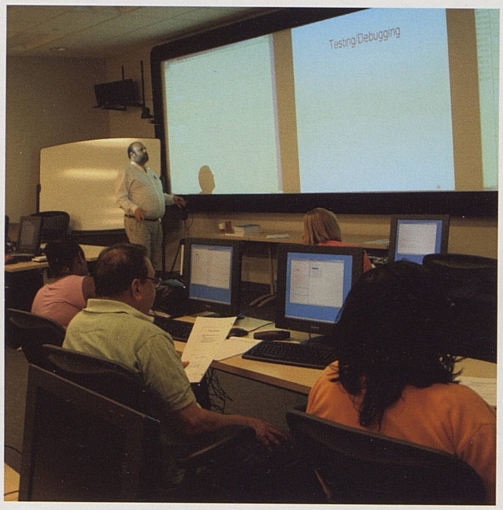
Summer Institute In Bioinformatics
(For Minority-Serving Institutions)

Bioinformatics Internship Program

Methods And Applications Of Hybrid QC/MM Simulations To Biomolecular Systems

Parallel Programming and Cluster Computing

TeraGrid New User Training



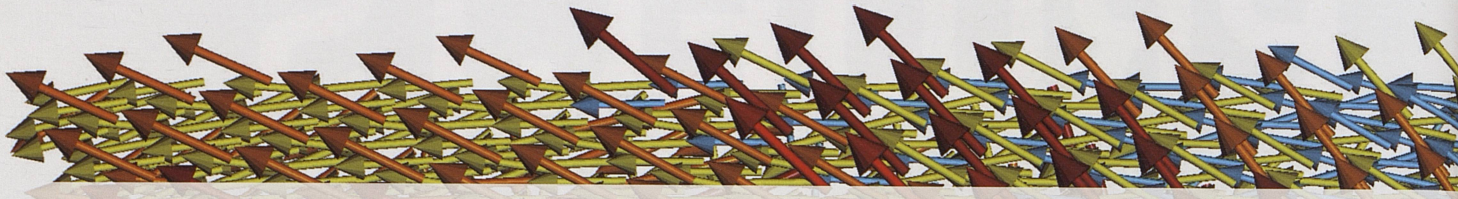
A workshop underway in the PSC Computer Training Center, the David W. Deerfield II Training Center, equipped with 30 "dual-boot" workstations and a projector for overhead display of the instructor's desktop.

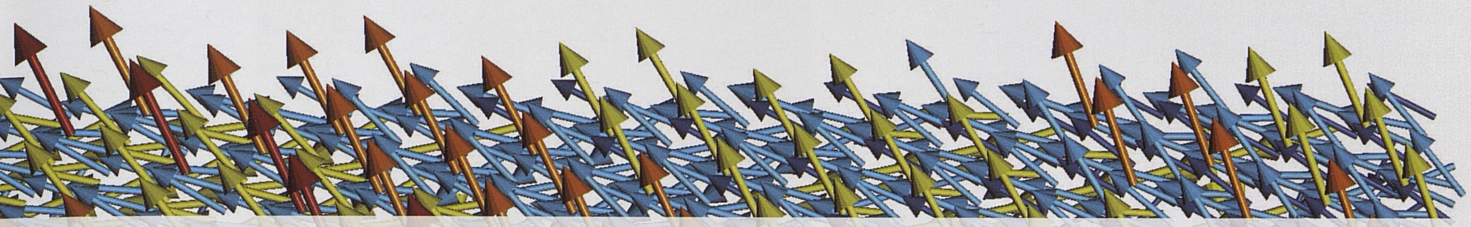
PSC'S DIRECTORS AND MANAGERS

PSC'S DIRECTORS AND MANAGERS (LEFT TO RIGHT), WHO OVERSEE THE DAY-TO-DAY OPERATIONS OF THE CENTER
Wendy Huntoon, director of networking; David Moses, executive director; Sergiu Sanielevici, director, scientific applications & user support; David Kapcin, director of financial affairs; Bob Stock, PSC associate director; Cheryl Begandy, director, education, outreach & training. Not pictured: Janet Brown, manager of networking; John Kochmar, manager of high-performance computing facilities; Nick Nystrom, director of strategic applications; J. Ray Scott, director of systems & operations; Richard Raymond, manager of user support; Joel Stiles, director of NRBSC.

PROJECTS *in* SCIENTIFIC COMPUTING 2010

PROJECTS 2010 CONTENTS





32

STRUCTURE OF PROTEINS AND NUCLEIC ACIDS

WHEN CELLULAR BONES SOFTEN

ACTIN FILAMENT REMODELING BY ACTIN DEPOLYMERIZATION FACTOR/COFILIN

Greg Voth, University of Chicago

Jim Pfaendtner, University of Washington

36

QUANTUM CHEMISTRY

NATURAL REACTION

QUANTUM CALCULATIONS OF REACTION PATHWAYS IN SESQUITERPENE BIOSYNTHESIS

Dean Tantillo, University of California, Davis

40

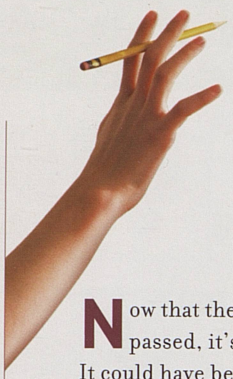
EVOLUTION & STRUCTURE OF THE UNIVERSE

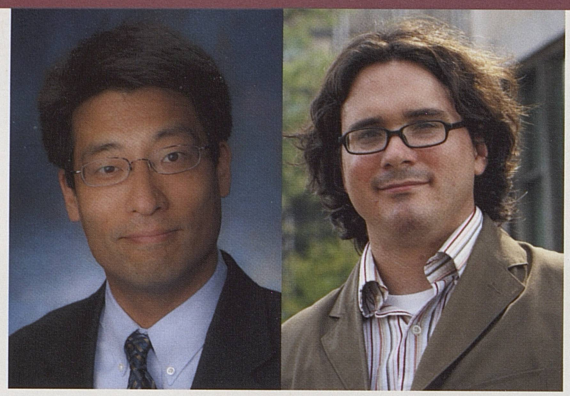
IN SEARCH OF BULGELESS DWARFS

THE ORIGIN OF BULGELESS DWARF GALAXIES AND DARK MATTER CORES

Fabio Governato & Thomas Quinn, University of Washington

IN PROGRESS





Bruce Lee
University of Pittsburgh
Graduate School
of Public Health

Shawn Brown
Pittsburgh Supercomputing
Center

Now that the 2009 H1N1 flu epidemic has passed, it's possible to feel some relief. It could have been much worse. Nevertheless, despite this novel virus strain turning out to be fairly mild relative to initial fears and prior flu pandemics such as 1918, which was 100 times more lethal, 2009 H1N1 was a major infectious disease outbreak — 12,000 deaths in the United States alone and thousands more worldwide.

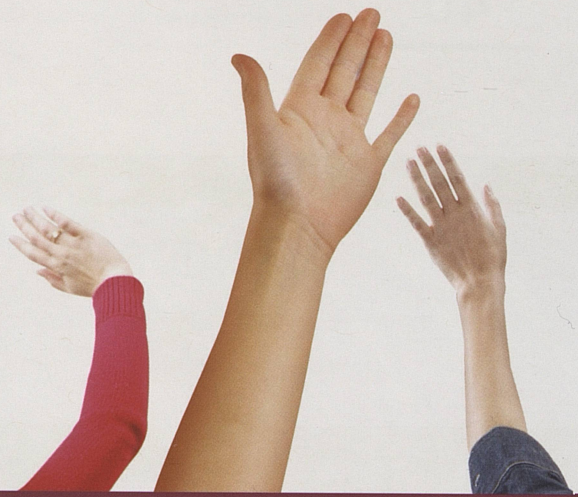
Realistically, it's a question of "when?" not "if" another virus will emerge as a far more serious public-health threat. And in that respect, part of the good news from 2009 is that epidemiological modeling, a powerful tool to assist decision makers, stepped into the fray. "Our models are a virtual laboratory to ask questions you can't ask with real populations," says Shawn Brown, PSC scientist and assistant professor in the University of Pittsburgh Department of Biostatistics. "We build a population, infect them with the flu, and then look at mitigation strategies, such as vaccinations or school closure, and see what effect it has."

Through the National Institutes of Health MIDAS (Models of Infectious Disease Agent Study) Center of Excellence, led by Donald Burke of the University of Pittsburgh Graduate School of Public Health (GSPH), Brown and collaborators from GSPH and other places used PSC's shared memory system, Pople, to model the spread of H1N1 on a regional basis, both in Allegheny County (which includes Pittsburgh) and in the Washington, D.C. metropolitan area. They did this modeling during 2009, developing results in real time, in response to requests from health officials and policy makers, as the severity of the H1N1 outbreak remained in question.

They shared their findings with the Allegheny County Health Department and officials for the state of Pennsylvania as well as the U.S. Biomedical Advanced Research and Development Authority (BARDA), the U.S. Department of Homeland Security, and the President's Council of Advisors on Science and Technology (PCAST). For three weeks in the fall of 2009, Brown and his GSPH colleague Bruce Lee were in effect embedded with BARDA. "They presented us with scenarios," says Brown, "and we did the modeling. Supercomputing really helped. We were able to get rapid response to complex scenarios. We're still doing that today."

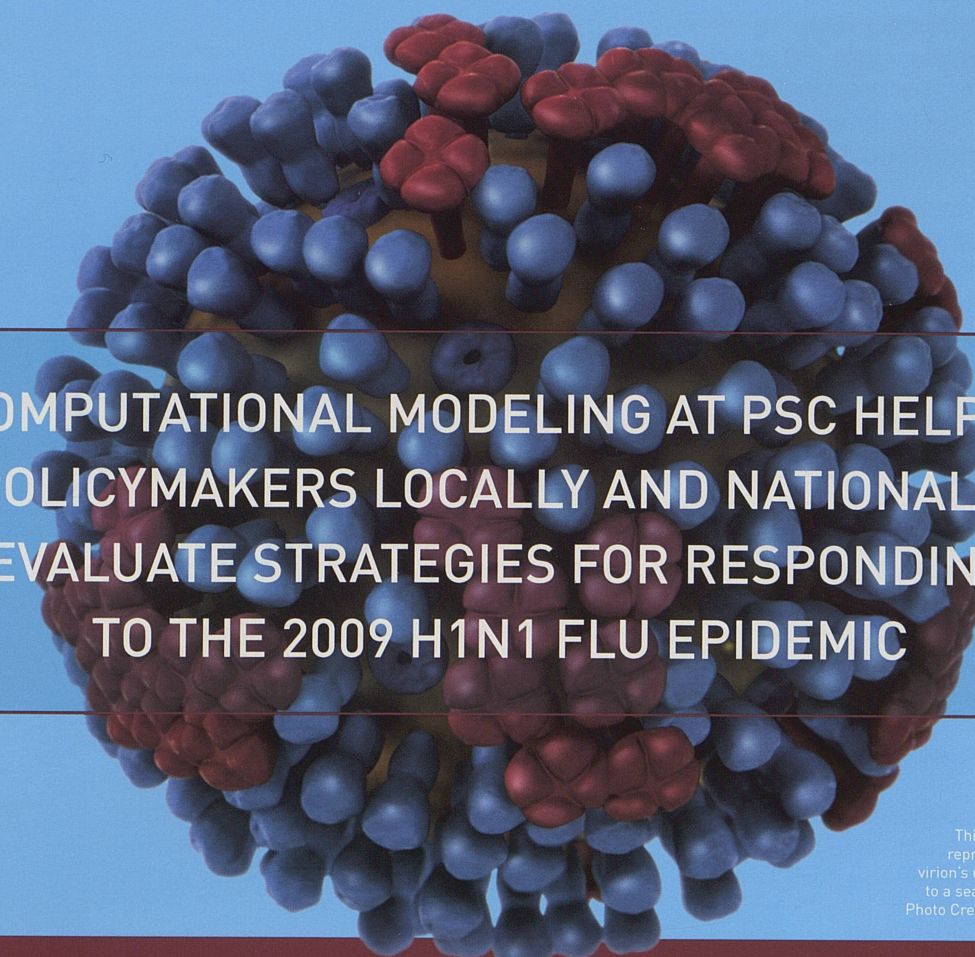
They found, for instance — in a study published in the *Journal of Public Health Management and Practice* (December 2009) — that to close schools less than two weeks may slightly increase infection rates, and that (contrary to Center for Disease Control recommendations) schools may need to be closed eight weeks or longer to have a significant impact.

They also used their model to investigate questions about vaccination priorities. Their findings — *Vaccine* (May 2010) — support recommendations by the U.S. Advisory Committee on Immunization Practices (ACIP) that priority be given to people at risk for severe complications. Prioritizing at-risk individuals, rather than only high transmitters (i.e., children), the modeling showed, may lead to slightly more cases of flu, but it reduces serious disease and death, and overall economic cost.



ASKING "WHAT IF?" ABOUT H1N1

COMPUTATIONAL MODELING AT PSC HELPED POLICYMAKERS LOCALLY AND NATIONALLY EVALUATE STRATEGIES FOR RESPONDING TO THE 2009 H1N1 FLU EPIDEMIC



This picture provides a 3D graphical representation of a generic influenza virion's ultrastructure, and is not specific to a seasonal, avian or 2009 H1N1 virus. Photo Credit: Illustrator: Dan Higgins, CDC

WORKING WITH AGENTS

Epidemiological modeling goes back to early 20th-century mathematical formulations that attempt to quantify the spread of epidemics by identifying the susceptible proportion of a population and specifying a rate of transmissibility. As susceptible people become infected and recover (or not), enough of the population eventually becomes immune and the epidemic passes. This fairly crude tool to estimate the length and severity of a disease outbreak has over the past two decades, with powerful computing and sophisticated software methods, gained complexity and greatly improved ability to reflect the reality of how infectious disease spreads.

Increased complexity is especially the case with "agent-based modeling" (ABM) — a relatively new approach that Brown and his colleagues used for their 2009 work on H1N1. ABM represents virtual persons as autonomous "agents" within a synthetic population built from the most accurate available data (such as the U.S. Census). As agents become infected with disease, their individual movements within the population — to work, school, cultural events, etc. — result in the virus being transmitted to other susceptible agents. Disease spread is based on algorithms that incorporate randomness, or *stochastic* processes.

"Disease spread is a stochastic process," says Brown. "It's not deterministic; you can't say with certitude when contact between an infected and susceptible agent will lead to infection. It's a statistically based outcome."

Other forms of epidemiological modeling, which include *compartmental modeling and network modeling*, are less detailed than ABM, and approximate certain aspects of a population and their interactions — making it possible to model larger populations and geographic regions, such as an entire nation. "All these models are valid," says Brown, "and all of them are useful. It just depends on what type of question you want to answer."

The Pittsburgh MIDAS group's ABM modeling incorporates disease data (how long infections last and recovery time), surveillance data (best available information on how many people are getting sick in real time) plus social and behavioral data. Families are assigned to households, children to schools, and agents to workplaces with commuting distance, location of hospitals and other demographic factors — developed from census data.

Because ABM represents an entire population inside the computer, it requires large amounts of memory. For the DC metropolitan area, MIDAS's ABM included 7.4 million people, requiring seven gigabytes of memory. "This is a shared-memory problem," notes Brown, referring to massively parallel systems, such as PSC's Pople, that allow each processor to access all the memory without message passing.

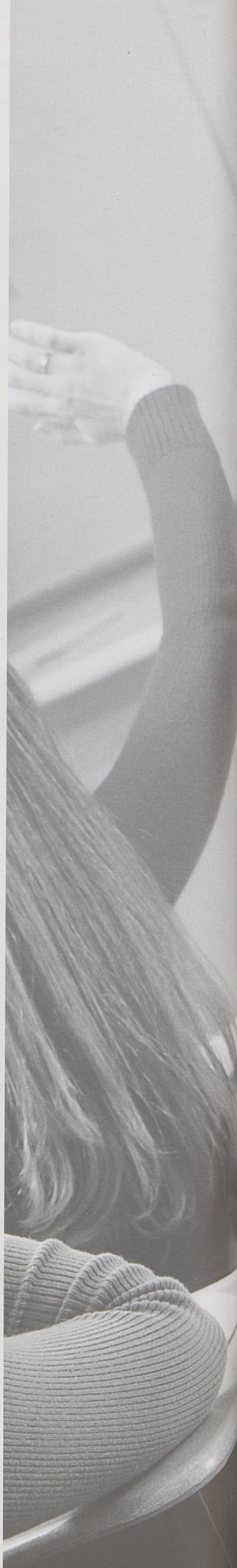
THE PITTSBURGH MIDAS NATIONAL CENTER OF EXCELLENCE

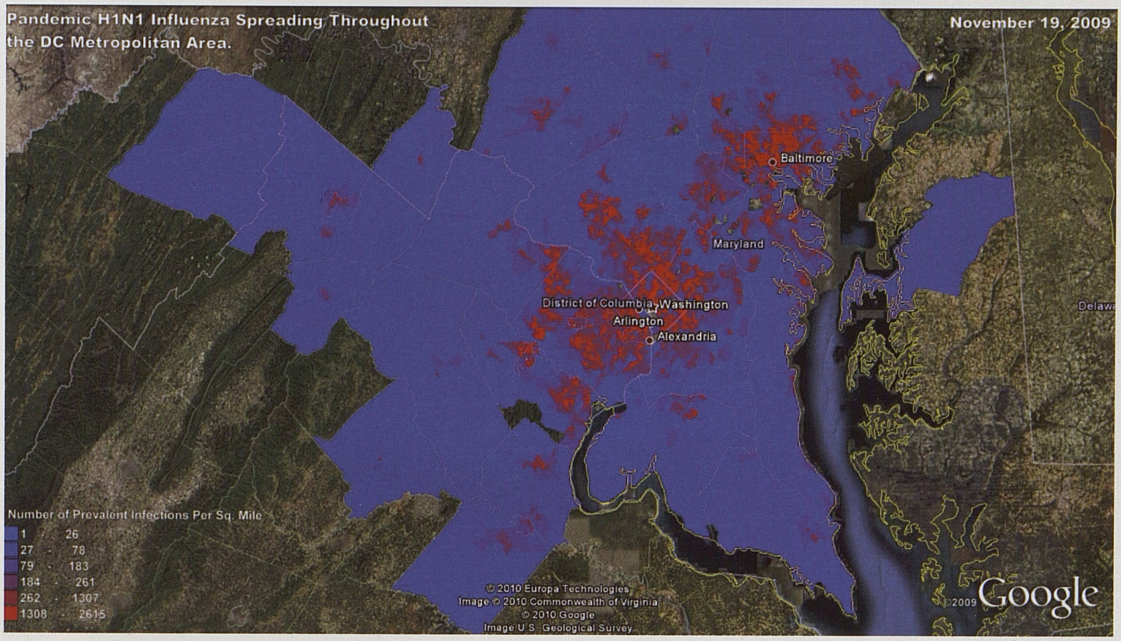
Rachel Bailey, MPH
Tina Assi
Shawn Brown, PhD
Donald Burke, MD
Phil Cooley, MS
Alona Fyshe
Josh Epstein, PhD
Ron Voorhees, MD
John Grefenstette, PhD
Bruce Lee, MD, MBA
Margaret Potter, JD
Roni Rosenfield, PhD
Sarah McGlone, MPH
Bill Wheaton, MS
Shanta Zimmer, MD
Rick Zimmerman, MD, MPH

SCHOOL CLOSURES & VACCINE PRIORITIES

It might seem obvious that to close schools would help to contain a flu outbreak — since children in contact with other children, who then bring it home to their families, is one of the primary ways that flu spreads through a community. Still, it's a step that imposes burdens on parents and, over time, economic costs on a community, as workers must either stay away from jobs or provide childcare. So, if you close schools to mitigate the spread of H1N1 (or some other flu), how long — in order to have optimum impact — should they remain closed?

This question arose during consultations with health officials of Allegheny County in the fall of 2009, and the MIDAS team addressed it with their ABM model. Their detailed simulations produced the unexpected finding that closing schools less than two weeks may actually prolong an epidemic. Short-duration school closures, they found, can increase transmission by returning susceptible students back to school in the middle of an epidemic when they are most vulnerable to infection.





SPREAD OF H1N1 IN THE WASHINGTON, DC AREA
From the epidemiological modeling of Shawn Brown and colleagues, this graphic shows infected individuals per square mile, coded by color (increasing from blue to red), at the peak of the 2009 H1N1 epidemic, overlaid on a Google earth contour map (green) of the area.

"Although closing schools may seem like a reasonable way to slow the spread of flu," says Lee, "we found it was not effective unless sustained for at least eight weeks. Closing schools quickly at the start of an outbreak was much less important than keeping them closed continually throughout the epidemic."

The study also found that identifying sick students individually and holding them away from school had minimal impact. And they found no significant differences in mitigating an epidemic between individual school closures and system-wide closure.

Later in 2009, spurred on by the initial limited availability of H1N1 vaccine, Brown and Lee, in collaboration with officials at PCAST and elsewhere, mounted a series of simulations looking in close detail at vaccine prioritization. With limited amounts of vaccine, what groups of people — children, elderly, caregivers, etc. — should be vaccinated first?

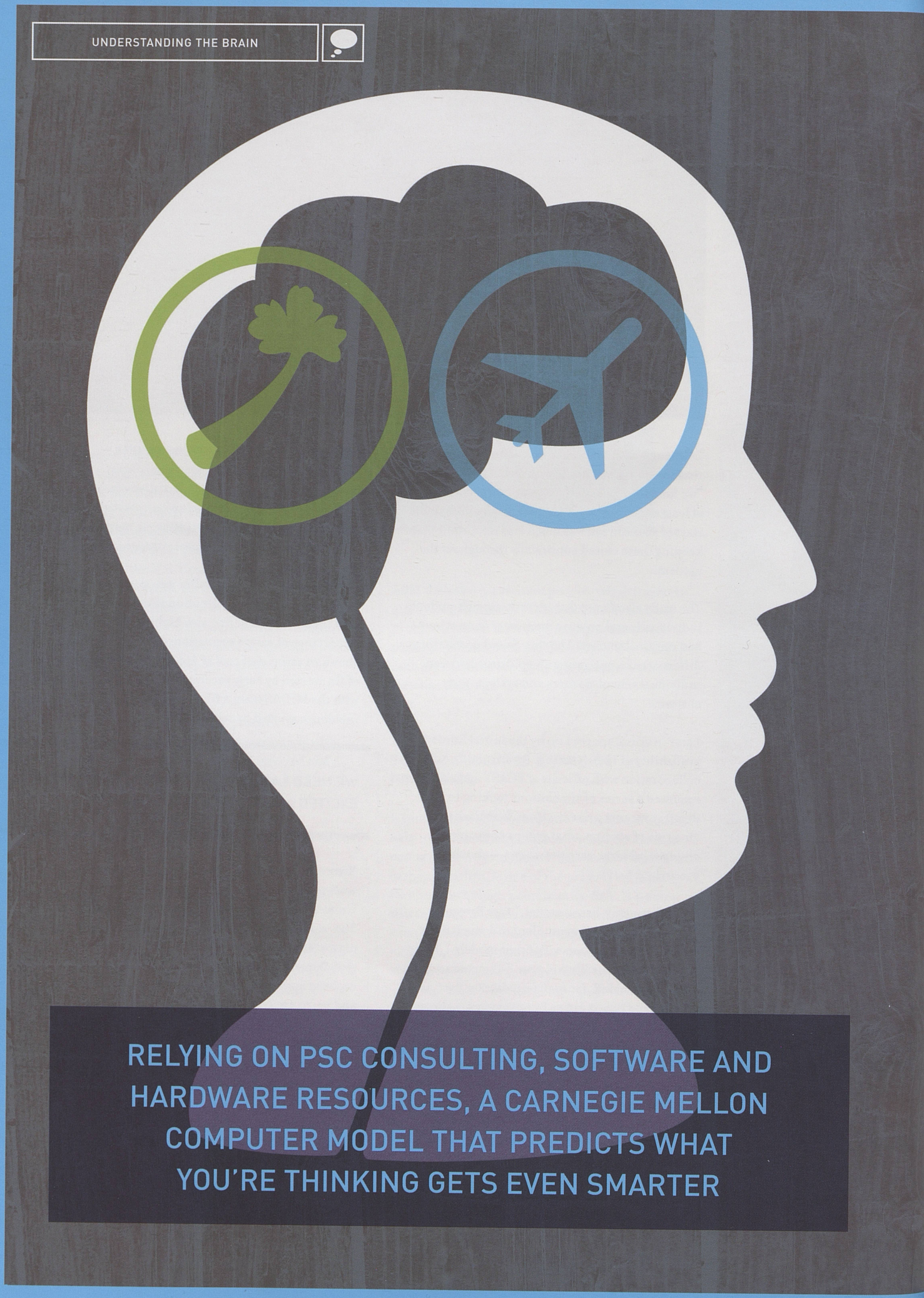
"With the agent-based model," says Brown, "we could explore this problem in a much more sophisticated way than with other models." Their ABM model of the Washington, DC metro area — which included, for instance, data on vaccine efficacy and how it varies by age groups — allowed the researchers to vary parameters and to look at many prioritization options. Recommendations of ACIP (which advises the U.S. government on immunization strategy) to prioritize groups at risk for severe reaction and hospitalization — mainly the elderly, pregnant women and families with newborns — along with the "high mixer" population of school-age children were borne out by the modeling.

The Pittsburgh MIDAS team also studied the effects of workplace vaccination, finding — as might be expected — that prioritizing workplaces with many employees higher than smaller job-sites reduces overall economic cost. Attention to economic cost with their ABM modeling also highlighted the importance of early vaccination; other research shows an associated cost of $100 for each incidence of flu averted by early vaccination, which compares with the MIDAS team's finding of $21,000 for each incidence averted by school closure.

WE NEED LARGE SHARED MEMORY, AND WE'RE EXCITED ABOUT THE UV SYSTEM AT PSC.

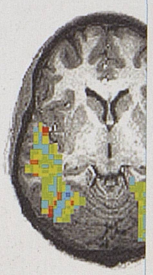
Brown and the MIDAS team are now working on scaling up their ABM model to cover the entire United States, incorporating a population of 300-million agents and requiring from 74 to 300 gigabytes of memory. "As we go to U.S. models and global models," says Brown, "this is very much a capacity application. We need large shared memory, and we're excited about the UV system at PSC."

MORE INFORMATION
www.psc.edu/science/2010/h1n1/



RELYING ON PSC CONSULTING, SOFTWARE AND HARDWARE RESOURCES, A CARNEGIE MELLON COMPUTER MODEL THAT PREDICTS WHAT YOU'RE THINKING GETS EVEN SMARTER

IT KNOWS WHAT YOU'RE THINKING



In his thesis, Rustandi acknowledged PSC scientists Raghu Reddy and Joel Welling. "The thesis research became tremendously more productive," says Rustandi, "in the beginning of 2009 when I finally had working access to Star-P running on Pople at PSC. Thanks to Joel Welling for facilitating access to Pople, and to Raghu Reddy for all his help early on resolving issues with Star-P and for being a liaison to the Star-P developers."

Tom Mitchell
Carnegie Mellon University

Indrayana Rustandi

How do you know what you know? How does your brain, for instance, hold onto the meaning of words, so you can form thoughts? Or — to focus the question more sharply — what parts of your brain become active when you read and think about a word? Which regions of your brain, which neurons, are activated, for instance, when you think about the noun *celery*?

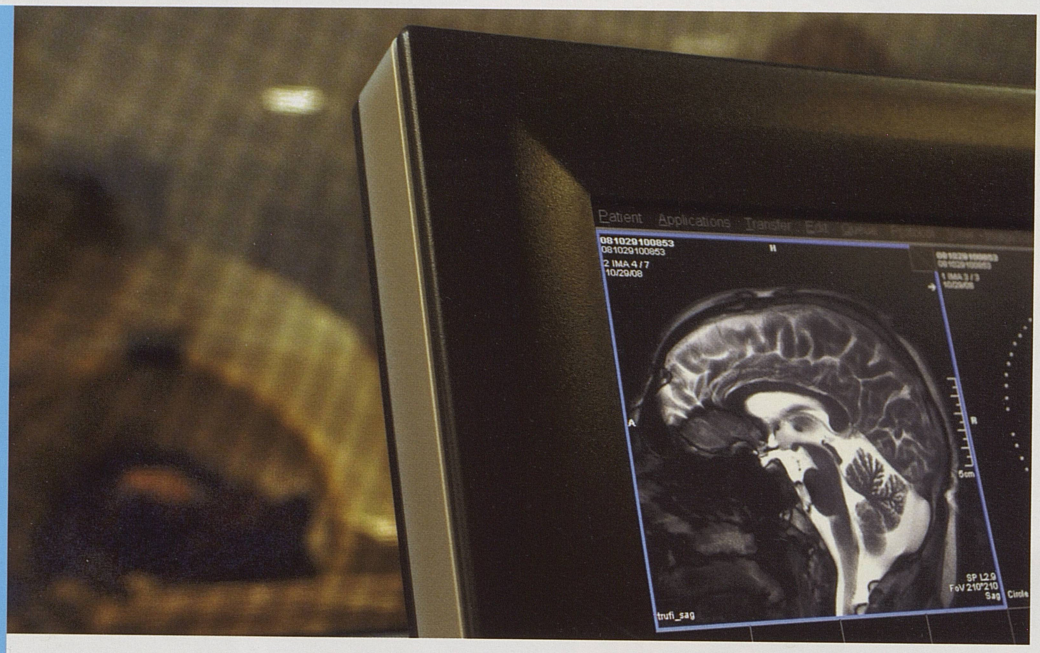
Now try a different noun. *Airplane*. Surely your brain does something at least a little differently than when you thought *celery*. Suppose there were a computer connected to your brain that could detect which brain cells are firing — a very smart, crypto-analytic computer that could tell from the patterns of activated neurons what you're thinking, a computer that can read minds. Science fiction?

Tom Mitchell, who chairs the Department of Machine Learning at Carnegie Mellon University — the first such department in the world — and his collaborators have shown that such a mind-reading computer, though a long reach from current reality, is within the realm of the possible. Their intriguing experiments with functional magnetic resonance imaging (fMRI) and the meaning of nouns — reported in *Science* (May 2008) — showed that a computer model can predict with 77-percent accuracy whether you're thinking *celery* or *airplane*.

Their paper, titled "Predicting Human Brain Activity Associated with the Meanings of Nouns," captured the attention of many scientists, and laypeople too — especially after Mitchell and his CMU colleague Marcel Just demonstrated their model's ability in an episode of "60 Minutes" (Jan. 4, 2009). Since then, the model has become even smarter.

For his Ph.D. thesis supervised by Mitchell and completed this year at CMU, computer scientist Indrayana Rustandi — now working on Wall Street — found a way to improve the model, and it can now predict whether you're thinking *celery* or *airplane* (or other choices between two nouns among a list of 60) with 85-percent accuracy. Working with PSC staff and making effective use of PSC's shared-memory SGI system, Pople, Rustandi designed an algorithm that makes it possible to integrate fMRI datasets, basically images of the brain in action, from different people, each with their own individual brain-activation patterns.

The result: the model is freed from the limitation of using fMRI images from a single individual to predict that individual's responses. "This improvement," says Mitchell, "that Indra was able to make using PSC resources is really important to this line of work."



A person undergoes brain scan in an MRI scanner (left) while his brain activity as he thinks about specific nouns displays on the monitor (right).

DOING THE TWO STEP

How can a computer read minds? As presented in *Science* — prior to Rustandi's work, the model implemented a two-step algorithm, the first step of which derives from research by computational linguists. The statistics of word association show that a word's meaning can be represented as a statistical relationship to words and phrases with which it commonly occurs. A noun like *breakfast*, for instance, often occurs in close association with the verb *eat* and also, less often, with *drink*.

Drawing from this idea, the model used a list of 60 concrete nouns (e.g., *celery*, *airplane*, *telephone*, *screwdriver*), and for each of them searched a huge lexical database (from Google) and gathered data on how each of the nouns correlated with a list of 25 sensory verbs (*see*, *hear*, *listen*, *taste*, *smell*, etc.). From this data, the model encoded each noun as a collection of statistical "semantic features."

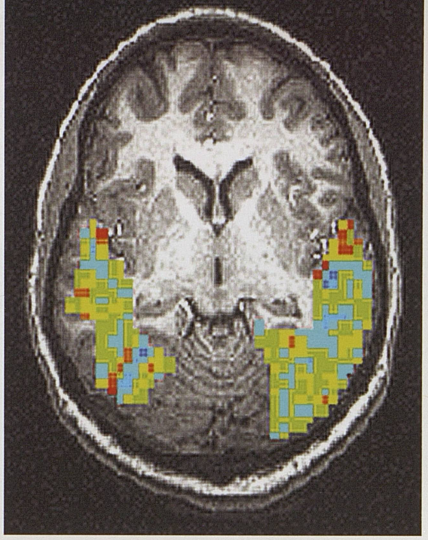
The model's second step relies on image data from fMRI experiments. For the results reported in *Science*, nine college-age participants viewed each of the 60 nouns, along with an associated picture, while in an MRI scanner. The researchers repeated this process six times for each participant, with the nouns in random order. For each participant, then, the model found a statistical mean fMRI image for each of the 60 nouns — building a separate dataset for each participant.

Using these fMRI datasets, the model then trains itself to predict an fMRI image associated with these semantic features. How do we know if the training worked? If the model has learned? The test, explains Mitchell, is to train it on 58 of the nouns, matching semantic features with fMRI activation patterns, then present it with fMRI images from the other two nouns in the list and have it decide which image — from nouns it doesn't yet know — goes with which noun. "It got that right 77-percent of the time in the original publication, and with some optimizations we got it up to 79-percent. That means the model is predicting something correctly about the neural activity in the brain, though not perfectly, and that's where we were."

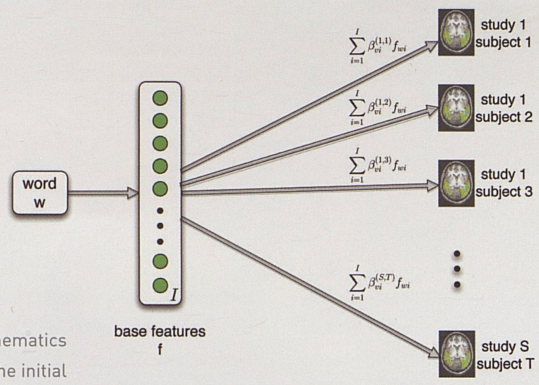
Looking for ways to improve the model, Mitchell and his colleagues wanted to overcome the limitation of separate datasets for each participant. "Because we use this kind of algorithm where you train it on data, we felt like the limiting factor was that we were data starved. The model is predicting something that involves 20,000 different locations in your brain, but we had only 60 words of data from each person." One possible direction was to combine fMRI datasets from different people, to give the model more information from which to train itself.

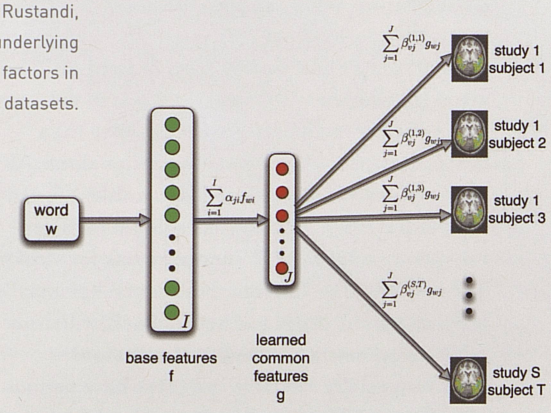
ONE BIG BRAIN

In 2009, Rustandi began working with PSC staff to get the model to run on PSC's Pople, which allowed the entire dataset — expanded to 20 human subjects — to reside in memory simultaneously, a significant advantage. The model, however, was written in MatLab, a serial processing software environment. To gain the benefit of parallelism and improved performance it can make possible, Rustandi turned to Star-P, proprietary software for which PSC holds a license, that allows a MatLab program to run on a parallel system such as Pople.



Images from fMRI show brain activation (colored dots) overlaid on a transverse slice of the corresponding structural brain image (grayscale).





These two schematics compare the initial predictive model (top) with the CCA approach developed by Rustandi, which finds underlying common factors in different fMRI datasets.

With computational tools in place, Rustandi explored various approaches to integrating datasets. A major problem he faced is that, like fingerprints, no two brains are alike. The simplistic approach is to try to register all the subjects' brains within a common spatial framework and pool all the data as if it were all from one brain. Rustandi tried several such approaches and achieved no significant improvement in prediction accuracy. "A major challenge in doing predictive analysis with fMRI data from multiple subjects and multiple studies," says Rustandi, "is having a model that can effectively account for variations among subjects' brains."

His breakthrough came with "canonical correlation analysis" (CCA), a statistical formulation to find combinations among multiple sets of variables that have maximum correlation with each other. "CCA isn't a new method," says Rustandi, "but the application to fMRI hasn't been widespread." Unlike the spatial-registration approaches, CCA has the advantage that it integrates the datasets without disturbing the essential distinctness of the data for each subject. The data doesn't have to fit into a mold.

CCA in effect, explains Mitchell, looks at the different datasets and mathematically replaces the semantic features derived from the 25 verbs with a combination of intensities computed from the fMRI images, and these replacement features are maximally correlated across the 20 brains. "It's a way of framing the question we were really interested in, which is 'What features can represent word meaning across all the fMRI datasets and are the best set of features possible in how they map to the fMRI data?' This algorithm that Indra developed based on CCA is a very nice way to solve that, but it's computationally intensive, and we needed PSC."

THE IDEA THAT THERE'S COMMONALITY IN HOW OUR BRAINS REPRESENT SIMILAR THOUGHTS IS IMPORTANT FOR UNDERSTANDING HOW THE BRAIN WORKS.

The improved accuracy with the CCA integrated dataset is substantial, adds Mitchell. "In terms of error, it's reduced from 21-percent to 15-percent." Rustandi removed a third of the room left for improving the model. In future work, Mitchell and his colleagues plan to apply the CCA method to integrate fMRI data from different studies, even when datasets, for instance, aren't from an identical set of nouns, thus further allowing fMRI studies to break free from limitations of sparse data relative to the complexity of phenomena studied, the human brain.

The idea that there's commonality among different people in how our brains represent similar thoughts, which the model tends to confirm, is important, Mitchell believes, not only for demonstrating what a computer model can do but also, more importantly, for understanding how the brain works. "It bodes well," he says, "for the feasibility of developing a real theory of how the brain represents things. If there's something in common, we can aspire to develop a unifying theory."

MORE INFORMATION

www.psc.edu/science/**2010/brainactivity/**

OFFB

Like a big kick in the chest — that's how people who have undergone defibrillation while conscious often describe it. Applied with "paddles" — as depicted in TV shows and movies — or by EMTs with AEDs (automated external defibrillators) or by implantable devices, defibrillation applies a brief burst of electrical current at energy levels many times higher than the human pain threshold.

"It's a huge shock, 130 to 360 joules externally, about seven internally," says computational scientist Elizabeth Cherry (as of this year at the Rochester Institute of Technology, formerly at Cornell). "Basically," says Cherry, whose work has focused on the heart's complex electrophysiology, "conventional defibrillation is for emergencies, most often ventricular fibrillation, which is almost always life-threatening and requires immediate resuscitation."

Among many heartbeat irregularities, from the occasional skipped beat to various flutters and throbs, ventricular fibrillation (VF) is in a class by itself for rapid fatal consequences. Other cardiac arrhythmias, nevertheless, present serious health problems that could be treated effectively with defibrillation — if it didn't involve serious pain and risk of lasting tissue damage.

With the goal of making that possible, Cherry and a large group of collaborators, including physicists Flavio Fenton (Cornell) and Jean Bragard (University of Navarra, Spain) and others at Cornell and the Max Planck Institute (Göttingen, Germany), have done experiments — correlated with simulations at PSC — that show the feasibility of defibrillation with much lower energy, less than one joule. Called *far-field antifibrillation pacing* (FFAP), their method appears to offer a realistic prospect for defibrillation below the pain threshold.

Many of their FFAP studies have focused on atrial fibrillation (AF), an arrhythmia in the heart's upper chambers. Affecting more than 2.2 million people in the United States alone, AF incidence increases with age and, unlike VF, can persist for years. Although AF seldom requires emergency treatment, it increases risk for stroke and heart failure. Current treatments — primarily antiarrhythmic drugs and sedated defibrillation in a clinical setting (called cardioversion) — aren't especially effective and often have serious side effects. "There's a big clinical hole," says Cherry, "in how to treat this type of disease."

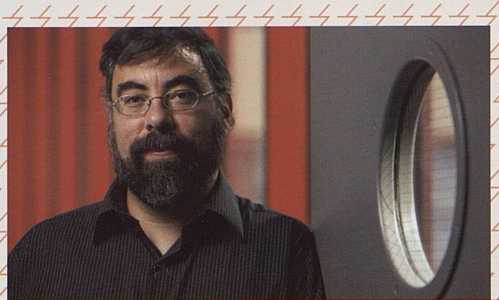
Their experiments with FFAP have shown a success rate comparable to conventional defibrillation, higher than 90 percent, in stopping AF and restoring normal heartbeat. Their computational simulations (using up to 500 processors on PSC's BigBen) confirm the theory underlying FFAP and allow the researchers to test many possibilities. They reported their findings last year in *Circulation* (August 2009), a leading cardiology journal, and newer publications are in press. "Based on human pain threshold values in the literature," says Cherry, "we believe our approach is currently at the threshold and that with optimization of parameters such as shock waveform and electrode placement, guided by computer simulation, we'll be below it."

BEAT hearts

A NOVEL LOW-ENERGY METHOD OF DEFIBRILLATION SHOWS PROMISE TO SUCCESSFULLY TREAT HEART ARRHYTHMIAS WITH LESS PAIN AND TISSUE DAMAGE THAN CURRENT DEFIBRILLATION METHODS

Flavio Fenton
Cornell University

Elizabeth Cherry
Rochester Institute of Technology



SCROLL WAVES IN A FAR FIELD

The steady lub-dub we call the heartbeat is regulated by a built-in pacemaker — the sinoatrial node. This cluster of cells at the top of the right atrium emits electrical pulses about once a second that excite nearby cells, generating a cell-to-cell electrical wave that in milliseconds becomes the smooth and concerted muscular contraction that pumps blood.

Arrhythmias occur as these propagating waves break apart and form "reentrant waves" — in effect, a mistimed pulse that loops back through the heart's muscle fibers like a dog chasing its tail. Called "spiral waves" or in three-dimensions "scroll waves," they meander through the heart and play havoc with its normal rhythmic pattern. VF is believed to involve multiple scroll waves, and AF may or may not require multiple waves. "A single wave might be tachycardia or flutter in the atria," says Cherry, "or it might be fibrillation."

Before the work of Cherry and her collaborators, research on low-energy defibrillation focused on a single spiral or scroll wave, investigating how a low-energy field could move the wave to the edge of tissue, where it would disappear (since with this kind of wave there's no reflection at edges). "People were looking at what we call 'pinned waves,'" says Cherry, "such as when there's scar tissue from a heart attack. A wave can confine itself within that tissue and circulate repeatedly."

Cherry and her colleagues took the low-energy approach much farther, starting from the same basic "far field" idea. "You have field electrodes," she explains, "so you're not injecting a current directly into the tissue, but letting the tissue respond to changes in the field around it." With a series of experiments, they are investigating how the low-energy field affects more complicated arrhythmias, beginning with AF.

Their experiments, reported in *Circulation*, showed that a series of pulses at low field strength (with an average energy of 0.24 joules) extinguished AF and restored normal rhythm in 69 of 74 trials, a 93-percent success rate. The energy required was 13-percent of that needed for cardioversion, and is below the human pain threshold of about 0.4 joules. Their newer work in a more clinical setting is consistent with these findings.

HOW IT WORKS

With computational simulations, Cherry and her co-workers corroborated their experimental findings and gained clearer understanding of how FFAP works. "With experiments," says Cherry, "we use dyes that bind to the tissue and fluoresce in proportion to changes in membrane potential, but you still can't see everything you want to see. With modeling and simulation studies, you can vary parameters and elucidate the changes systematically, which is hard to do in experiments."

COMPUTATIONAL SIMULATIONS CORROBORATED EXPERIMENTAL FINDINGS AND ENABLED BETTER UNDERSTANDING OF HOW FFAP WORKS.

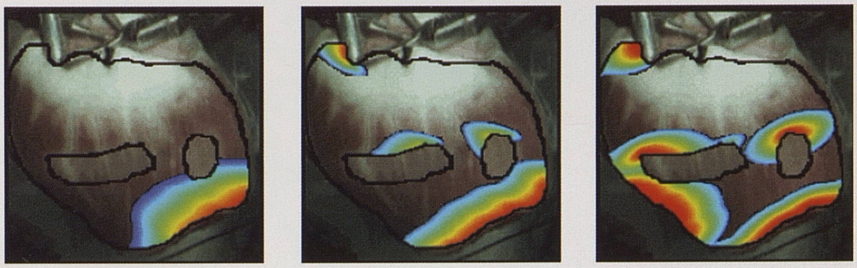
Their computational models, developed over a period of years, have allowed them to do basic research on how arrhythmias get started and evolve, both as a function of the electrical properties of heart cells and larger-scale heart anatomy. Because of the great range of scales involved — spatially (from a single cell to the full-size heart) and over time (microseconds to minutes) — these simulations require large amounts of computing. For their whole-organ model of cardiac electrical dynamics, they relied on PSC's BigBen until it was decommissioned this year. "This system worked well for us," says Cherry. "It was easy to use and convenient." They are now moving their code to Purdue's Steele system.

Their whole-heart model is well suited to simulating FFAP because of its characteristics as a "bidomain" model. Some models of current flow are "monodomain," a simpler mathematical formulation. "Basically you're representing the membrane potential with no extracellular medium where current can flow," says Cherry, "as opposed to a bidomain model, which represents intercellular space as distinct from extracellular space."

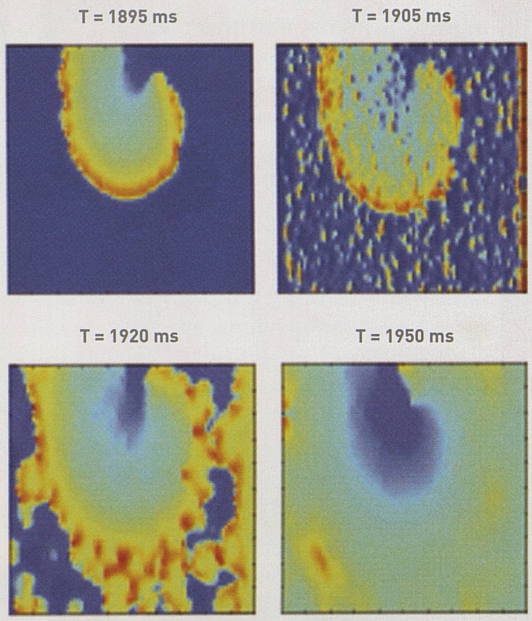
This difference, says Cherry, is not only critical to accurate modeling of FFAP but also is tied to how it works. "The idea of FFAP," says Cherry, "is that you apply a pulsed electric field, and as the current produced by this field encounters discontinuities — blood vessels, for instance, or collagen and other features of the extracellular matrix — it recruits these discontinuities as 'virtual electrodes.'"

These discontinuities — whose electrical effects can't be represented with a monodomain model — function essentially as physically implanted "real" electrodes would. They activate nearby cells, and with successive pulses, the far-field current activates increasingly larger regions of tissue, which synchronizes electrical activity throughout the heart and restores normal rhythm.

The FFAP findings of Cherry and her colleagues demonstrate feasibility and offer the prospect



VIRTUAL ELECTRODES
These graphics from simulation show realistic cardiac tissue with discontinuities created by two inexcitable obstacles at three different field strengths (threshold, low and medium, with color representing time). Increasing the field strength recruits more conductivity discontinuities as virtual electrodes.

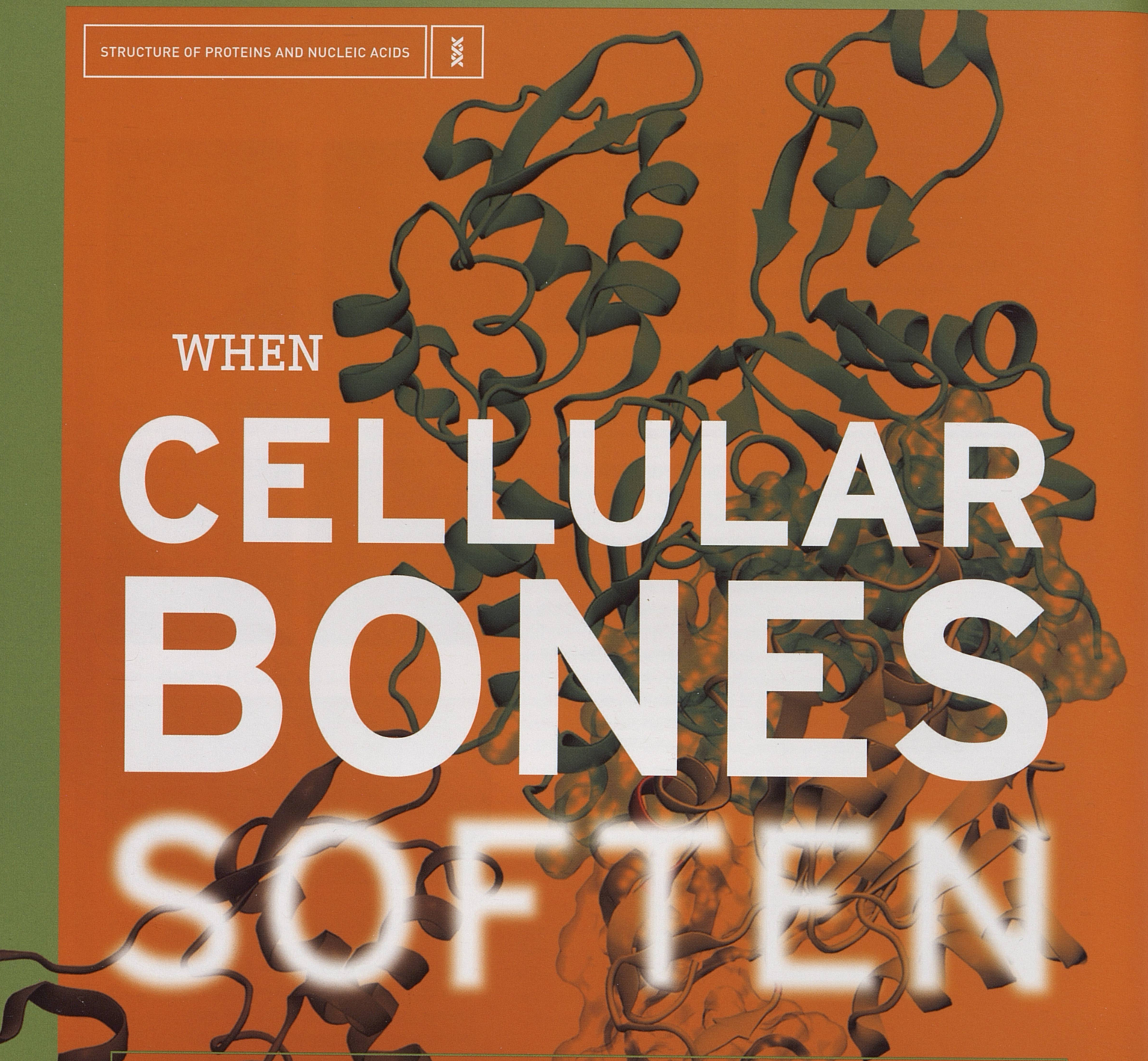


EXTINGUISHING A SPIRAL WAVE
These frames from simulation at four different times (ms = milliseconds, color as membrane potential) show a reentrant spiral wave (1895 ms) in a five-centimeter square sheet of cardiac tissue. The wave is extinguished (1950 ms) by application of a low-energy field. The frame at 1905 ms shows recruitment of "virtual electrodes" as a current pulse is applied and subsequent breakup (1920 ms) of the reentrant wave.

of more effective treatment for AF. Along with antiarrhythmic drugs, which work with varying success, and — less frequently — catheter intervention to isolate a problematic region of atrial tissue, standard therapy for AF is cardioversion. "You can usually restore rhythm, but the more you've had atrial fibrillation, the longer it's been there, the more difficult it is to keep it from coming back."

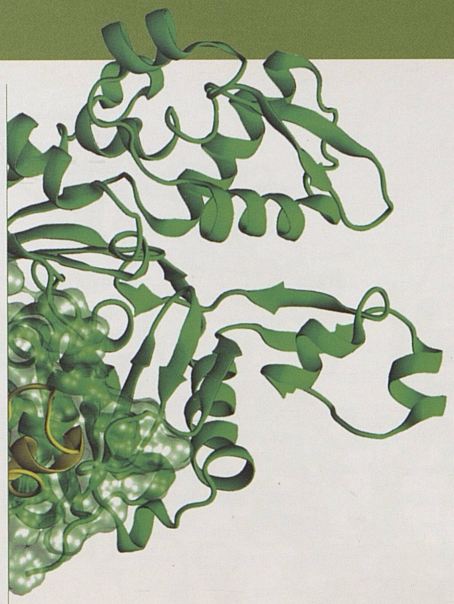
Like cardioversion, FFAP could be applied externally, but also could function as an implantable device, which Cherry envisions as a potential large step forward in AF treatment. "As an implantable, it would be available to work on demand whenever it senses fibrillation, not causing any perceptible pain — even if there might be some mild discomfort. The response would be immediate whenever this pathological rhythm occurs. And the less your heart is experiencing this arrhythmia, the easier it should be to keep it from returning."

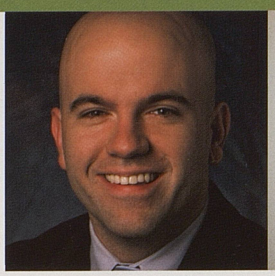
MORE INFORMATION
www.psc.edu/science/2010/offbeatheart/



WHEN CELLULAR BONES SOFTEN

COMPUTATIONAL SIMULATIONS SHOW HOW SLIGHT STRUCTURAL CHANGE IN A UBIQUITOUS PROTEIN TRANSLATES TO DRAMATICALLY ALTERED FLEXIBILITY OF THE CELLULAR SKELETON, FUNDAMENTAL KNOWLEDGE THAT COULD HELP THE SEARCH FOR NEW ANTI-CANCER DRUGS





Jim Pfaendtner
University of Washington

Gregory Voth
University of Chicago

Actin — so named because it *activates* muscle cells — is the most abundant protein by mass in the body. Basically, where there's life there's actin. In evolutionary terms, it's highly conserved — present to some degree in the most simple to complex species of "eukaryotic" organisms — any living thing made of cells with a nucleus.

Perhaps foremost among actin's important functions is structural engineering. As bones are to the body, actin is to cells. Just beneath the cellular membrane, actin forms networks of filaments that are a primary component of the *cytoskeleton*, an undergirding molecular structure that maintains cellular shape, protects the cell and — with its ability to dynamically extend and contract — enables cellular motion.

"One of the most important things in cell biology is the behavior of the cytoskeleton and what proteins do to drive this behavior," says University of Chicago chemist Greg Voth. "We've learned a lot over the years about actin, how it forms polymers and then changes its flexibility properties and depolymerizes. This cycle is what cells use to create motion, change shape and to do many other things, and it's at the core of cellular behavior."

The fundamental processes by which individual actins link with each other (polymerize) to form filaments and, in turn, break apart (depolymerize) have been found to be involved in various cancers, in particular breast cancer, and better understanding could point the way to new cancer therapies. "There's a lot of interest in cancer that goes along with this kind of fundamental cellular biology," says Voth. "People would like to find drugs that target actin and certain actin binding proteins in cancer cells. Our work isn't to design drugs, but you can't design drugs of this kind until you understand these fundamental steps."

Over the past few years, with an extensive series of computational simulations, Voth and biomolecular engineer Jim Pfaendtner of the University of Washington have uncovered previously unknown details of these fundamental processes. Using TeraGrid resources at Texas, Tennessee and Pittsburgh, they first looked at how one small part of actin's structure changes during metabolism — a question raised by laboratory experiments. Their results — reported in *Proceedings of the National Academies of Sciences* (August 2009) — clarified this question, and pointed the way toward further work.

They followed with very large-scale simulations involving multiple actins combined into filaments, looking at complicated interactions between actin filaments and other biomolecules. Their findings from this work —*Journal of Molecular Biology* (February 2010) and *PNAS* (April 2010) — show that these interactions, which involve the same small part of actin they first looked at, radically change the properties of the filaments, making them structurally more flexible — in effect softer — initiating the structural change that leads to the breaking apart of the filaments.

"We've been able to show," says Voth, "how structural changes of the protein at a relatively small scale, things that you can't see in typical structural biology experiments, translate at a larger scale to global changes in the flexibility of the actin filaments that can lead them to sever and break into pieces. We've been able to use these big calculations to reveal at the molecular level what's giving rise to some of the experimentally observed properties."

COLLABORATION MAKES THINGS HAPPEN

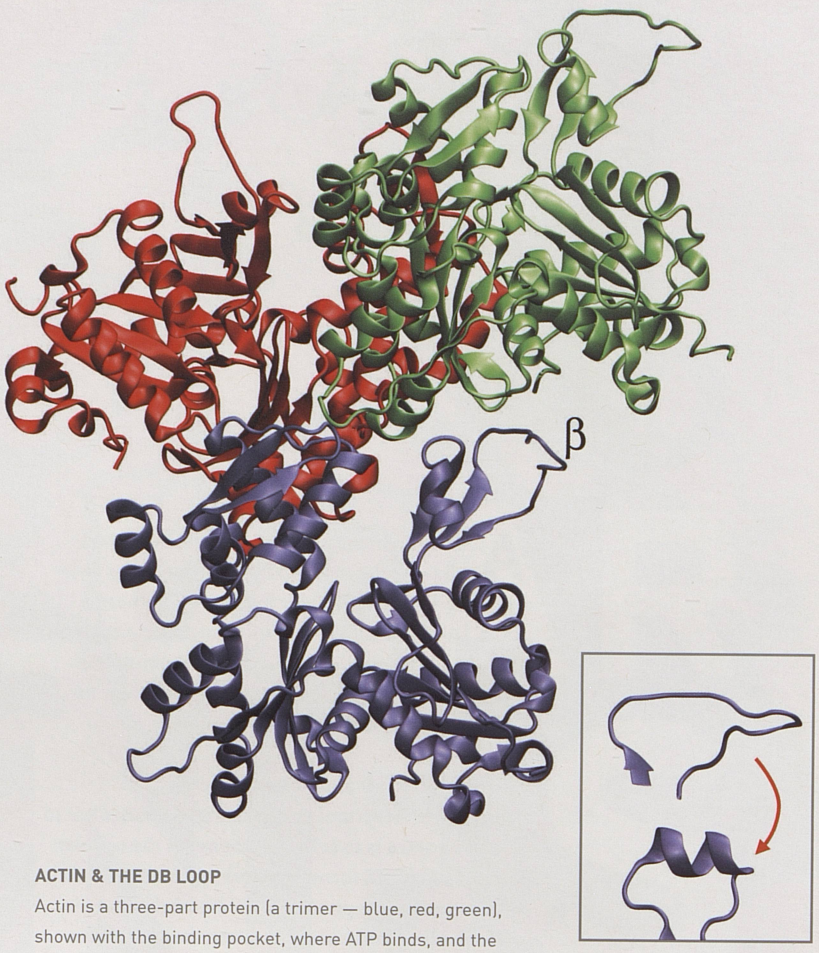
In all this work, Pfaendtner and Voth collaborated closely with molecular biologists — Thomas Pollard and Enrique De La Cruz at Yale — who have done pioneering experimental studies of actin. "Our computational work," says Pfaendtner, "complements experiment, which often offers hypotheses that are based on static pictures, sometimes of low resolution. With our simulations, we can look at the movement of individual atoms and say, 'Yes, this is happening, or no, this isn't happening,' and then corroborate our findings with our experimental collaborators."

Following from discussions with Pollard, Voth and Pfaendtner attacked the question of how actin's structure differs when it's bound with ATP (adenosine triphosphate) from when the ATP reacts to form ADP (adenosine diphosphate), a fundamental reaction of cellular metabolism. The question involved one small loop — called the "DB loop," a length of eight amino-acid groups — within actin. Some studies showed that the DB loop, which is disordered in ATP-actin, forms a helix in ADP-actin. Other work didn't show the helix.

From other actin research, it was apparent that the reaction of actin-bound ATP to ADP leads to softening of the actin filaments, and the larger, interesting question was whether a change from disordered to helical structure in the DB loop might be a switch for this overall softening. Voth and Pfaendtner's simulations show that a helical DB loop is the stable, low-energy form for ADP-actin and provide convincing evidence that the helix forms as part of the ATP to ADP reaction.

The main tool for these simulations was NAMD, a molecular dynamics (MD) program developed to run efficiently on massively parallel systems. MD simulations track the movements in time of each atom in a molecule — 5,700 atoms in one actin protein (50,000 atoms including the water environment), and NAMD has been a powerful tool in protein studies. "Development of this code and the associated force-field we use has been a boon to this field," says Pfaendtner. "We do things routinely now that not long ago we couldn't have imagined doing."

For the DB loop simulations, Voth and Pfaendtner used the Ranger system at TACC, where PSC scientist Phil Blood, working through the TeraGrid ASTA program (Advanced Support for TeraGrid Applications) had optimized NAMD to run 20 to 30-percent faster than was previously possible, using up to 2,000 processors simultaneously. Even with NAMD and Ranger, however, it wouldn't have been feasible to get useful results, because the folding of the DB loop into a helix occurs over a microsecond or longer of biological time, too much, even with the most powerful tools, to fully simulate with MD.



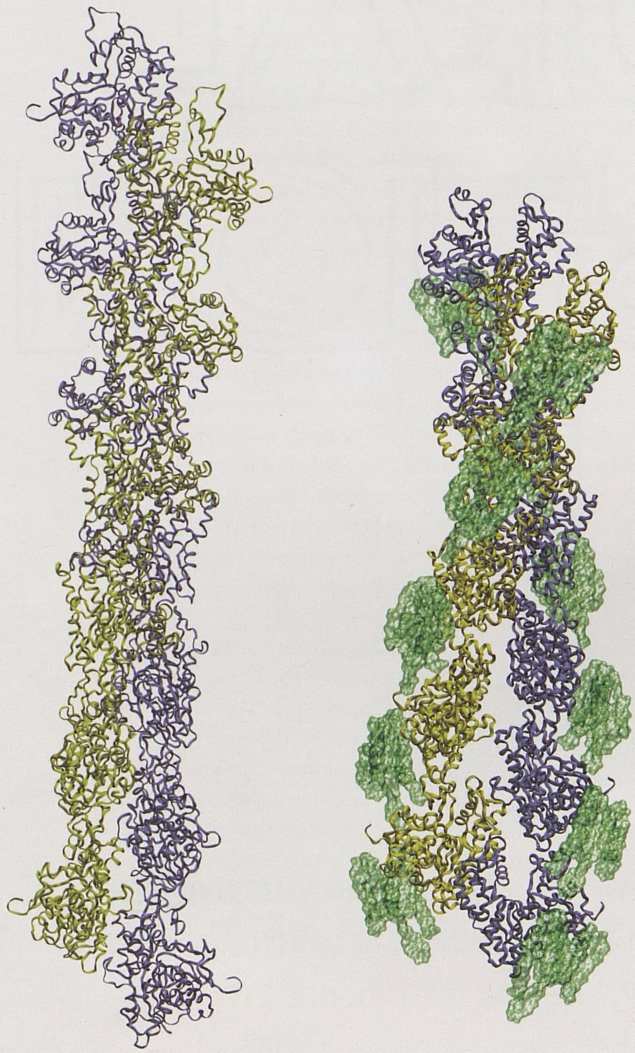
ACTIN & THE DB LOOP
Actin is a three-part protein (a trimer — blue, red, green), shown with the binding pocket, where ATP binds, and the DB loop (labeled β). The inset shows the DB loop in both its disordered and helical form.

Pfaendtner addressed this problem with an innovative approach called "metadynamics" — a powerful, mathematically sophisticated algorithm that is applied, in effect, on top of MD and which makes it possible to arrive at the stable, free-energy structure of a folding event without the complete atom-by-atom MD. Through an NSF-funded international collaboration, Pfaendtner worked in the research group of acclaimed physicist Michele Parrinello, at the Swiss Federal Institute of Technology (ETH Zurich), which developed metadynamics.

Through this collaboration, Pfaendtner applied metadynamics to actin simulation, which enabled the finding that the DB loop's transformation from unstructured to helical is part of the ATP to ADP reaction. "We have this picture now," says Pfaendtner. "It had been suggested by experiment, but the experiments were conflicting, and we were able to clarify the details."

The next question, says Pfaendtner, was "What does it mean for the mechanical properties of the

cells that this loop folds?" With Kraken at NICS and BigBen at PSC, Pfaendtner and Voth built on groundbreaking 2005 work by Voth and Jhih-Wei Chu (University of California, Berkeley) and simulated 13 actin proteins linked as an actin filament, about 500,000 atoms, for a total of about 500 nanoseconds, very large simulations. With these studies, which modeled actin filaments in several different configurations, the researchers interpreted experimental data showing that a result of the ATP to ADP reaction is softening of the filament. Their data, furthermore, taken together with the prior simulations, indicate that folding of the DB loop causes the filament to reduce its "persistence length," in effect a softening of the filament.



ACTIN FILAMENTS

Actin filaments in standard form (left) form a double helix (yellow/blue strands). When ADF (green) binds with actin (right), the structure changes.

The next round of simulations, this time with BigBen at PSC, included a protein called cofilin or ADF (actin depolymerization factor) for short. "We were again looking at actin interactions," says Pfaendtner, "but trying to understand at the molecular level 'How do these filaments actually fall apart?'" Experiments had shown filament softening related to binding with ADF, but the how and why of it weren't understood. Once again, with very large simulations of 13-actin filaments, Voth and Pfaendtner added new detail to the picture of the actin polymerization-depolymerization cycle.

PERSPECTIVE GOING FORWARD

Their results show that ADF-binding changes the orientation of the DB loop, moving it away from the central axis of the filaments, which in effect, loosens contacts between adjacent actins, hence an overall softening. "TeraGrid has enabled this experiment-based collaboration," says Pfaendtner, "and through it we've seen the experimental community becoming enthusiastic about what we can do with simulations."

IT REMAINS TO BE SEEN TO WHAT EXTENT THE DB LOOP PRESENTS A TARGET FOR DRUG THERAPIES AIMED AT STOPPING CANCER.

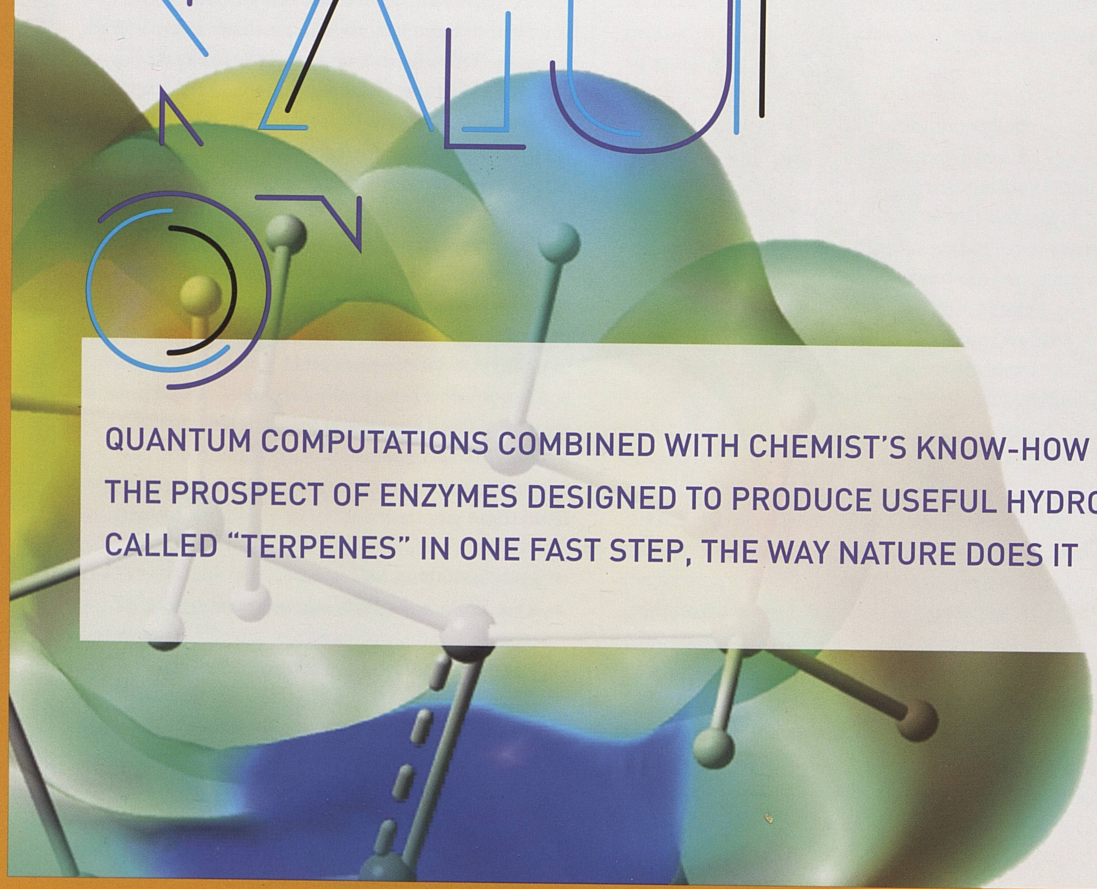
Through several years of synergy between experiments and simulation, this work has arrived at new understanding about a small structure within actin, the DB loop, as a major player in regulating cytoskeletal dynamics and structure. It remains to be seen to what extent this loop may present a target for drug therapies aimed at stopping cancer, which are, in effect, cells whose growth process has gone out-of-control.

"Ultimately," says Voth, "we want to develop multi-scale models and take all these molecular properties and put them into simpler models to understand these phenomena at the cellular scale, but we've been really surprised to find how relatively small changes at the molecular scale propagate to properties of the cytoskeleton. This is very fundamental knowledge that we're getting from these simulations."

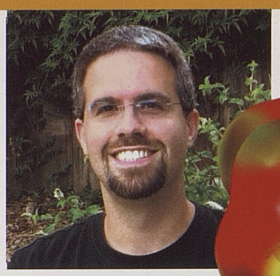
MORE INFORMATION

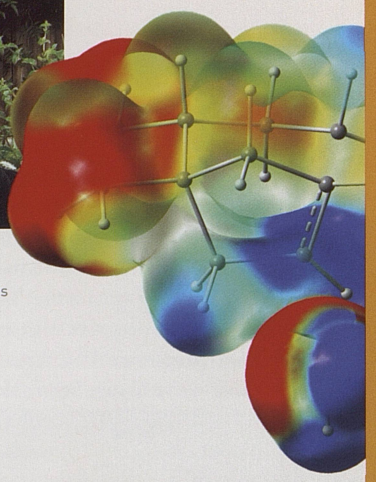
www.psc.edu/science/**2010/cellularbones/**

NATURAL REACTION



QUANTUM COMPUTATIONS COMBINED WITH CHEMIST'S KNOW-HOW OFFER THE PROSPECT OF ENZYMES DESIGNED TO PRODUCE USEFUL HYDROCARBONS CALLED "TERPENES" IN ONE FAST STEP, THE WAY NATURE DOES IT





Dean Tantillo
University of California, Davis

Parsley, sage, rosemary and thyme — four spices from an old song. Although the anonymous lyricist from hundreds of years ago wasn't an organic chemist, it happens that each of these spices is related to a large class of hydrocarbon compounds called *terpenes*. The essential oils of violets and roses, peppermint, pine trees (hence "turpentine"), eucalyptus, oranges, and also frankincense and myrrh — all of them contain terpenes that are among thousands found in nature, many entering our lives as pungent fragrances and flavors.

"Most terpenes are plant-derived and have interesting, pleasant flavors or smells," says University of California, Davis chemist Dean Tantillo, "but they have activities all over the map. Some have the ability to fight cancer."

The terpene taxol, a widely used anti-cancer drug, effective because it inhibits cell division, exemplifies Tantillo's interest in these compounds. Taxol comes from the bark of the Pacific Yew tree, and when chemists first isolated it in 1967, they faced huge obstacles in making it available as a pharmaceutical. Early efforts required 1,200 kilograms of bark to produce 10 grams of taxol. Estimates by the mid-1980s were that it would take 360,000 yew trees a year to meet the anticipated need for ovarian cancer alone, an amount impossible to sustain. Creative chemists eventually sidestepped this problem by developing ways to make taxol from available small molecules or a precursor chemical in yew-tree needles. The latter approach exploits nature's ability to produce taxol's complex core architecture, but the former involves a long sequence of chemical steps, and for most complex terpenes, chemists have been forced to follow this difficult approach.

In nature, explains Tantillo, the core of a complicated terpene is assembled by an enzyme catalyst in one step. "If you could harness the enzyme to do it for you, synthesizing a complex terpene would be much more efficient." Tantillo's research focuses on understanding these reactions, with the aim of unraveling a central mystery: Nature uses hundreds of different enzymes to produce hundreds of different terpenes, but all from only a handful of starting compounds. How does a given enzyme make an abundance of only one product when hundreds are possible?

"It's rare that biology uses the same compound to make hundreds of different complex things so efficiently," says Tantillo. "If we can understand why one enzyme is producing one terpene and a different enzyme is producing a different terpene, we should be able to apply that knowledge to rationally redesign the enzymes. The goal would be to take a readily available enzyme that makes one complicated organic molecule that perhaps has an interesting structure but no useful biological activity, understand how that enzyme works, and change it in ways that will cause it to produce a different molecule that is highly valued."

IT'S RARE THAT BIOLOGY USES THE SAME COMPOUND TO MAKE HUNDREDS OF DIFFERENT COMPLEX THINGS SO EFFICIENTLY.

With demanding quantum computations on Pople, PSC's SGI Altix system, Tantillo has explored the detailed, atomic-level chemistry of terpenes. His recent papers from this work — *Nature Chemistry* (August 2009) and *JACS* (March 2010) — report unanticipated findings that alter the standard view of terpene-forming reactions and point toward new understanding of the mysterious chemistry at play in producing terpenes in nature.

PENCIL, PAPER & SUPERCOMPUTER

Terpenes are derived biosynthetically from units of isoprene, a biological building block that consists of five carbons — four in a row and one on the side — and eight hydrogen atoms, C_5H_8. The basic formula of terpenes is to multiply that, $(C_5H_8)_n$ — where n is the number of linked isoprene units.

Much of Tantillo's recent work has focused on "sesquiterpenes," a large family of terpenes formed from three isoprenes, 15 carbon and 24 hydrogen atoms, rearranged through reactions into complex configurations of carbon rings, fused together or linked by straight hydrocarbon chains. (Tantillo has also examined taxadiene, a diterpene, 20 carbons and 32 hydrogens, that's a precursor to taxol.) The starting compound for this work is farnesyl diphosphate (FPP), a straight chain of three isoprenes ending with a diphosphate group, from which nature constructs hundreds of different sesquiterpenes.

For his computational modeling of sesquiterpene reaction pathways, Tantillo relies on GAUSSIAN03,

software developed originally by Carnegie Mellon University chemist John Pople, who won the 1998 Nobel Prize in Chemistry for this work — and who, coincidentally, is the namesake of PSC's supercomputing system that Tantillo's team uses, usually eight processors at a time. "Pople has been the best computer at PSC for us," says Tantillo, "because we've always been able to access it without a long wait. Also, it has many processors [384 dual-core], and although we don't use all that many at a time, we sometimes are running tens of simultaneous calculations with eight processors each, so we need access to a lot of computing power."

Within the framework of GAUSSIAN03, Tantillo relies on density-functional theory (B3LYP and other methods), a much-used approach that makes it possible to resolve with reasonable accuracy quantum calculations that would otherwise be unsolvable on a useful timescale. Tantillo and his co-workers first write the structures of FPP and the desired sesquiterpene end-product on paper, and then make educated guesses at the intermediate structures and transition states that connect them along the reaction pathway. Then, essentially, they build models of these structures in the computer and turn to GAUSSIAN03 to do the quantum heavy lifting.

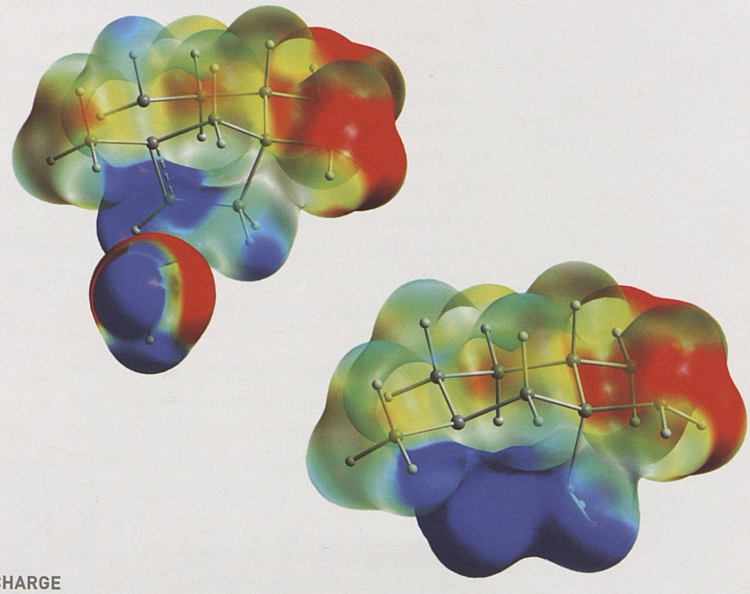
"You have to be a good chemist to do this kind of work — knowledgeable and logical, but also open-minded," says Tantillo. "You start with a carefully chosen guess at each structure — based on your experience and trained intuition. Then you use the quantum mechanics calculations to refine that. Sometimes, though, the results push you in unexpected directions."

For each proposed structure, the software accounts for, among other things, attraction between electrons and nucleus within an atom and repulsive forces between electrons and between atomic nuclei of the molecule. After what can be hundreds of iterations, it arrives at a low-energy structure and provides the geometry and energy to the researcher.

"In general," says Tantillo, "to find a transition state structure can take from hours to weeks, usually on the order of days. So in a five-step pathway, multiply that by five. From start to finish, a project that leads to one of our papers usually takes months."

CHANGES IN SURPRISING PLACES

Ultimately, an intrinsic reaction coordinate (IRC) plot of the energy versus molecular structure for the entire pathway from reactant to product is constructed. The IRC plot maps out the topography of the energy surface, like a mountain range with transition states at different peaks, and the reactant, intermediates, and product molecules in valleys — minimum energy regions — down the mountainsides.



CHARGE
A comparison between the transition state in a terpene-forming reaction (right) and a stable molecule meant to mimic it. The blue parts of the surfaces are the most positively charged.

Their calculations with FPP reacting to form various sequiterpenes often showed that the reaction path avoided expected intermediate steps. "The IRC calculations are super important for us," says Tantillo. "One of the things we discovered is that when you write out a reasonable mechanistic proposal it might have, say, six steps, but we often find that two or three of the steps that any self-respecting chemist would write down actually merge into one step. That was unexpected."

This unexpected finding led them to some careful double-checking. Following a pathway down from a transition state to the intermediate to which it is directly connected using the IRC procedure, they are able to show convincingly that they have not missed intermediate molecules. They were also surprised to find, however, that changes to molecular structure that in most reactions would take place simultaneously were, for the sesquiterpene reactions, spread out along the reaction path between a transition state and the intermediates connected to it.

"Let's say you have three events combining into one," says Tantillo. "As the transition state is approached, you might see the geometry of the molecule changing in a way that corresponds to only one of those events, like the formation of a new carbon-carbon bond. But as you walk down

TANTILLO'S TERPENE TEAM

Young Hong
Pradeep Gutta
Michael Lodewyk
Selina Wang
Dan Willenbring
Greg Ho
Dustin Nouri
Mihaela Bojin

the surface from the transition state to the next intermediate, you see the geometry changing in different ways that correspond to the other two events. A lot of the interesting stuff that is happening — the bonds that are being made or broken — is often happening not near the intermediate or even near the transition state structure but somewhere in between, and that's part of the energy surface that's often not well characterized."

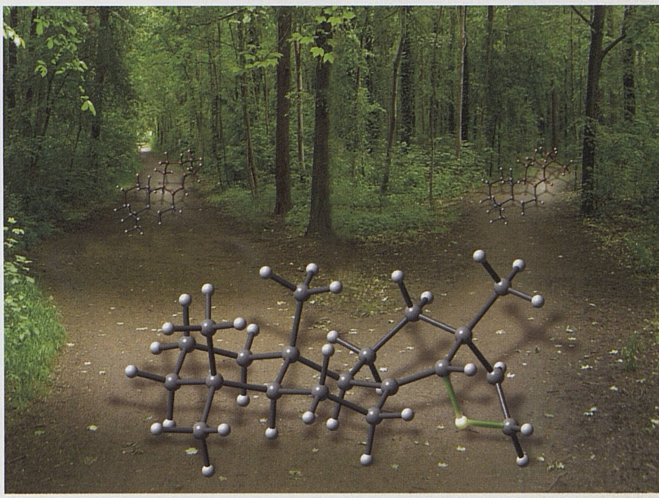
Tantillo believes that this finding might be a key to the main puzzle of terpene chemistry. "The process we have discovered avoids some intermediate structures that would lead to undesirable products," he explains. The relative lack of intermediates, each of which could act as a fork in the road of the reaction pathway, directs the reaction toward only a few products instead of many. "That's the biggest general theme that has come out of this research so far."

THIS FINDING MIGHT BE A KEY TO THE MAIN PUZZLE OF TERPENE CHEMISTRY.

Tantillo envisions producing hydrocarbons with carbon skeletons as complex as that in taxol simply and quickly, with just one enzymatic step, as nature does it. "Our attitude is that if we can answer the question 'Why does one terpene-synthesizing enzyme make only one of the 400 or so possible terpene products,' then we ought to be able to put that knowledge to use and say 'OK, well, let's change these few parts of the enzyme active site, and it ought to produce this different terpene.' Successfully redesigning an enzyme, a very complicated thing to do, would be the ultimate test of our models." (TP)

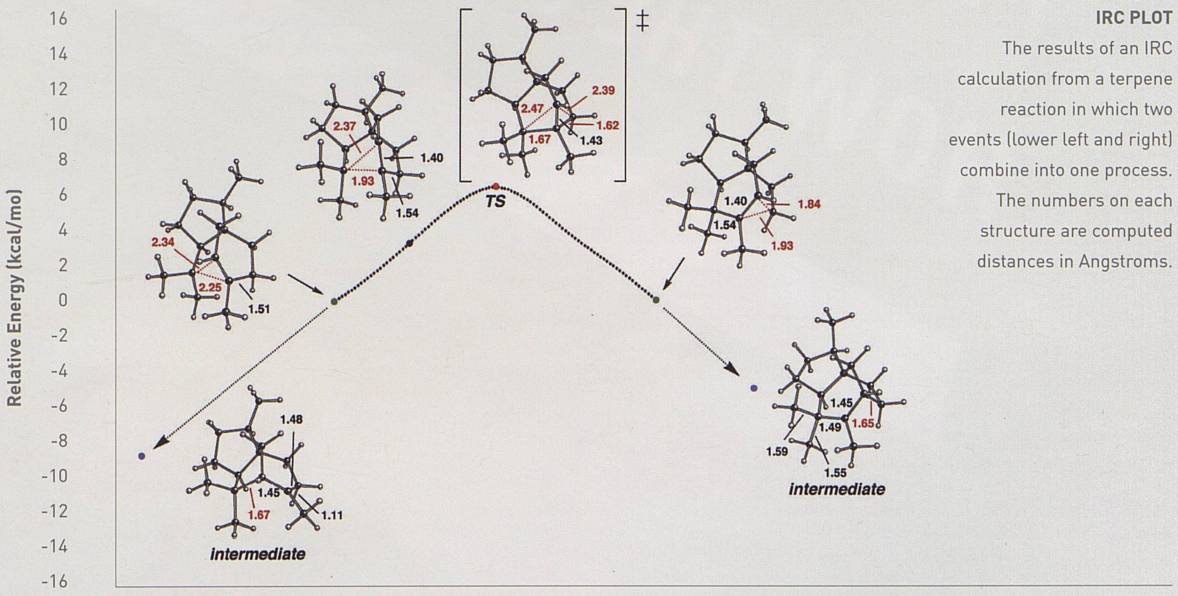
MORE INFORMATION

www.psc.edu/science/**2010/terpene/**



REACTION PATHWAY

Ball-and-stick representations of a transition state and two possible terpenes formed from it at a "fork in the road." The green "sticks" in the transition state correspond to bonds being made or broken.

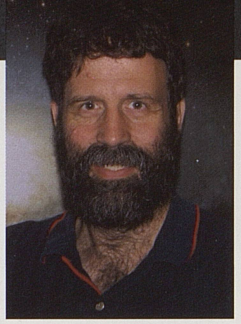


IRC PLOT

The results of an IRC calculation from a terpene reaction in which two events (lower left and right) combine into one process. The numbers on each structure are computed distances in Angstroms.

IN SEARCH OF BULGELESS DWARFS

SIMULATIONS OF DWARF GALAXIES AND INTENSIVE DATA ANALYSIS RESOLVE A MAJOR PROBLEM WITH THE REIGNING THEORY OF HOW THE COSMOS EVOLVES



Fabio Governato
University of Washington

Tom Quinn
University of Washington

Photo: Sunny Facer

Diminutive in size, with about one-percent the number of stars in the Milky Way, dwarf galaxies have posed a big problem for cosmological theory. Astronomers who look at them — many are known to orbit the Milky Way — know they are "bulgeless," with a distribution of stars that's more or less flat on edge, slightly humped like a frisbee. Astrophysicists who run computational simulations that test the reigning "cold dark matter" (CDM) model of how galaxies form, however, have been seeing bulges.

A spherical central bulge of stars is common in larger galaxies — the Milky Way has one, but when CDM simulations of dwarf galaxies have them, as they have for about 15 years, it suggests a problem with the model, which has in nearly every other way shown excellent agreement with observations. Similarly, dwarf galaxies from CDM simulations have a centrally concentrated distribution of dark matter, the invisible matter that comprises the largest part of the universe's mass, again inconsistent with observed dwarfs. "Basically we have a model that's really good at explaining a lot of what's going on in the universe," says University of Washington astrophysicist Fabio Governato, "and then there's these two sore points — bulges and dense dark-matter halos in the dwarf galaxies."

"This failure is potentially catastrophic for the CDM model," wrote Governato and his University of Washington colleague Tom Quinn and an international team of collaborators in their paper — *Nature* (January 2010) — reporting simulations that convincingly resolve this dwarf galaxy problem. A big part of the solution was GASOLINE, astrophysics simulation software developed over a 15-year period at PSC by Quinn, James Wadsley of McMaster University, and Joachim Stadel of the University of Zurich. "We made very good use of BigBen [PSC's Cray XT3 system, decommissioned this year]," says Quinn, "both in developing this code and in earlier simulations."

By improvements to the way GASOLINE represented some of the physics involved and with higher resolution than had before been possible, the researchers more realistically captured the processes of star formation and evolution, including the violent star death and spectacular gas outflow phenomena of supernovae. Totaling more than a million hours of TeraGrid computing (mainly at NICS and TACC), their simulations showed that not only dark matter, but also stars and gas influence the structure of dwarf galaxies. The result is bulgeless dwarf galaxies — with density properties of dark matter and stars that agree well with observed dwarf galaxies.

"Realistic dwarf galaxies are thus shown to be a natural outcome of galaxy formation in the CDM scenario," wrote the researchers. Or as Governato says, "CDM lives to fight another day."

STARBURST IN A DWARF IRREGULAR GALAXY (LEFT)
This image from the Hubble Space Telescope, about 8,000 light-years across, shows a burst of star-forming activity in dwarf galaxy NGC 1569, 11 million light-years away in the constellation Camelopardalis. This turbulent environment is fed by supernova explosions that spew out material and trigger further star formation.

Image courtesy of NASA, ESA, Hubble Heritage (STScI/AURA).

SUPERNOVA OUTFLOWS

"It was a massive computational project," says Governato. "This kind of research wasn't possible just three years ago. We took advantage of the fact that computers are getting faster and faster."

To keep the amount of computing within limits, the researchers simulated a cubic volume of space 25 megaparsecs on edge (about 475 million-trillion miles). With the typical distance between galaxies being one megaparsec (19 trillion miles), this volume gives a representative sample of galaxy formation. GASOLINE has the ability to adapt resolution according to regions where stars and galaxies form, with a "tree code" — which automatically sprouts more "branches" in active areas of the cube.

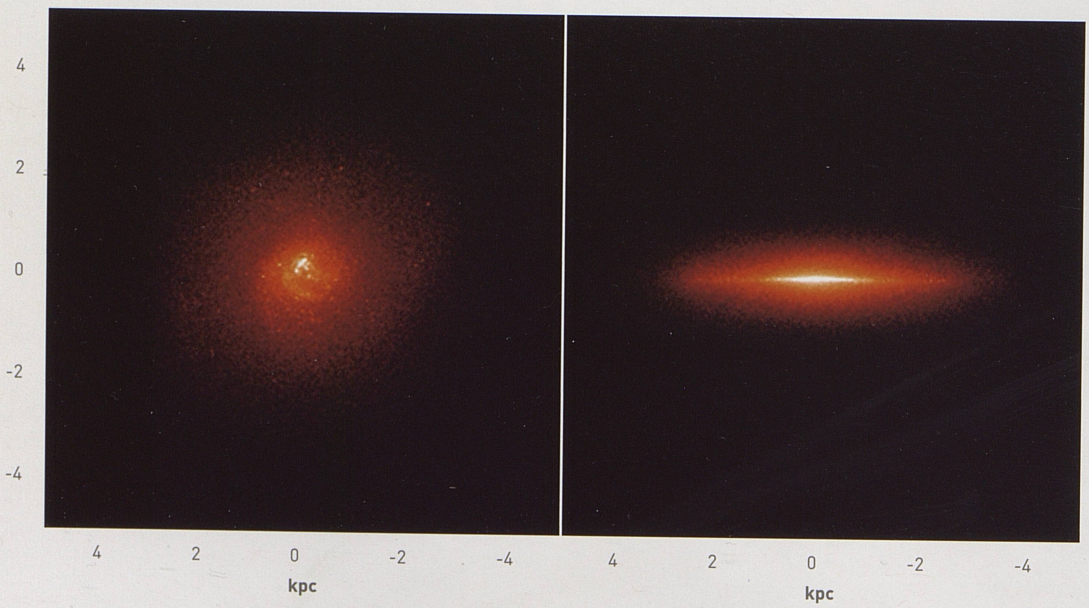
To look at the details of galaxy formation, the team focused on some of the sites within the cube where mass congregated. "We picked galaxy formation sites from a lower resolution simulation," says Quinn, "and then went back and did a higher resolution simulation where we selectively sampled the region where chosen galaxies formed." In these regions, they simulated with resolution up to 100 parsecs — the most detailed picture of galaxy formation to date.

By implementing star birth and supernova explosions, the simulations produced bulgeless dwarf galaxies while adhering to the physics of the CDM model. "The massive stars explode and form supernovae," says Governato, "and these giant stellar explosions have a dramatic effect on the distribution of hydrogen and helium in the galaxy. The star explosions produce what we call 'outflows,' which blow most of the gas away from the center of galaxies. There's not much left to form stars, so the bulge doesn't form. It's not like you remove it — it doesn't form in the first place. This is the key result of our simulations."

Along with bulges, previous models of dwarf galaxies also suffered from what's called the "cusp" problem — high density of dark matter in the central region. The researchers had no reason to expect their simulation would resolve this, but a colleague analyzing their data found that the dark-matter density profile was flat throughout the central volume of the dwarf galaxies. This was initially a puzzle, since dark matter, which responds only to gravity, is unaffected by "wind" from a supernova explosion, and wouldn't blow away in the outflow.

"At that point we started scratching our heads," says Governato. "This is where it gets fun, where the computer gives you a hint but it doesn't come with an explanation." What they eventually realized is that the luminous matter acts as a binding agent for dark matter. "When most of the gas in the central region of the galaxy is blown away, there's nothing left to hold the dark matter; it expands, and the dark matter density near the center decreases."

As a result, the simulations solved two problems of the CDM model: the observed lack of bulges and the flat distribution of dark matter in real dwarf galaxies. "Usually this is a good sign for science," says Governato, "when one result explains more than one thing; a complex system resolves to a simple process."



PROPERTIES OF A SIMULATED GALAXY
These two frames from Governato and colleagues' simulation of a dwarf galaxy at high resolution show light distribution face-on (left) and edge-on.

ARTIFICIAL OBSERVATIONS VIA SHARED MEMORY

To further verify their findings, Quinn, Governato and their colleagues have turned to PSC's shared-memory system, Pople, for "artificial observations" of the dwarf galaxies formed in their simulations. "After we run the simulations, we want to look at them as an observer would," says Quinn. "What would this galaxy look like if you looked with the Hubble space telescope at five different wavelengths, five different colors? To do that we have to trace how the starlight propagates through the dust and the gas of the galaxy."

For this analysis, they use a program called SUNRISE (a Monte Carlo ray-tracing code developed by Patrik Jonsson), which takes the data generated by the simulations and produces images of the light distribution based on the masses and ages of the stars, their chemical compositions, and how they are distributed in the galaxy. The galaxy's light distribution might make it look blue, or red, or ultraviolet.

This kind of data analysis is best suited to a shared-memory computing system. "Photons are flying all over the place," says Quinn. "You can't easily domain decompose and expect reasonable communications between domains. To handle the photons from a single dwarf galaxy simulation requires about 50 gigabytes of memory, and you really need to have the whole simulation in shared memory." The improved processor speed and larger shared memory of the new SGI Altix UV system (p. 4) at PSC will help with the larger datasets and artificial observations produced by this group and others.

BLACKLIGHT, THE NEW SGI ALTIX UV SYSTEM AT PSC, WILL HELP WITH LARGE DATASETS PRODUCED BY THIS GROUP AND OTHERS.

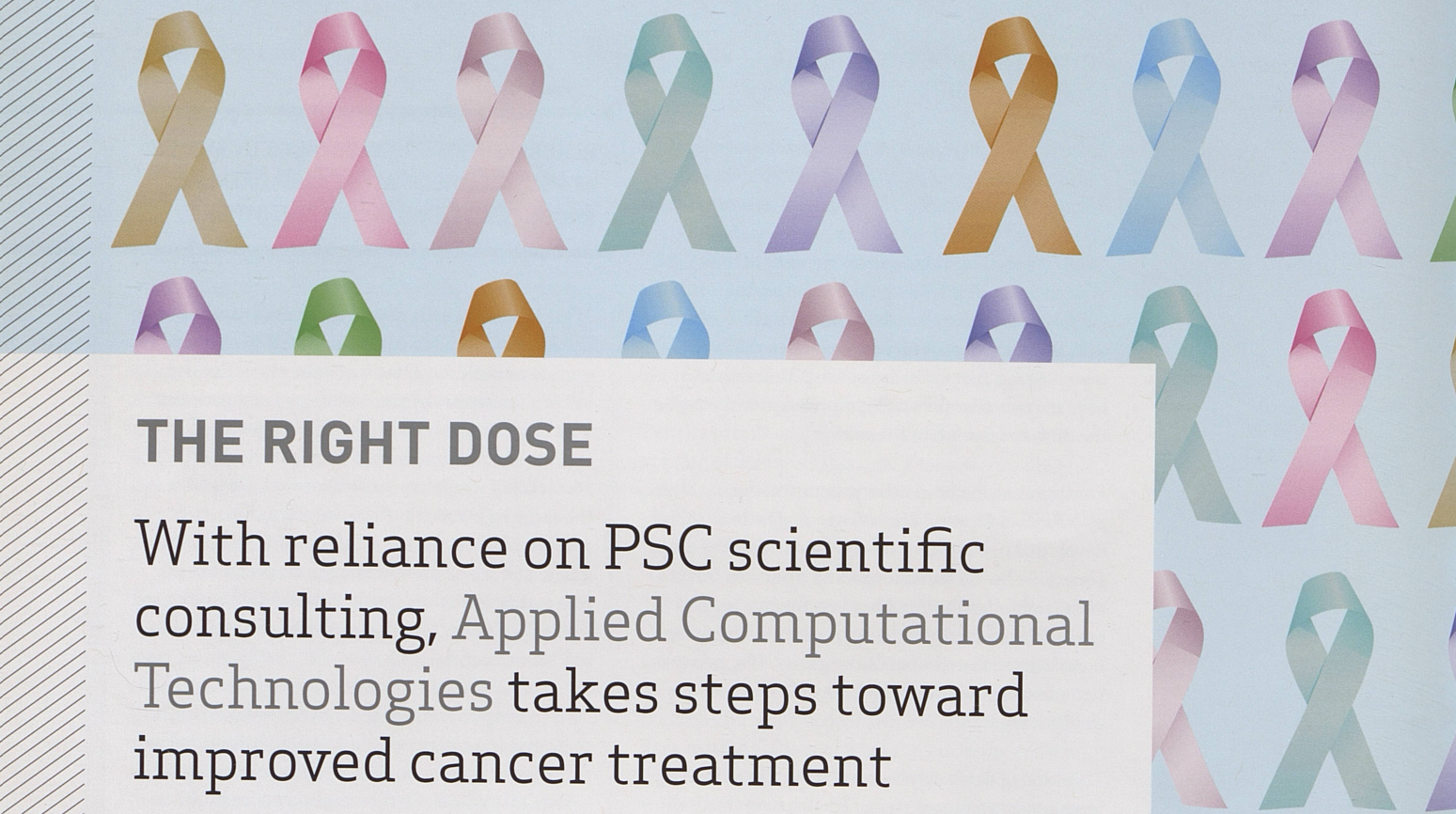
"This is what's really new in our field," says Governato. "You don't just run the simulation, you want to analyze the data so that they look like data out of a telescope, because when you compare data with observers you want to compare apples with apples. We find that our artificial galaxies look remarkably similar to those observed. They have the same light distribution, they form the same amount of stars, the same light profile without the bulge, and based on how fast the stars and the gas move, they have the same amount of dark matter as the real galaxies. And this confirms the results of our Nature paper."

"We've learned two things," he continues. "One is that you can create realistic galaxies within what we think is the correct model for structure formation — the CDM model. So it removes a major problem for this model. And second it shows that we've started to understand what physical processes shape the distribution of mass at the center of galaxies, namely these large gas outflows caused by supernova explosions." (TP)

MORE INFORMATION
www.psc.edu/science/**2010/dwarfgalaxy/**

WHICH GALAXY IS REAL?
A galaxy from Governato and colleagues' simulation (left) appears in all respects identical to a real galaxy (right) and background image from the Sloan Digital Sky Survey Collaboration.

Image courtesy of Chris Brook (The Jeremiah Horrocks Institute at the University of Central Lancashire) and Patrik Jonsson (Center for Astrophysics, Harvard).



THE RIGHT DOSE

With reliance on PSC scientific consulting, Applied Computational Technologies takes steps toward improved cancer treatment

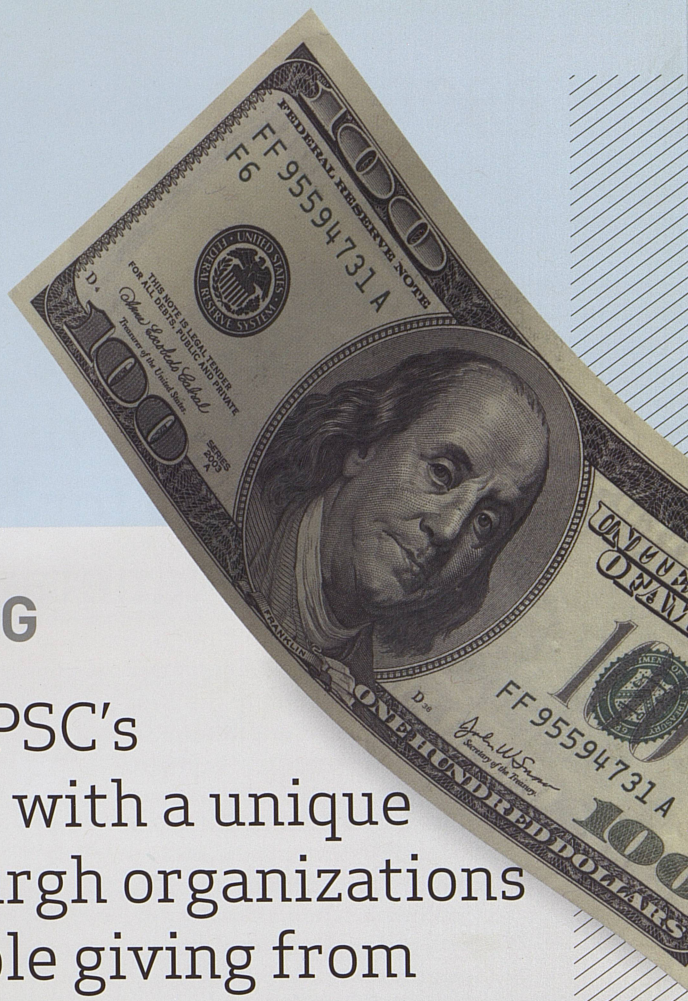
During 2010, about 1.4 million Americans will be diagnosed with cancer and about 65-percent of them will receive radiation therapy. Applied Computational Technologies (ACT), a software development company in Windber, Pennsylvania, works to make radiation therapy better for patients and easier for professionals to make good decisions. They have developed a sophisticated software tool, called ProACTive, to dramatically improve the radiation oncologist's ability to implement the best possible course of radiation treatment.

ProACTive derives from an innovative mathematical approach to the physics of radiation dose calculation, and in recent years, ACT has worked closely with PSC scientists Nick Nystrom and Joel Welling to refine their approach. "We're developing a next generation dose calculator for radiation treatment of cancer," says ACT co-founder and vice-president Jay McClatchey. "It requires a high degree of mathematics and particle physics capability. We need help from very smart people, and PSC has been a great resource."

Current radiation-treatment software is a compromise between speed and accuracy. ProACTive eliminates this trade-off, and in 2010 testing it has calculated accurate treatment doses more than 170 times faster than other software in this field. In July at the American Association of Physicists in Medicine conference in Philadelphia, ACT presented results from rigorous benchmark tests that, says McClatchey, were superior to the top commercial systems and scientific codes.

"The key benefits of improved planning accuracy," says McClatchey, "are local tumor control and reduced complications in normal tissue. This is especially true when the cancer is close to organs. Improved accuracy can enable an increase in the dose prescription and physicians may be more comfortable that normal tissue complications can be kept within acceptable limits."

ACT's goal is to position itself to transfer ProACTive to a radiation treatment system provider. Currently in final stages of prototype testing, ProACTive is expected to dramatically increase the precision of radiation dose calculation, which means more effective cancer therapy and additional lives saved.



THE JOY OF GIVING

Researchers use PSC's BigBen and Pople with a unique dataset of Pittsburgh organizations to model charitable giving from private donors

While it's often true that giving makes you feel good, the "warm glow" of contributing to a worthy cause goes only so far in attracting private donations to charitable and cultural organizations. Far more effective for organizations that depend on wealthy private donors are exclusive, high-profile events, such as dinner parties, that convey social status and offer visibility with other affluent people. That's the main finding from computational modeling using PSC resources by Holger Sieg, J. M. Cohen Term Chair in Economics at the University of Pennsylvania's Department of Economics.

"Individuals with high levels of wealth or who tend to support political candidates — people who may have greater incentive to seek occasions for social networking — place a much higher value on the private benefits associated with their giving levels," said Sieg. "These individuals are key to the sustainability of nonprofit organizations, so it is more important than ever to appeal to what motivates such donors to give."

Sieg and University of Pittsburgh Ph.D. student Jipeng Zhang compiled an extensive dataset from publicly available donor lists for 10 large Pittsburgh nonprofits: Pittsburgh Ballet Theater, Children's Museum, City Theater, Pittsburgh Opera, Phipps Conservatory, Pittsburgh Public Theater, Pittsburgh Symphony, Western Pennsylvania Conservancy, Pittsburgh Zoo and the PPG Aquarium. To analyze the data, they used a method new to philanthropy research, classifying individual tiers of giving as "prices" associated with different bundles of benefits.

Their modeling approach — "repeated discrete choice with multiple choice occasions" — is flexible, says Sieg, and has many other potential applications where consumers demand multiple units of different products. Estimating these models is computationally intensive. Initially, the researchers ran their model many times using 300 BigBen processors. For subsequent revisions, they relied on Pople, PSC's SGI Altix shared-memory system. "This would have taken forever with a high-end desktop system," says Sieg, "but the computational burden is easily feasible with the current generation of supercomputers available at the PSC."

They found that charities relying heavily on private benefits to attract donors would see lower giving if these benefits were eliminated. In addition, the nearly eight-percent of individuals who supported multiple charities would decline significantly if there were no high-value private benefits.

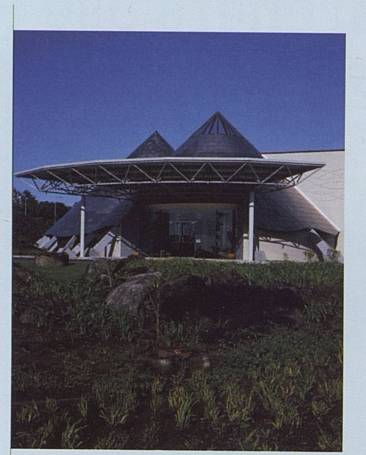
VISUALIZING STORMS

PSC scientific visualization shows storms in 3D for NSF workshop at Imiloa Planetarium

To create a visualization of a hurricane brewing its swirl of energy so that it can be projected in 3D stereo in the spectacular Imiloa Astronomy Center at the University of Hawaii in Hilo, Hawaii — that's the task that PSC visualization specialist Greg Foss undertook in 2009. The image shown here, from that visualization, is from a simulation of Hurricane Ike, the third major hurricane of the 2008 Atlantic hurricane season, and the costliest hurricane ever to make landfall in the United States.

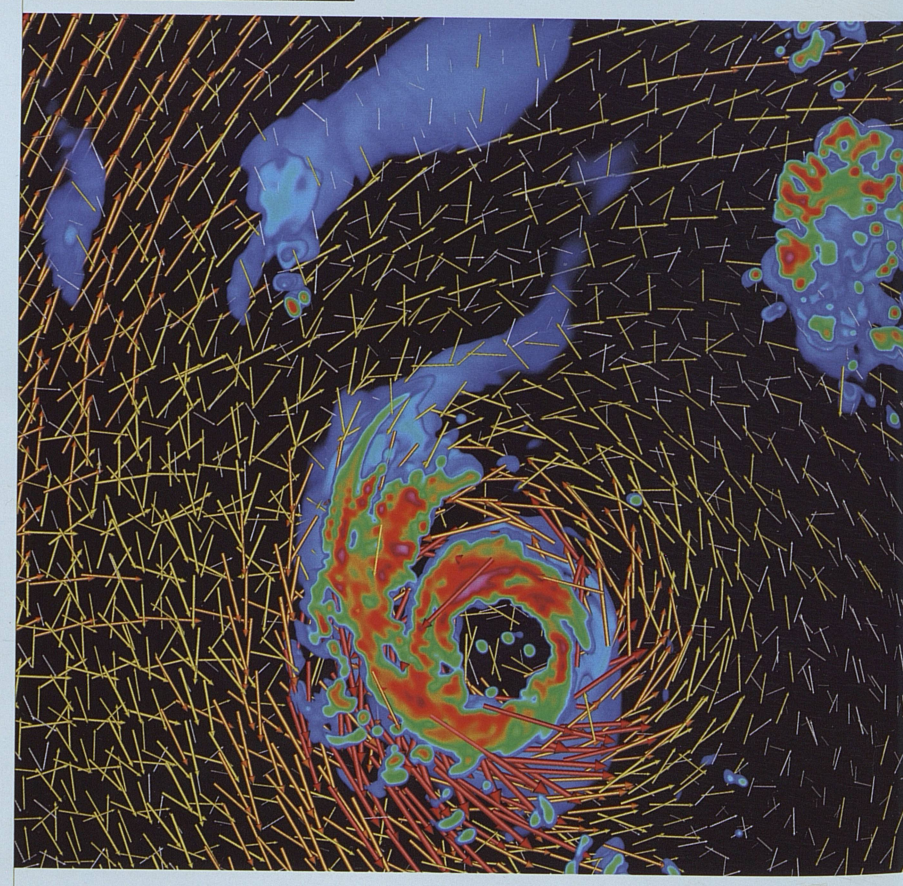
Foss created this and other animations for a November 2009 National Science Foundation workshop at the Imiloa Planetarium titled "Data Visualization: Taking the Presentation of Methods and Results to the Next Level." The workshop highlighted an NSF EPSCoR (Experimental Program to Stimulate Competitive Research) award to the University of Hawaii system to develop advanced data manipulation and visualization in Hawaii. The world-class 120-seat Imiloa Planetarium, with a 52-foot diameter dome, was the first in the world to have 3D stereo capability.

Simulations on PSC's BigBen and Pople systems, using the Advanced Regional Prediction System software from the Center for Analysis and Prediction of Storms at the University of Oklahoma, Norman, produced the data. The Hurricane Ike simulation ran on 64 processors of PSC's shared-memory SGI Altix system Pople. Foss used the visualization software VisIt to create the animation, which represents a domain 1600 kilometers on each side and 16.8 kilometers vertically. The arrows represent wind direction and velocity (increasing in magnitude with color from yellow to red). The bright coloring of the clouds corresponds to radar reflectivity, which shows the amount of water vapor in the atmosphere (increasing from blue to green, yellow and red).

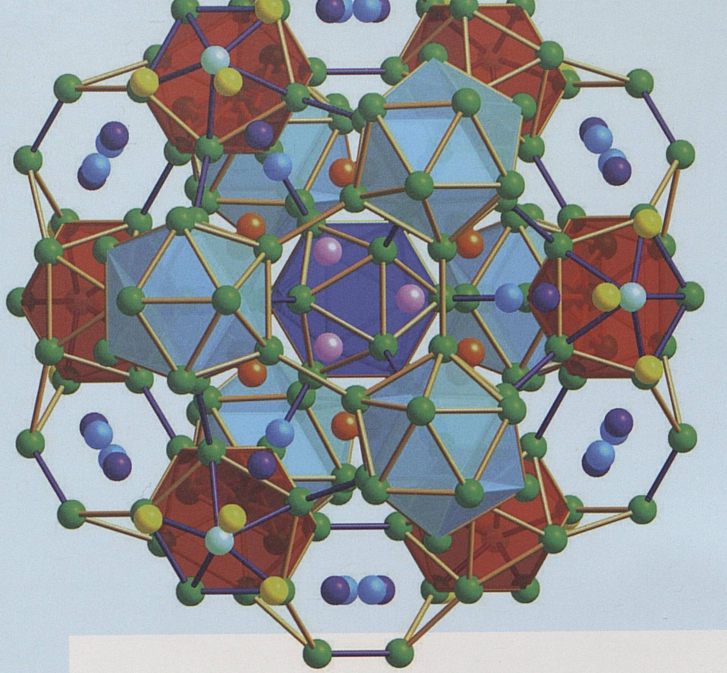


Imiloa Planetarium, University of Hawaii, Hilo

Photo Credit: Macario



Hurricane Ike



QUANTUM LEAP TO VIENNA

PSC connection leads Carnegie Mellon student to Vienna to work on quantum physics software

MAX HUTCHINSON
PSC's director of Scientific Applications and User Support, Sergiu Sanielevici, notes that Hutchinson is an "early explorer" of PSC's new SGI UV system, Blacklight (p. 4). "Max has excellent knowledge of computing systems," says Sanielevici, "and how to program them to productively generate new scientific knowledge."

To work on quantum physics software — even if it involves an expenses paid sojourn in Vienna — isn't the usual undergrad spring break. For Carnegie Mellon physics student Max Hutchinson, however, it's a natural extension of his involvement, since high school, with PSC and TeraGrid educational outreach. From a supercomputing award at the Pittsburgh Science Fair to awards at the 2008 and 2009 TeraGrid conference student competitions, to working as an intern in PSC's Strategic Applications Group, Hutchinson has made the leap from student competitions to real world scientist. This year he began working in the research group of Carnegie Mellon physicist Michael Widom, which led him to Vienna on spring break to work on widely used quantum physics software called VASP (Vienna Ab-initio Simulation Package).

"His PSC experience taught him a great deal about machine architecture and the link with efficient algorithms," says Widom of Hutchinson, "and through him this experience is improving my own ability to solve interesting problems." Some of Hutchinson's work involved GPU (graphics processing unit) systems. The TeraGrid includes several GPU systems, which have processors specialized for graphics rendering. Because of this expertise, Hutchinson was invited to the University of Vienna over spring break, where he worked on porting the VASP software to test on a GPU system. He will apply this method to study the complex crystal structure of the element boron, the stable form of which (called "beta" boron, represented in the graphic shown here) Widom and collaborators confirmed using VASP.

In addition to the study of boron, Widom's work with VASP has included new understanding of the liquid-liquid transition in supercooled silicon, reported in *Physical Review Letters* (February 2009). He has relied on PSC's Pople system for many of his VASP projects, since its shared-memory architecture allows these quantum calculations to run more efficiently than they would otherwise. In this work, Widom has often collaborated closely with PSC scientist Yang Wang.

PITTSBURGH SUPERCOMPUTING CENTER

The Pittsburgh Supercomputing Center is a joint effort of Carnegie Mellon University and the University of Pittsburgh together with Westinghouse Electric Company. It was established in 1986 and is supported by several federal agencies, the Commonwealth of Pennsylvania and private industry.

PSC GRATEFULLY ACKNOWLEDGES SIGNIFICANT SUPPORT FROM THE FOLLOWING:

The Commonwealth of Pennsylvania

The National Science Foundation

The National Institutes of Health

The National Energy Technology Laboratory

The National Archives and Records Administration

The U. S. Department of Defense

The U. S. Department of Energy

Chesio Communications

Cisco Systems, Inc.

Cray Inc.

DataDirect Networks

DSF Charitable Foundation

Microsoft Corporation

Silicon Graphics, Inc.

The Buhl Foundation

Bill and Melinda Gates Foundation

EDITOR/WRITER: Michael Schneider, PSC

CONTRIBUTING WRITER: Tim Palucka

GRAPHICS RESEARCH, PHOTOGRAPHY DIRECTION & PROJECT COORDINATION: Shandra Williams, PSC

TRANSCRIBING: Ryan Omecene, PSC

PHOTOGRAPHY: Tim Kaulen, Photography & Graphic Services at Mellon Institute.

GRAPHICS: Thanks to the researchers and to Shawn Brown, Greg Foss, Joel Stiles, Art Wetzel and Troy Wymore of PSC for graphics.

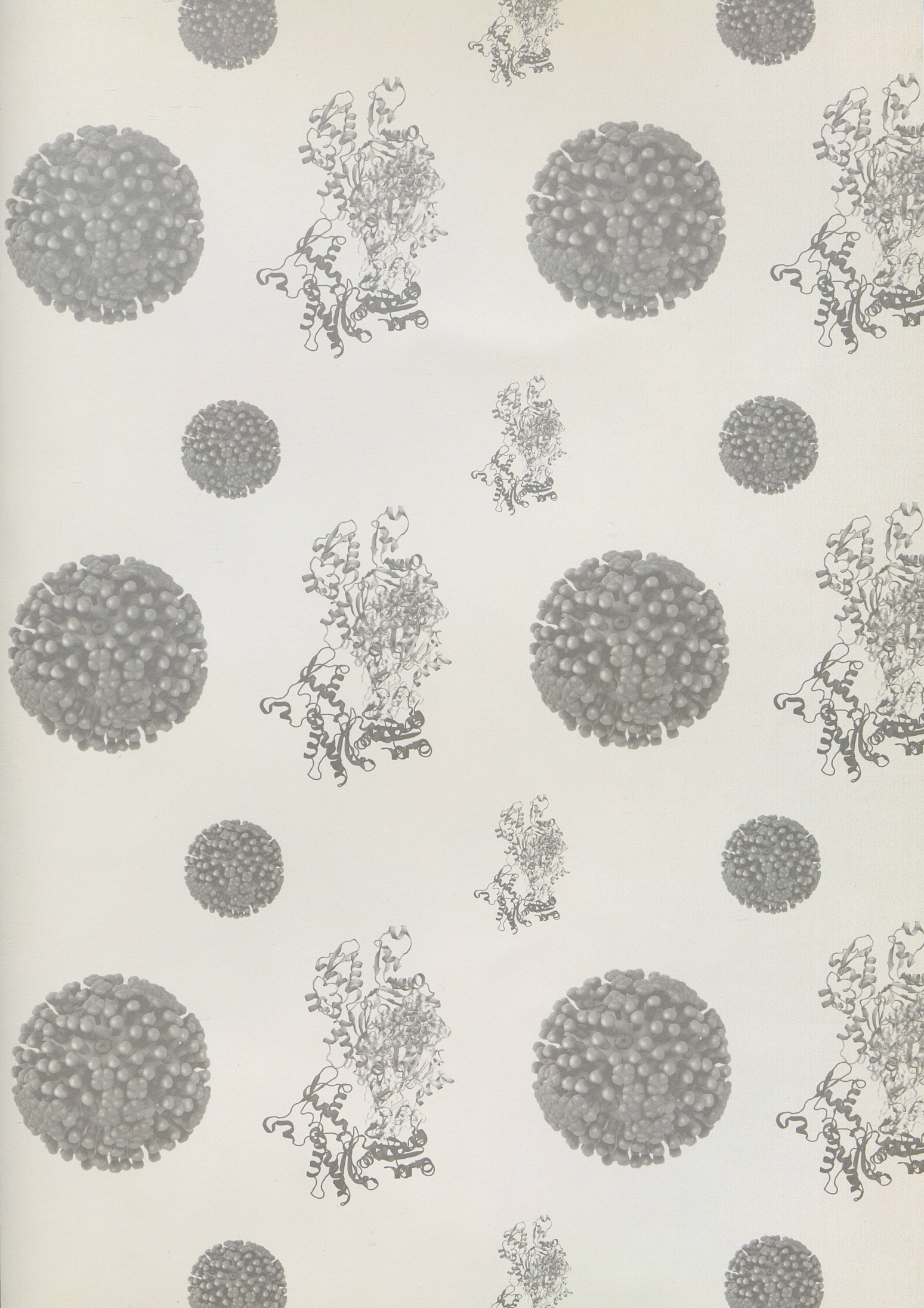
COVER GRAPHIC: Vectors of Hurricane Ike. Rendered by Greg Foss, PSC visualization specialist, from simulations of Hurricane Ike (see p. 46).

DESIGN: Wall-to-Wall Studios, Inc.

PRINTING: Broudy Printing

MIX
Paper from responsible sources
FSC® C013022
FSC www.fsc.org

Printed on Sappi Flo Paper, a premium sheet with 10 percent post-consumer waste fiber, with vegetable-based inks.



PITTSBURGH SUPERCOMPUTING CENTER
300 S. CRAIG STREET
PITTSBURGH, PENNSYLVANIA 15213